

Technical University of Košice



**Faculty of Electrical Engineering
and Informatics**

Electrical Engineering and Informatics

14

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

ISBN 978-80-553-4407-2

2023

Electrical Engineering and Informatics 14
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: July 2023 Language: English, Slovak
Printing: 50 pieces CD Pages: 486

Editorial board chairman: Karol Kyslan

Proceedings reviewers:

Samuel Andrejčík	Kristína Machová
František Babič	Branislav Madoš
Ján Bačík	Ján Magyar
Anna Biceková	Stanislav Marchevský
Samuel Bucko	Miroslav Michalko
Eva Chovancová	Jozef Onufer
Dávid Cymbalák	Jakub Oravec
Miloš Drutarovský	Luboš Ovseník
Juraj Ďurišin	Lukáš Pancurák
František Ďurovský	Ján Paralič
Jozef Dziak	Marek Pástor
Jaroslav Džmura	Daniela Perduková
Peter Fecilak	Jaroslav Petráš
Pavol Fedor	Emília Pietriková
Ján Genčí	Eugen Šlapak
Peter Girovský	Branislav Sobota
Peter Gnip	Miroslav Sokol
Renát Haluška	Zuzana Sokolová
Daniel Hládek	Ján Staš
Jozef Juhár	William Steingartner
Ondrej Kainz	Roman Vápeník
Marianna Koctúrová	Ján Vaščák
Štefan Korečko	Marcel Vološin
Karol Kyslan	Ján Zbojovský
Milan Lacko	Jaroslava Žilková
Marian Mach	

Editors: Karol Kyslan
Emília Pietriková
Lukáš Pancurák

Contents

Dávid Cymbalák, Roman Vápeník, Timotej Sobota, František Jakab <i>Optické rozpoznávanie textu na báze Tesseract vo video sekvenciách v rozsiahlych video archívoch</i>	7
Dávid Cymbalák, Roman Vápeník, Nikola Móricaová, František Jakab <i>Multimedia Archive Services with Effective Categorization and Intuitive Search</i>	12
Dávid Cymbalák, Miroslav Murin, Roman Vápeník, František Jakab <i>Využitie STT na extrakciu textu vo video sekvenciách v rozsiahlych multimediálnych archívoch</i>	17
Roman Vápeník, Dávid Cymbalák, Nikola Móricaová, František Jakab <i>Video archív s efektívnym odporúčaním obsahu a kategorizácie na báze tagov a tituliek</i>	22
Kristína Machová, Mário Horváth <i>Aktívne učenie v analýze sentimentu</i>	26
Ján Skybjak, Marián Mach <i>Numerická optimalizácia inšpirovaná vlčími svorkami</i>	30
Dmytro Lahunov, Marián Mach <i>Cielové kontradiktórne útoky na klasifikačné hlboké modely</i>	35
Kristína Machová, Michal Janok <i>Analýza webových trendov</i>	40
Samuel Titko, Ján Vaščák <i>Scolopendra: Softvérové rozšírenie hexapodnej robotickej platformy</i>	45
Matúš Smolko, Ján Vaščák <i>Návrh softvérového rámca pre komunikáciu V2X v oblasti robotických rojov</i>	51
Luboš Ovseník, Jakub Oravec, Samuel Andrejčík, Martin Schlachta <i>Analýza mikropásikových anténnych polí pre hybridné FSO/RF systémy</i>	57
Luboš Ovseník, Jakub Oravec, Samuel Andrejčík, Ján Bujňák <i>Návrh modelu systému „Radio over Fiber”</i>	62
Martin Nguyen <i>Detecting Melanoma in Images Using Neural Networks: A Hybrid Application for Early Diagnosis</i>	68
Ján Vaščák <i>Návrh navigačného systému na báze neuro-fuzzy prístupu s využitím cloud computing-u</i>	73
Luboš Ovseník, Jakub Oravec, Samuel Andrejčík, Ján Giňovský <i>Návrh mobilnej aplikácie pre vizualizáciu poveternostných podmienok v monitorovanom priestore</i>	79
Luboš Ovseník, Jakub Oravec, Samuel Andrejčík, Richard Pribula <i>Návrh web rozhrania pre vizualizáciu poveternostných podmienok v monitorovanom priestore</i>	85
Petra Jankurová, Ján Genči <i>Identifikácia kandidátov na neologizmy na báze slovenského internetového doménového priestoru</i>	89
Oliver Lohaj, Ján Paralič, Ján Jevčák <i>Meranie aspektov použiteľnosti systémov pre podporu rozhodovania lekárov</i>	96
Lenka Dobranská, Tomáš Adam, František Babič	

<i>Porovnanie klasifikačných modelov z pohľadu interpretovateľnosti</i>	102
Jozef Doci, Tomáš Adam, František Babič <i>FEI's digital assistant in augmented reality</i>	107
Iveta Gladišová <i>Nekonvenčná metóda bezstratovej kompresie statických obrazov</i>	112
Branislav Sobota, Miriama Mattová, Soňa Matulová, Štefan Korečko <i>Virtual reality environment as a basis for humanoid robot training and testing system</i>	118
Branislav Sobota, Miriama Mattová, Sebastián Dimun, Štefan Korečko <i>Virtuálna škola v prostredí kolaboratívnej virtuálnej reality</i>	124
Kristína Machová, Miroslav Plavec <i>Rozpoznávanie hoaxov modelovaním slovníka ich autorov</i>	129
Anna Biceková, Dávid Ujhelyi <i>Metódy dátovej analytiky pre predikciu platov v oblasti športu</i>	134
Jozef Humeník, Jaroslav Džmura <i>Vplyv elektromobility na životné prostredie</i>	140
Anna Biceková, Alena Čubiňáková <i>Návrh systému na podporu výučby</i>	145
Jakub Oravec, Ľuboš Ovseník, Samuel Andrejčík <i>Možnosti použitia postupností OEIS pri štatistickom testovaní obrazových šifrovacích algoritmov</i>	149
Jakub Oravec, Ľuboš Ovseník, Samuel Andrejčík <i>Hardvérový návrh a realizácia laboratória sietovej steganografie</i>	153
Stanislav Levendovský, Michal Vrábel <i>Detection of Events in Data from DIMS Experiment</i>	158
Tomáš Kmec <i>Cold atmospheric plasma treatment of substrate for heating resistive film</i>	164
Samuel Oklamčák, Tomáš Tkáčik, Ján Jadlovský <i>Inovácia materiálov pre predmety priemyselnej informatiky so zameraním na manažérske informačné systémy – výukový portál</i>	170
Ján Vaščák <i>Reinforcement learning a radiálne siete pre potreby riadenia pohybu</i>	176
Kristián Haborák, Dušan Medved' <i>Vplyv prosumerov na distribučnú sieť</i>	182
Martin Kriššák, Dušan Medved' <i>Budúcnosť vodíkových technológií</i>	188
Lívia Kováčová, Branislav Madoš <i>Transforming a Gantt Chart Into an Equivalent Hierarchical Data Structure</i>	193
Ján Vaščák <i>Návrh simulátora fuzzy kognitívnych máp pre potreby navigácie vozidiel</i>	199
Lubomír Rusnák, Marián Mach <i>SAT kódovanie kardinalitných ohraničení pomocou binárneho stromu</i>	204

Filip Juríni, Dušan Medved' <i>Návrh výroby palíva pre palivový článok</i>	210
Katarína Chavková, Branislav Madoš <i>Information System for Voting Support Using Blockchain Technology</i>	215
Roland Jansky, Anton Baláž, Branislav Madoš <i>Vzdelávanie prostredníctvom hier</i>	221
Jozef Majkut, Anton Baláž, Emília Pietriková <i>Gamifikácia v oblasti bezpečnosti IKT</i>	229
Matej Vysokai, Anton Baláž <i>Vzdialená správa hardvéru IP KVM</i>	237
Vladimír Szomosi, Dušan Medved' <i>Možnosti využitia elektrolyzéra pre elektroenergetické účely</i>	242
Miroslav Andraš, Dušan Medved' <i>Porovnanie AC a DC ostrovnej prevádzky nn siete</i>	248
Martin Sedlačko, Branislav Madoš <i>Design and Implementation of a CASE Tool for Reverse Engineering</i>	253
Marek Fedor, Daniela Perduková <i>Riadenie pohonu s asynchrónnym motorom na báze inverzného fuzzy modelu</i>	259
Martin Kopecký, Tomáš Tkáčik, Ján Jadlovský <i>Modifikácia informačného systému DARMA pre experiment ALICE CERN</i>	267
Marek Belej, Branislav Madoš, Anton Baláž <i>Návrh a programová implementácia sociálnej siete</i>	272
Dávid Bodnár, František Ďurovský <i>Modelovanie Superkapacitorov pre Aplikácie Hybridných Zásobníkov Energie</i>	278
Peter Girovský, Daniel Kormanský <i>Návrh metódy pre meranie zdravia stromov</i>	282
Peter Girovský, Jaroslava Žilková, Peter Nemergut <i>Modulárne robotické rameno</i>	288
Peter Girovský, Jaroslava Žilková, Milan Jano <i>Robotické rameno pre triedenie objektov</i>	293
Pavol Fedor, Peter Radváni, Daniela Perduková <i>Model turbíny malej vodnej elektrárne s fuzzy modelom účinnosti</i>	298
Pavol Smoleň, František Ďurovský <i>Fyzikálny model kontinuálnej linky</i>	305
Ján Báčik, Pavol Fedor, Daniela Perduková <i>Vývoj výskumnej hardvérovej platformy na báze multikoptéry</i>	314
Maroš Baumgartner, Ján Papaj <i>Využitie Technológie Blockchain v Procese Smerovania v Mobilných Viacpreskokových Sietach</i> .	317
Maroš Baumgartner, Ján Papaj, Matúš Sabol <i>Využitie Blockchainu v procese smerovania vo viacpreskokovej sieti</i>	323

Maroš Baumgartner, Ján Papaj, Peter Sárosi <i>Smerovanie vo viacpreskokovej D2D sieti a Blockchain</i>	329
Kristína Machová, Alexandra Pejková <i>Strojové učenie v rozpoznávaní rôznych foriem antisociálneho chovania</i>	335
Šimon Gans, Ján Molnár <i>Compensation of the waveform distortion of transformers to match the desired output voltage</i> ..	340
Natalia Kurkina, Ján Papaj <i>Analýza výkonnosti smerovacieho protokolu AODV a jeho modifikácií pre MANET siete</i>	346
Kristián Šimon, Natalia Kurkina, Ján Papaj <i>Analýza smerovacích protokolov na báze strojového učenia pre MANET siete</i>	351
Daniel Ďurča, Natalia Kurkina, Ján Papaj <i>Tvorba datasetov pre strojové učenie v MANET</i>	356
Peter Girovský, Patrik Molnár <i>Návrh konštrukčnej a elektrickej časti robotického ramena</i>	362
Silvester Strigáč, Tomáš Tkáčik, Ján Jadlovský <i>Aplikácia mobilného robota MiroSot v úlohe vyhýbania sa prekážkam</i>	367
Marek Pástor, Jaroslava Žilková, Juraj Jevčák <i>Možnosti simulácie strát výkonových polovodičových súčiastok v prostredí MATLAB/Simulink</i> .	373
Jaroslava Žilková, Peter Girovský <i>Vizualizácia priemyselného robota</i>	379
Renát Haluška, Erik Švenk, Matúš Pleva <i>Rozpoznávanie Používateľov v Mobilnej Aplikácii Pomocou Snímačov Odtlačkov Prstov</i>	384
Renát Haluška, Dominik Molčanyi, Matúš Pleva <i>Mobilná Aplikácia pre Biometrické Bezpečnostné Systémy</i>	389
Peter Janda, Renát Haluška, Stanislav Marchevský, Slavomír Pilár <i>Bunkové siete na báze bezpilotných lietajúcich prostriedkov</i>	394
Zuzana Sokolová, Maroš Harahus, Ján Staš, Jozef Juhár <i>Tvorba korpusu textov pre úlohy detekcie nenávistných prejavov, ofenzívneho jazyka a analýzy sentimentu</i>	399
Marián Mach <i>Explorácia v učení posilňovaním</i>	403
Renát Haluška, Peter Janda, Stanislav Marchevský, Slavomír Pilár <i>Detekcia Požiaru s Využitím Bunkových Sietí</i>	408
Maroš Harahus, Zuzana Sokolová, Daniel Hládek, Matúš Pleva <i>Anotačný nástroj korpusu textov pre úlohy opráv gramatických chýb v slovenskom jazyku</i>	414
Jakub Dzuřovčín, Ján Magyar <i>Multi-Task Learning for Personalization</i>	418
Daniel Hreha, Ján Magyar <i>Plánovanie trasy drona v inteligentnom priestore</i>	423
Marek Tóth, Ján Magyar	

<i>Lokalizácia drona v inteligentnom priestore</i>	428
Kristína Machová, Branislav Kovalčín <i>Modelovanie Šírenia Nedôveryhodných Správ o Pandémii CoViD-19 na Sociálnych Sieťach</i>	432
Adrián Marcinek, Marek Pástor <i>Viacportový menič založený na MAB</i>	438
Daniel Marcin, Milan Lacko <i>Batériové systémy elektromobilov a predstavenie modulárneho systému</i>	444
Lukáš Pancurák, Karol Kyslan <i>Synchronizácia trojmotorového pohonu</i>	449
Jaroslava Žilková, Ján Kaňuch <i>Pretlakový plnič PET fľaš</i>	455
Daniel Gordan, Tomáš Basarik, Viktor Šlapák <i>Meranie indukčností vinutí elektrických motorov s použitím samovyvažovacieho mostíka a mikrokontroléra</i>	460
Ján Kaňuch, Peter Girovský <i>Troj-statorový diskový synchronný motor s permanentnými magnetmi</i>	464
Jaroslava Žilková, Marek Pástor <i>Priame momentové riadenie AM</i>	473
Daniel Kolibár, Filip Gurbál <i>Hodnotenie dôveryhodnosti dát pohybových aktivít študentov</i>	478
Author's Index	484

Optické rozpoznávanie textu na báze Tesseract vo video sekvenciách v rozsiahlych video archívoch

¹Dávid CYMBALÁK, ²Roman VÁPENÍK, ³Timotej SOBOTA, ⁴František JAKAB

^{1,2} Spoločné pracovisko UVP Technicom a Centra vedecko-technických informácií SR, Technická univerzita v Košiciach, Slovenská republika

^{3,4} Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹david.cymbalak@uvptechnicom.sk, ²roman.vapenik@cvtisr.sk,
³timotej.sobota@student.tuke.sk, ⁴frantisek.jakab@tuke.sk,

Abstrakt — Článok skúma využitie Tesseract OCR na rozpoznávanie textu v rozsiahlych multimediálnych archívoch. Vhodnou bázou pre jeho adaptáciu bol rozsiahly archív vizuálnych záznamov Národnej Teleprezentačnej infraštruktúry obsahujúci tisíce videí prednáškového charakteru. Výsledkom výskumu bol návrh a implementácia riešenia s prepojením podporných knižníc a nástrojov pre spracovanie videa, agregáciu snímok a multikrokovú filtráciu. Vytvorené riešenie s použitím Tesseract OCR sa ukázalo ako efektívny a presný nástroj na rozpoznávanie s vysokým výkonom pri spracovaní veľkého množstva vstupných dát v krátkom čase, čo ho robí ideálnym nástrojom pre rozpoznávanie textu v rozsiahlych multimediálnych archívoch.

Kľúčové slová — video OCR, multimediálny archív, Tesseract

I. ÚVOD

V dnešnej digitálnej dobe je spracovanie väčšieho množstva multimediálnych dát a ich efektívna archivácia pre dotknuté organizácie veľkou výzvou. Tieto archívy môžu obsahovať videá, fotografie, dokumenty, zvukové nahrávky a ďalšie súbory, ktoré je potrebné efektívne spracovať a kategorizovať. Článok sa venuje najmä jednému z kľúčových problémov - extrakcii textu pri spracovaní videí z týchto rozsiahlych multimediálnych archívov. Manuálne prepisovanie textu z videí je časovo náročné a náchylné na chyby. V súčasnosti preto existuje široká škála nástrojov na rozpoznávanie textu v obraze a videách, avšak väčšina z nich má obmedzenú presnosť a nie sú plne adaptované pre spracovanie veľkého množstva videí. Úlohou našich výskumných aktivít je práve zameranie sa na adaptáciu OCR nástrojov pre rozsiahle databázy vizuálnych video záznamov a otestovanie úspešnosti rozpoznania textu pri nasadení.

II. ANALÝZA TECHNOLOGIÍ OPTICKÉHO ROZPOZNÁVANIA TEXTU PRE VYUŽITIE VO VIDEU

Nástroje a technológie na rozpoznávanie textu v obraze a videách môžu mať rôzne vlastnosti a presnosť v závislosti na použitých algoritmoch, trénoch a kvalite vstupných dát [1]. Výber správnej technológie pre rozpoznávanie textu závisí na viacerých faktoroch, vrátane požiadaviek na presnosť, rýchlosť spracovania, podporu jazykov a ďalšie funkcie. Každá z technológií má svoje výhody a nevýhody a správna voľba závisí od konkrétnych požiadaviek na rozpoznávanie textu. V súčasnosti sú na trhu v tomto segmente najpoužívanejšie tieto technológie na rozpoznávanie textu [2]:

Google Cloud Vision je služba od spoločnosti Google, ktorá umožňuje rozpoznávanie textu a iných objektov v obrázkoch a videách pomocou strojového učenia. Táto služba umožňuje detekciu a rozpoznávanie textu v rôznych jazykoch a umožňuje aj možnosť trénoch modelov na základe vlastných dát.

Amazon Rekognition je služba od spoločnosti Amazon, ktorá okrem rozšírenej podpory jazykov pri rozpoznávaní textu, umožňuje aj detekciu a rozpoznávanie iných objektov v obrázkoch a videách.

OpenCV (Open Source Computer Vision Library) je open-source knižnica pre spracovanie obrazu, ktorá umožňuje rozpoznávanie textu vo videách pomocou širokej škály algoritmov

a technológií. Táto knižnica okrem primárnych nástrojov spracovania obrazu umožňuje aj vlastné tréovanie modelov pre rozpoznávanie textu.

Textract je nástroj od spoločnosti Amazon, ktorý umožňuje extrakciu textu z rôznych typov súborov, vrátane obrázkov a videí. Podobne ako Amazon Rekognition má tento nástroj silnú podporu rozpoznávania textu v rôznych jazykoch.

Tesseract OCR je open-source rozpoznávač textu v obraze a dokumentoch s dlhou históriou, ktorý je vyvíjaný už od roku 1985. Tesseract OCR je považovaný za jednu z najpresnejších technológií na rozpoznávanie textu v obrazoch a videách a má veľké množstvo používateľov aj v oblasti spracovania multimediálneho obsahu.

OCRopus je open-source knižnica pre rozpoznávanie textu v obraze a dokumentoch, ktorá sa vyznačuje vysokou presnosťou a robustnosťou. Táto knižnica je založená na moderných technikách strojového učenia a umožňuje aj tréovanie modelov pre rozpoznávanie textu v rôznych jazykoch.

Kraken je ďalšia open-source knižnica s obdobnými vlastnosťami ako OCRopus, no vyniká vysokou presnosťou a rýchlosťou spracovania.

Abbyy FineReader je komerčný nástroj pre rozpoznávanie textu v obrazoch a dokumentoch, ktorý sa vyznačuje množstvom funkcií na spracovanie a úpravu textu, ako aj možnosť vlastného tréovania modelov.

A. Nástroj Tesseract a jeho adaptácia

Výber adaptácie nástroja Tesseract OCR ako riešenia na problém rozpoznávania textu v rozsiahlych multimediálnych archívoch má množstvo výhod primárne reflektujúcich vysokú presnosť a rýchlosť spracovania. Tesseract OCR je taktiež navrhnutý tak, aby bol ľahko rozšíriteľný a prispôsobiteľný pre rôzne jazyky a typy dokumentov. Používa moderné techniky strojového učenia a algoritmy na rozpoznávanie textu, ktoré znižujú počet chýb pri rozpoznávaní textu v obraze. Možnosti konfigurácie pre prispôbenie nástroja konkrétnym potrebám umožňujú nastaviť parametre jazykového základu, zvoliť optimálnu konfiguráciu pre rôzne typy vstupných dokumentov a určiť, ako má byť spracovaný rozpoznávaný text a generované výstupné súbory. V závislosti na veľkosti a zložitosti multimediálneho archívu môže byť potrebné použiť aj iné doplnkové nástroje a technológie pre pred-spracovanie, filtráciu obrazu alebo agregáciu snímok videí. Nástroj Tesseract je vhodné nasadiť s pred-spracovaním obrazu z videa v dvoch štandardných rovinách: s použitím nástroja FFMPEG alebo s použitím knižníc OpenCV.

B. Tesseract a vyhľadávanie línií v obraze

Algoritmus na vyhľadávanie línií ako súčasť nástroja Tesseract [3] bol navrhnutý tak, aby vedel rozpoznáť aj zošikmenú stranu bez toho, aby ju musel narovnávať, čím zabraňuje zníženiu kvality obrazu. Kľúčovou časťou procesu je filtrovanie spojených znakov (blob) a tvorba línií za predpokladu, že analýza rozloženia strany zhruba poskytla textové oblasti s približne rovnakou veľkosťou textu. Jednoduchý percentilový filter výšky odstraňuje iniciály a zvislé dotýkajúce sa znaky. Následne medián výšky aproximuje veľkosť textu v regióne, takže je možné bezpečne odstrániť spojené znaky a teda eliminovať len interpunkčné znamienka a šum [4].

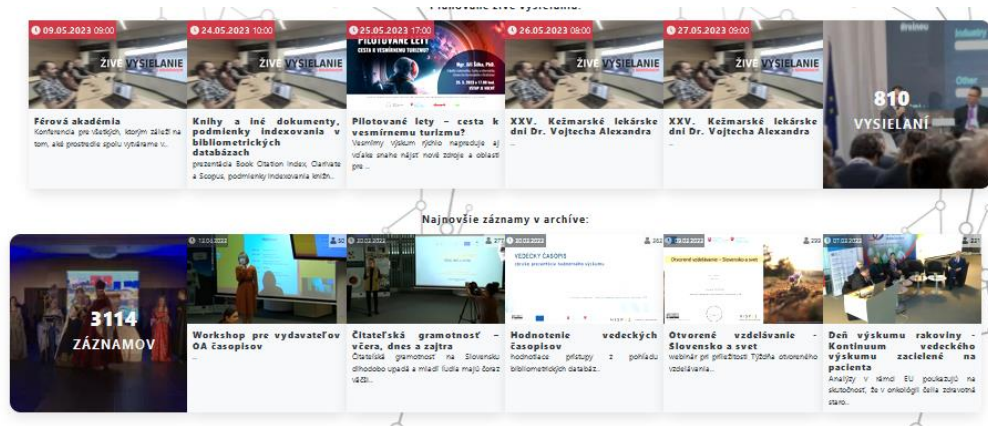
Filtrované znaky majú väčšiu pravdepodobnosť, že sa hodia do modelu neprekrývajúcich sa paralelných, ale sklonených línií. Usporiadanie a spracovanie znakov podľa x-ovej súradnice umožňuje priradenie znakov jedinečnej textovej línií, pričom sleduje sklon strany so zníženým rizikom priradenia k nesprávnej textovej línií v prítomnosti šikminy. Keď sú filtrované znaky priradené k líniám, na odhad spodnej hrany textu sa používa najmenší medián štvorcov [5] a filtrované znaky sa vkladajú späť do príslušných línií. Posledným krokom v procese vytvárania línií je zlúčenie znakov, ktoré sa aspoň z polovice prekrývajú vodorovne, následne umiestnenie diakritických znamienok spolu s vhodnou základňou a správne asociovanie častí niektorých rozbitých alebo poškodených znakov [4].

III. NÁVRH A IMPLEMENTÁCIA RIEŠENIA PRE ROZPOZNÁVANIE V ROZSIAHLOM ARCHÍVE VIZUÁLNYCH ZÁZNAMOV

A. Určenie vhodnej bázy pre nasadenie riešenia optického rozpoznávania textu vo videu

Cieľom je teda prispôbiť riešenie optického rozpoznávania textu pre obrovskú základňu vizuálnych záznamov. V tomto prípade je možné riešenia adaptovať na rozsiahlych video archívoch Národnej teleprezentačnej infraštruktúry (NTI) pre podporu výskumu, vývoja a transferu technológií na Slovensku [6]. NTI bola pôvodne vytvorená v rámci Centra vedecko-technických informácií SR v roku 2015. Počas existencie NTI [7] boli realizované video

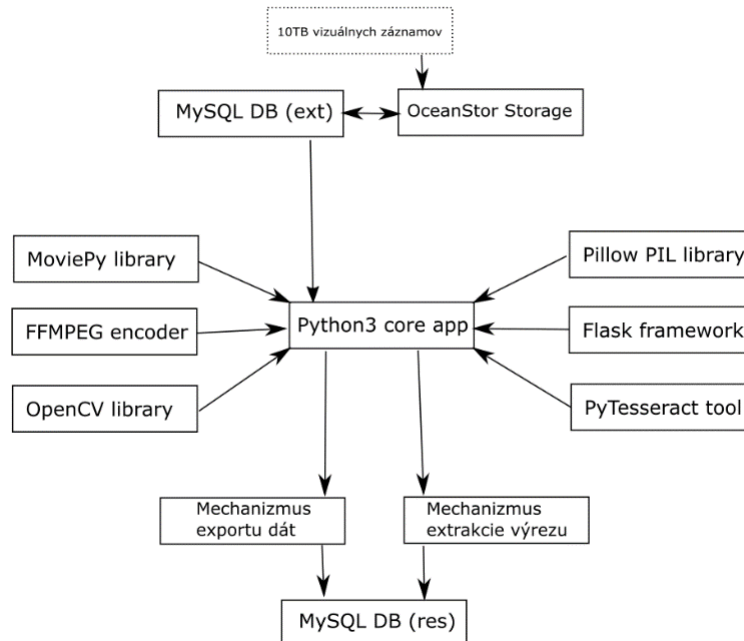
konferencie v objeme stoviek tisíc spojení, tisíce manažovaných živých vysielaní a bolo vytvorené a archivované tisíce záznamov prístupných pre stovky tisíc divákov. NTI multimediálny archív je tvorený katalógom záznamov (Obr. 1) s podsekciami s konkrétnym vizuálnym záznamom obsahujúcim súvisiace záznamy a informácie s integrovaným webovým prehrávačom VOD s efektívnou distribúciou zdroja videa vo formáte MP4 k príjemcovi. Spracovanie a distribúcia vizuálnych záznamov boli navrhnuté pomocou inštancie VOD streamovacieho servera s prepojením na mySQL databázu.



Obr. 1 Multimediálny archív NTI obsahujúci video záznamy vedeckých podujatí

B. Implementácia prepojenia riešenia s existujúcim archívom a nasadenie Tesseract s podpornými knižnicami a viackrokovou filtráciou

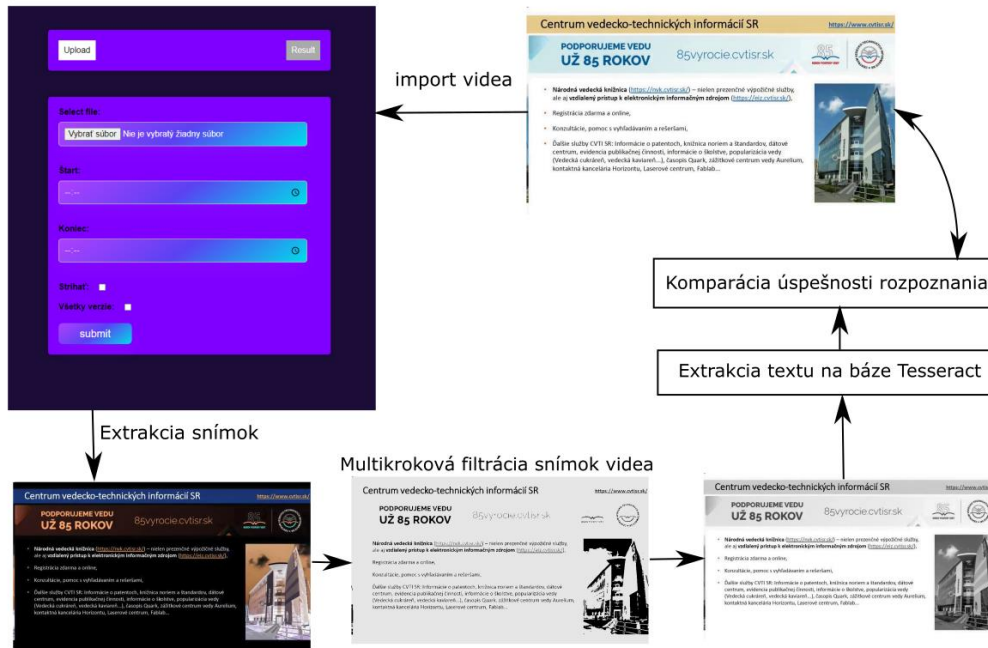
Navrhnuté riešenie pozostáva s vytvorenia aplikácie v jazyku Python s využitím viacerých externých knižníc a realizácii prepojenia na existujúce databázy rozsiahleho archívu a mechanizmy extrakcie a exportu dát (Obr. 2).



Obr. 2 Prepojenie prvkov riešenia pre rozpoznávanie textu vo videu z rozsiahlej databázy existujúceho archívu vizuálnych záznamov

Databáza existujúceho rozsiahleho archívu obsahuje umiestnenie zdrojových MP4 videosúborov umiestnených na OceanStor úložisku. Knižnica MoviePy sa využíva core funkciou `cut_video`, kde sa extrahuje relevantná sekvencia videa určená pre spracovanie. OpenCV v tomto návrhu slúži na čítanie videosnímkov z videosúboru alebo kamerového prenosu a realizáciu multikrokovej filtrácie agregovaných snímok videa (Obr. 3). V tomto prípade je parametrom cesta k videu v prepojenej existujúcej databáze. V následnom kroku sa na automaticky vybraných agregovaných snímkach s použitím OpenCV `cv2` vykoná multikroková filtrácia videa začínajúca realizáciou operácie NOT na každom pixeli obrazu. Ďalším krokom je konverzia z RGB farebného priestoru na odtiene šedej. Tento prevod je často užitočný na zjednodušenie úloh

spracovania obrazu, ktoré vyžadujú iba informácie o jase alebo intenzite pixelov a nie informácie o ich farbe. Nasledujúca technika prahovania (thresholding) realizuje priradenie hodnôt pixelov vo vzťahu k zadanej prahovej hodnote. Pri prahovaní sa každá hodnota pixelu porovnáva s prahovou hodnotou a nastaví sa na 0, ak je menšia alebo 255, ak je to opačný prípad.



Obr. 3 Sekcia riešenia pre import videa s multikrokovou filtráciou a komparáciou úspešnosti rozpoznania textu

K sekundárnemu spôsobom extrakcie snímok z videa slúži inštancia FFMPEG enkódera. Metóda `image_to_string` z knižnice `pytesseract` prijíma vstupný obrázok ako argument a vykonáva operácie, ako je škálovanie, prahovanie a binarizácia, aby pripravila obrázok na OCR. Potom na obrázok aplikuje algoritmy OCR na rozpoznanie textu, ktorý sa v ňom nachádza a vráti reťazec obsahujúci rozpoznávaný text. Text je následne párovaný so zdrojovým vizuálnym záznamom a uložený do internej `mysql` databázy. Doplnkovou časťou riešenia je mechanizmus pre extrahovanie výrezu obrázku z videa, ktorý obsahuje relevantné informácie zistené nasadením textového rozpoznávania. Využitelnosť tohto prístupu je najmä pri extrakcii prezentačných snímok pri videu prednáškových aktivít zachytávajújúcich plátno a prezentujúceho v miestnosti. Je dôležité poznamenať, že extrahovanie výrezov obrázkov z videa môže byť náročné, ak sa obrázok, ktorý chceme získať, nachádza v rýchlo sa meniacom prostredí alebo je umiestnený v časti videa, kde sa nachádza veľa pohybu. V takýchto prípadoch môže byť potrebné použiť špeciálne techniky a nástroje, ako sú napríklad sledovanie objektov v reálnom čase, ktoré umožňujú získať presnejšie výsledky. V extrahovaní výrezov obrázkov z videa môže byť tiež využitá technológia strojového učenia, ktorá umožňuje trénovať modely, ktoré sú schopné automaticky identifikovať a extrahovať dôležité informácie, vzory a objekty z videa.

IV. ZÁVER

Experimentálna implementácia opísaná v tomto článku poukazuje na to, že Tesseract OCR je efektívnym a presným nástrojom na rozpoznávanie textu v rozsiahlych multimediálnych archívoch. Adaptácia na rozsiahlom archíve vizuálnych záznamov NTI ukazuje, že využitie Tesseract OCR s prepojením na sekundárne nástroje FFMPEG a `openCV` má vysokú presnosť aj v prípade, keď je text umiestnený na rôznych pozíciách, s rôznou veľkosťou písma alebo s rôznymi typmi písma. Testované riešenie dokáže spracovať veľké množstvo textu v krátkom čase, čo ho robí ideálnym nástrojom nie len pre účely spracovania videa vo veľkom objeme, ale aj v iných príbuzných oblastiach vyžadujúcich efektívne rozpoznanie textu v reálnom čase.

Výsledky tejto práce tiež ukazujú, že je dôležité používať viacero nástrojov a technológií s nájdením správneho prístupu v ich prepojení, čo prispieva k presnejšiemu a rýchlejšiemu spracovaniu informácií v rozsiahlych systémoch.

Budúce rozšírenie riešenia je uvažované v oblastiach ako sú rozpoznávanie tváre, rozpoznávanie emócií alebo v detekcii objektov záujmu vo videu.

POĎAKOVANIE

Táto publikácia vznikla vďaka podpore v rámci Operačného programu Integrovaná infraštruktúra pre projekt: „Výskumné centrum pre analýzu a ochranu dát - II. Etapa“, Kód ITMS2014+: 313021W479, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.

LITERATÚRA

- [1] JAIN, Pooja; TANEJA, Kavita; TANEJA, Harmunish. Which OCR toolset is good and why: A comparative study. *Kuwait Journal of Science*. 2021, roč. 48, č. 2.
- [2] HEGGHAMMER, Thomas. OCR with Tesseract, Amazon Textract, and Google Document AI: a benchmarking experiment. *Journal of Computational Social Science*, 2022, 5.1: 861-882.
- [3] SMITH, R. A simple and efficient skew detection algorithm via text row accumulation. *Proceedings of 3rd International Conference on Document Analysis and Recognition*. 1995, roč. 2, 1145–1148 vol.2.
- [4] SMITH, R. An Overview of the Tesseract OCR Engine. In: *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*. 2007, zv. 2, s. 629–633. Dostupné z doi: 10.1109/ICDAR.2007.4376991.
- [5] ROUSSEEUW, Peter J; LEROY, Annick M. *Robust regression and outlier detection*. John wiley & sons, 2005.
- [6] JAKAB, F., et al. Aggregation and Evaluation of Communication Data of Large Telepresence Infrastructure based on Computing Nodes on Critical Endpoints. In: *2022 20th International Conference on Emerging eLearning Technologies and Applications (ICETA)*. IEEE, 2022. p. 260-265.
- [7] CYMBALAK, Dávid, et al. Effective collaborative environment with live broadcasts using Slovak National telepresence infrastructure. In: *2017 15th International Conference on Emerging eLearning Technologies and Applications (ICETA)*. IEEE, 2017. p. 1-6.

Multimedia Archive Services with Effective Categorization and Intuitive Search

¹Dávid CYMBALÁK, ²Roman VÁPENÍK, ³Nikola MÓRICOVÁ, ⁴František JAKAB

^{1,2} Joint office of UVP Technicom and Slovak Centre of Scientific and Technical Information, Technical University of Košice, Slovak Republic
^{3,4} Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

¹david.cymbalak@uvptechnicom.sk, ²roman.vapenik@cvtisr.sk,
³nikola.moricova@student.tuke.sk, ⁴frantisek.jakab@tuke.sk,

Abstract — This paper discusses multimedia archive and its services. Specifically, efficient services are analyzed such as searching in multimedia, classification and categorization of multimedia content based on deep neural network. Also, there is discussed system of intelligent recommendation in archives based on YouTube like web portals. If the video in archive contains subtitles, then search methods or tagging can be also used from extracted text. Data in multimedia archives can be represented by audio, text, video, animation and so on. Considering video archives there is need to know some video methods for compression and decompression. The main goal of this paper is to acquaint the reader with available options and possibilities in developing an intelligent video archive.

Keywords — multimedia archive, neural network, deep learning, efficient information services

I. INTRODUCTION

Multimedia archive is a technology used to archive data. These data can be represented by text, audio, animation, and other information. Considering video archive, it is necessary to deal with video and the best options for delivering to the user. These options are methods for encoding, decoding, storing, encapsulating video and in the end, transferring it over the Internet. The archive is also associated with various services over its information such as storage, classification, categorization, recommendation, search, etc. With the increasing numbers of information and data in archives is necessary to move on from manual work, which can involve manually adding information by administrators to databases for further use, to more efficient automated work. Neural network and depth learning can provide this benefit. The advantage is that the neural network can learn, which can also be used in multimedia archives. Therefore, this article will deal with the mentioned topics in relation more efficient information services in multimedia archives.

II. MULTIMEDIA ARCHIVE AND ITS SERVICES

Multimedia as a combination of text, audio, graphics, video etc. are intended for information, entertainment, backup and stored in the multimedia archives, which are often presented on the Internet. Archiving is a process in which inactive information and data are securely stored for a longer time. It may happen that the data is accidentally deleted, and archive provides the data protection. Modern intranet technology provides easy management of large databases with different types of data (documents, books, pictures, video recordings, digital audio and video, streams, movies, dynamic websites, advertisements, etc.). Multimedia archive allows creation of central database with all kinds of these data and decentralized data access. Thanks to flexibility and the ability to adapt, the multimedia archive is extraordinarily advantageous for use in a wide range of organizations. Archives are easy to use, often contain search boxes, which makes it easier to work with. They also provide various services on information such as storage, retrieval, classification, categorization, recommendations and so on [1].

A. Searching in multimedia content

The main goal of searching systems is helping the user to find videos that he wants to watch and maximize his long-term satisfaction. With the increasing number of users, attention is on search, which depends on what the user wants to see and the results. In this process the user enters a query, and the tool returns results with match. Demand may be represented by one or more terms and phrases. The important role has searching methods based on neural network which consist of user's demand, query results and generation of target vector according to demand. The demand is based on a multilayer perception and methods include predicting vector representation of target expression based on the deep neural network semantic model [2].

The searcher is trying to return the relevant information on request. This requirement is usually a query, a group of words related to a topic. For example, the web search engine will find upon a request the web sites that contain one or more words in the query. This process is characterized by text-oriented searching. The tool returns links of web pages and short description provided by the site [3].

B. Classification of multimedia content

Many depth learning algorithms focus on three data domains: images, speech, and text. These data domains are used in applications to classify images, text and speech recognition. Another interesting data representation is video. Platforms like YouTube collect huge amounts of data to support research of deep learning. One type of this learning that is used to classify images is Convolutional Neural Network (CNNs). The main reason for not using deep learning is that the images can be large, and the position of the detected object may change. CNN is a neural network that uses convolutional layers based on mathematical operation of convolution. This layer is represented by smaller filters that are applied gradually to the entire image for generating output image [4].

The Convolutional Network can also be used for the dataset of videos. Sports-1M is a dataset made up of 1 million videos on YouTube. The problem may be the huge size of the data set, which makes it impossible loads to local memory. The solution to this problem is using the library for analyzing URLs with dynamic download videos and video's override that are currently in memory. For acceleration, parallel computing is used that works on a different machine than trains the model. The classical convolutional network can be modified to consider time dependencies in the videos. Previous images are stacked on top of the colored axis channel and the input of two image frames in the video has the shape $224 \times 224 \times 6$. There are 3 different strategies for combining input frames shown in Fig. 1. The single-frame model is an example of video classification by simple aggregation of prediction between frames. The late fusion model combines images by grouping the first and last frames in the clip. The early fusion has larger coherent segment from clip and model of the slow fusion, in the combined layers, processed 4 overlapping adjacent segments in the convolution layer. Experimentation has found most success with the slow fusion strategy [5].

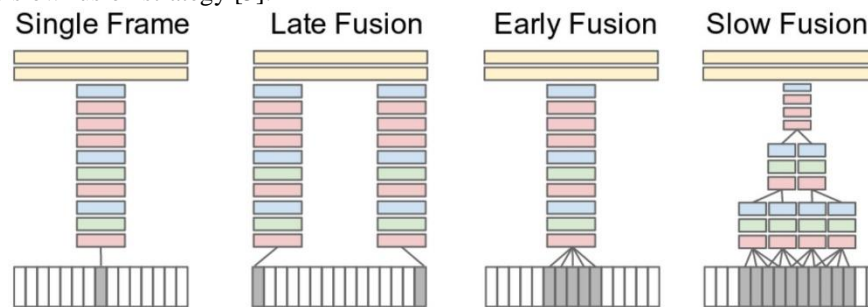


Fig. 1 Model of image classification [5]

CNN training [6] on many datasets takes a long time, even if is available fast CPU. Effort of multifunctional convolution neural network is to change the architecture that achieves faster time without compromising performance. Reduction layers and neurons in each layer achieves lower performance. Instead of reducing the size of the network, the improvement is training on images with lower resolution. The multi-resolution architecture includes two separate processing streams in space as shown. The input is frame of the size 178×178 . The content stream forms the frame in half of the original spatial resolution (89×89) and fovea stream centres of 89×89 in original resolution. Both streams are composited from convolution (red), normalization (green) and association (yellow). In that way the entire size of the input is half. This design has an advantage in presenting live video from the camera, where the main subject is usually in the center [5].

C. Categorization of multimedia content

Automated creation of video categories on a large web collection, such as YouTube, is a challenging task. However, such categories are necessary, because of their importance for indexing and searching purposes. Existing solutions related to automated categorization have been focused especially for a small number of videos or just for a specific area, such as sport. One of the main problems with this categorization is the lack of training data and manually describing video is time-consuming. The solution to these problems is a completely new approach that combines multiple data sources for a varied categorization of YouTube videos. As shown Fig. 2, trained data are from a small number of manually described videos (50 per category) further disseminated by adding captions to related videos, gathering data using Internet search engines (such as Google video search engine), and including data from other domains. To use this data, a method of merging them was developed, three discriminatory random fields using conditional random fields. Each node is linked to a binary description and most likely output training models are used as local information for nodes. Continuous models are then connected with fusion to the final classifier [7].

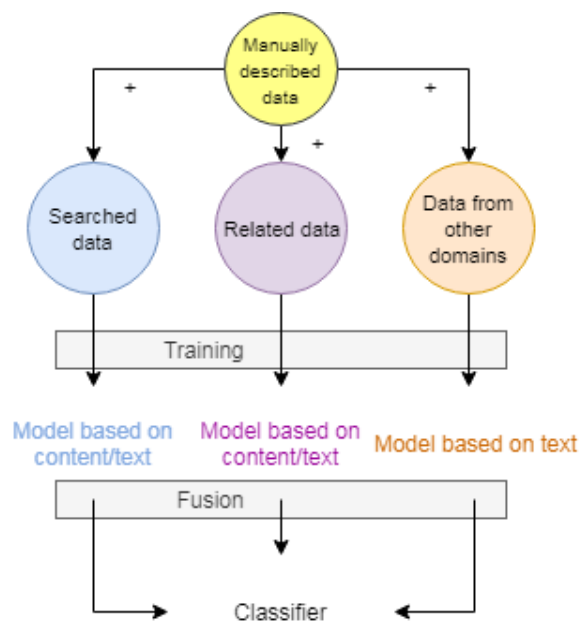


Fig. 2 General framework of proposed solution [7]

D. System of intelligent recommendation in multimedia archives

Nowadays, multimedia archives require mechanisms to display information for users through their interests. Recommendation systems are one of the key methods for searching and displaying information. YouTube is one of the largest platforms for creating, sharing, and discovering videos. Video recommendation system of YouTube is responsible for providing personalized content to more than a trillion users. Recommendation videos are extremely difficult due to three factors [8]:

- Scope - there are many known algorithms that work on smaller problems but fail on larger.
- Video upload time - recommendation system should be responsive enough to newly added content and older.
- Noise - from a historical perspective, user behavior on YouTube is very difficult to predict, especially diversity and various unobservable external factors.

YouTube has done fundamental move to using deep learning as a universal solution for almost all learning-related problems. The recommendation system is built on Google Brain that was recently made available as open-source software, named TensorFlow [9]. It provides a flexible framework for experimentation with various deep neural network architectures. Models learn about a billion parameters and are trained on a huge number of examples. The structure of the recommendation system is composed of two neural networks as is shown in Fig. 3. One network is used for generating candidates and the second to rank them. The network of candidate generation receives input events from history of user activity on YouTube. Its output is a small subset of videos. Network for sorting will decide which of these videos end up as recommended

for user. Operations are based on assigning points to individual videos, according to objective function using a rich set of properties describing video and the user. Videos with the highest score are then presented to the user[9].

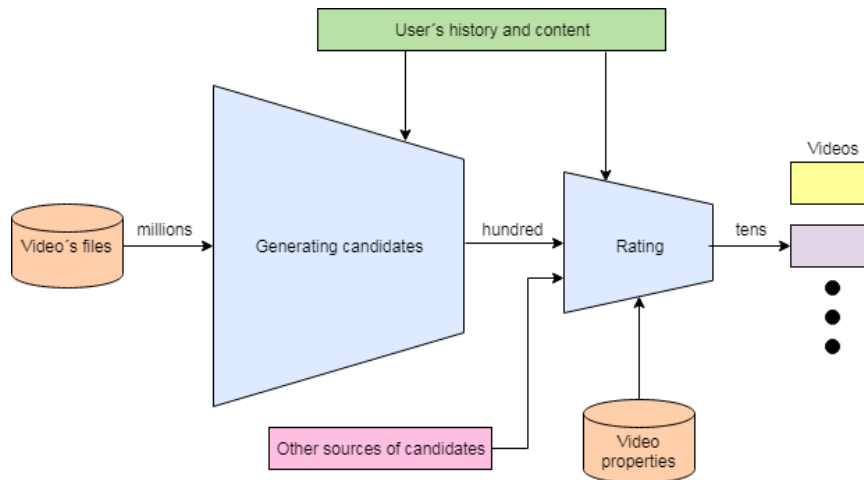


Fig. 3 Recommendation system [8]

E. Searching methods and algorithms for automatic text categorization

If the video contains subtitles, search methods can also be used in text. One of these methods is also text classification, which means assigning a suitable predefined category to text, and can be used for storing, structuring, showing and so on. Text categorization is flexible and a dynamic form of rule-based structuring that uses techniques to automatically assign text to a predefined category and plays a big role in organizing and displaying multimedia information. Texts can have different forms. They are often represented by a group of words where each word or the phrase appears one or more times in the text. They may also contain unnecessary words that make searching more difficult. There is need a reduction where is selected just a subset of the words from the whole text.

Another approach to classifying text is manual categorization, where somebody manually assigns text categories. This approach achieves qualitative results, but it takes time. Automatic categorization uses machine learning, natural language processing and other techniques for automatic text processing. This method is faster but also more costly. Automatic classification can be divided into 3 different types of system [10]:

- Rule-based systems - define the rules under which the system is searching. The rule contains pattern, based on system searches in text and then act.
- Machine learning system - is based on machine learning methods, when the classifier is created automatically based on the properties that are extracted from a set of manually classified documents. This is controlled learning because it is necessary to know the categories that were manually created.
- Hybrid systems – the combination between Rule-based systems and Machine learning system,

Nowadays, multimedia systems allow the user to add a word or a small phrase to an object. This process is called tagging and allows you to tag, or add key words to, video, audio, article. These tags can be used to search for similar objects. Such marking may be a problem in the natural language processing. Tagging in a sequence of words try to predict the tag for each word. The examined model of multi-task training is based on performing multiple sequence labelling tasks in one language. The multi-language tagging model is based on performing the same task over a set of multiple languages.

Based on this, a general model was created, which can be applied in both cases. This model uses depth gate repeated units on characters and word levels and applies conditional random field to create a structured prediction. Network composed of such units is a type of recurrent neural network, first introduced for machine translation. At the character level, these units capture morphological information and at the verbal level, learn N-gram patterns and semantics. This model can handle multi-task and multi-language training in a uniform way by sharing network architecture and model parameters between tasks and languages [11].

III. CONCLUSION

The main goal of this article was to introduce efficient information services in multimedia archives. The task is to acquaint the reader with available options in developing an intelligent video archive. This work described intelligent searching, classification, and categorization of multimedia content. We have briefly described how it works for example in YouTube. We also introduced a system of intelligent recommendation in multimedia archives. As was mentioned, if the video contains subtitles, then search methods or tagging can also be used in text. Considering video archives, first we needed discuss video digital representation and transmission in the environment of multimedia archives. Our future work will build on this article and in developing such a multimedia archive will be used some of mentioned methods and services.

ACKNOWLEDGEMENT

This article was written thanks to the generous support under the Operational Program Integrated Infrastructure for the project "Research Centre for Data Analysis and Data Protection - Stage II", Project No. 313021W479, co-financed by the European Regional Development Fund.

REFERENCES

- [1] *What is Data Archiving?* URL:<https://www.fdiinc.com/what-is-dataarchiving/>.
- [2] Xinyu Wang et al. *Searching method and searching apparatus based on neural network and search engine*. US Patent App. 15/625,460. Máj 2018.
- [3] Lawrence E Page. *Method for searching media*. US Patent 9,684,676. Jún 2017.
- [4] Shuiwang Ji et al. "3D convolutional neural networks for human action recognition". In: *IEEE transactions on pattern analysis and machine intelligence* 35.1 (2012), s. 221–231.
- [5] Andrej Karpathy et al. "Large-scale video classification with convolutional neural networks". In: *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. 2014, s. 1725–1732.
- [6] Matthew D Zeiler a Rob Fergus. "Visualizing and understanding convolutional networks". In: *European conference on computer vision*. Springer. 2014, s. 818–833.
- [7] ZheshenWang et al. "Youtubecat: Learning to categorize wildweb videos". In: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE. 2010, s. 879–886.
- [8] Paul Covington, Jay Adams a Emre Sargin. "Deep neural networks for youtube recommendations". In: *Proceedings of the 10th ACM conference on recommender systems*. ACM. 2016, s. 191–198.
- [9] Martin Abadi et al. "TensorFlow: Large-scale machine learning on heterogeneous systems, 2015". In: *Software available from tensorflow. org* 1.2 (2015).
- [10] Charu C Aggarwal a ChengXiang Zhai. *Mining text data*. Springer Science & Business Media, 2012.
- [11] Zhilin Yang, Ruslan Salakhutdinov a William Cohen. "Multi-task crosslingual sequence tagging from scratch". In: *arXiv preprint arXiv:1603.06270* (2016).

Využitie STT na extrakciu textu vo video sekvenciách v rozsiahlych multimedialných archívoch

¹Dávid CYMBALÁK, ²Miroslav MURIN, ³Roman VÁPENÍK, ⁴František JAKAB

^{1,3} Spoločné pracovisko UVP Technicom a Centra vedecko-technických informácií SR, Technická univerzita v Košiciach, Slovenská republika

^{2,4} Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹david.cymbalak@uvptechnicom.sk, ²miroslav.murin@student.tuke.sk,
³roman.vapenik@cvtisr.sk, ⁴frantisek.jakab@tuke.sk

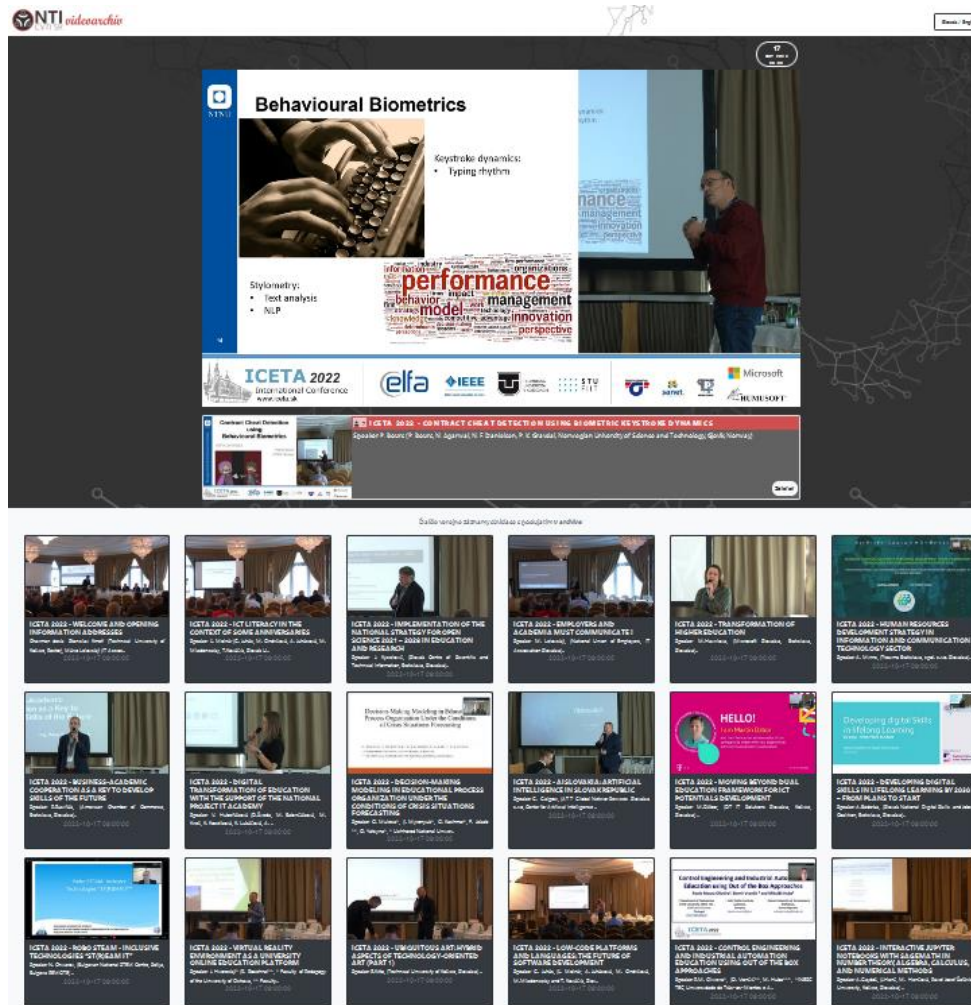
Abstrakt — Cieľom tohto vedeckého článku je prezentovať využitie technológií Speech-to-Text (STT) pri spracovaní rozsiahlych vizuálnych záznamov v multimedialnom archíve videí. Navrhnuté riešenie pracuje s rozsiahlou bázou videí existujúceho multimedialneho archívu NTI, ktorý slúžil ako testbed prostredie pre STT spracovanie. Experimentálne prepojenie nástrojov FFmpeg pre re-streaming a pseudo-streaming existujúcich aj nových vizuálnych záznamov v archíve bolo prepojené s technológiami spracovania SoX a NAudio, čo umožnilo efektívne spracovať audio stopy s riešením VOSK na rôznych jazykových modeloch. Výsledné riešenie extrahujúce reč do textu ponúka možnosti ďalšieho prepojenia systému rozpoznávania textu pre indexovanie a vyhľadávanie v multimedialnom archíve a otvára priestor pri tvorbe špecializovaných aplikácií tejto technológie v rôznych oblastiach.

Kľúčové slová — STT, speech-to-text, multimedialný archív, VOSK

I. ÚVOD

V dnešnej dobe sú multimedialne archívy videí neoddeliteľnou súčasťou mnohých organizácií a spoločností, ktoré potrebujú ukladať a zdieľať videá z rôznych zdrojov. Tieto videá môžu obsahovať dôležité informácie, ktoré môžu byť využité pre vzdelávanie, výskum, obchodné účely alebo zábavu. Avšak správa týchto archívov pozostávajúca s dopĺňania textových informácií, popisov, titulkovania, kategorizácie a pod. môže byť časovo náročná s pribúdajúcim objemom obsahu a vekom samotného archívu. Prehľadávanie a získavanie informácií z rozsiahlych multimedialných archívov môže byť časovo náročné a niekedy aj neefektívne. V oblasti správy a organizovania multimedialných archívov sa pravidelne objavujú nové technológie, ktoré umožňujú najmä efektívnejšie vyhľadávanie a správu videí. Existujúce metódy pre indexovanie a vyhľadávanie videí sú často zamerané na vizuálny obsah videa, ale ignorujú jeho zvukovú stopu. Naopak tento článok sa bude venovať najmä jednej zo stabilne inovujúcich sa technológií v oblasti spracovania zvukovej stopy - Speech-to-Text (STT), ktorá umožňuje konverziu hovoreného slova na text. V ponímaní multimedialných archívov ňou vieme dosiahnuť automatickú extrakciu textu z rečového obsahu audio stopy videa. Extrahovaný text z videa je okrem funkcie titulkovania následne možné použiť pre efektívnejšie indexovanie, kategorizáciu, metatagy a vyhľadávanie videí v multimedialnom archíve.

Ako testbed prostredie pre spracovanie rozsiahlych video záznamov bude slúžiť multimedialný archív Národnej teleprezentačnej infraštruktúry (NTI) [1] pre podporu výskumu, vývoja a transferu technológií na Slovensku, ktorá bola pôvodne vytvorená v roku 2015 Centrom vedecko-technických informácií SR v roku. Počas takmer 8 ročnej existencie NTI bolo realizovaných niekoľko tisíc živých vysielaní, čo vytvorilo rozsiahly archív multimedialných záznamov prístupných pre stovky tisíc divákov [2]. Multimedialný archív NTI je tvorený verejným aj privátnym katalógom záznamov (Obr. 1) s podsekciami obsahujúcimi popisné informácie, metatagy a zaradenie do kategórií. Tieto informácie sa počas dlhoročnej prevádzky multimedialneho archívu spracovali manuálne, avšak s možnosťami nasadenia technológie STT je práve toto prostredie vhodné na automatizáciu v kategorizácii, indexácii a tvorby popisov k videám v archíve.



Obr. 1 NTI a multimedialný archív s tisícami záznamov obsahujúcich rozsiahle videozáznamy so zvukovou stopou vhodnou na STT spracovanie

II. TECHNOLÓGIA STT A NÁSTROJE SPRACOVANIA AUDIA Z VIDEO STREAMU

Ako už bolo úvodom spomenuté Speech-to-Text technológia umožňuje previesť rečový obsah videa na textový formát. Táto technológia využíva metódy strojového učenia, ktoré umožňujú rozpoznávanie reči a konverziu do textovej podoby. Existuje niekoľko algoritmov, komerčných a open-source STT nástrojov. Medzi najznámejšie komerčné STT nástroje patria Google Cloud Speech-to-Text, Amazon Transcribe, Microsoft Azure Speech Services a IBM Watson Speech-to-Text. Open-source STT nástroje zahŕňajú napríklad Kaldi, Mozilla DeepSpeech, VOSK a Pocketsphinx.

Proces konverzie zvukového obsahu na text v STT systéme sa zvyčajne skladá z niekoľkých krokov. Prvým krokom je zber zvukového signálu, ktorý môže byť získaný z rôznych zdrojov, ako sú napríklad mikrofóny, nahrávacie zariadenia alebo internetové streamy. Nasleduje predspracovanie, ktoré upravuje zvukový signál náchylný na rôzne šумы a nepresnosti, ktoré môžu negatívne ovplyvniť proces rozpoznávania reči. Predspracovanie teda môže zahŕňať rôzne filtre, normalizáciu a odstránenie šumu. Po predspracovaní sa vstupný zvukový signál rozpoznáva a konvertuje do textovej podoby pomocou rôznych algoritmov, ako sú napríklad HMM-GMM [3], DNN [4] a LSTM [5]. Nasleduje dekódovanie, ktoré sa používa na výber najpravdepodobnejšieho výstupu z množiny hypotéz generovaných vybraným rozpoznávacím algoritmom. Výsledný text môže byť ďalej upravený pomocou rôznych postprocesovacích krokov, ako sú napríklad korekcia gramatiky, normalizácia textu alebo segmentácia do viet.

A. Aplikácia SoX a knižnica NAudio

Z ponúkaného spektra nástrojov pre primárne spracovania audia boli vybrané ako vhodné pre nasadenie do nášho riešenia práve aplikácia SoX a knižnica NAudio.

SoX je open-source nástroj fungujúci v príkazovom riadku, ktorý poskytuje množstvo funkcií pre prácu s audio súborami. SoX umožňuje konvertovať audio súbory do rôznych formátov, meniť vzorkovaciu frekvenciu a bitovú hĺbku, pridať efekty ako echo, zosilniť, normalizovať, filtrovať šum a mnoho ďalších. SoX taktiež umožňuje spojenie, rozdelenie audio súborov, generovanie

signálov a podporuje mnoho formátov audio súborov a kodekov [6].

Knižnica NAudio je open-source knižnica pre prácu s audio stopou v jazyku C#. Knižnica NAudio umožňuje čítať a zapisovať rôzne formáty audio súborov, spracovávať a prehrávať audio signály a realizovať efekty ako echo, zosilnenie, normalizácia a mnoho ďalších. Knižnica NAudio ponúka aj funkcionality pre nahrávanie zvuku zo vstupného zariadenia alebo simulovaného audio streamu [7].

Hlavným rozdielom medzi nástrojom SoX a knižnicou NAudio je spôsob ich použitia. SoX je vhodný pre použitie v príkazovom riadku alebo v skriptoch, knižnica NAudio je určená pre použitie v jazyku C# a je vhodná pre využitie priamo v aplikáciách.

B. Nástroj pre rozpoznanie reči VOSK

VOSK je STT nástroj s otvoreným zdrojovým kódom, ktorý je navrhnutý na presné a efektívne rozpoznávanie reči v rôznych jazykoch. VOSK je založený na modeloch strojového učenia, ktoré sú trénované na rozpoznávanie reči a následné preklady a extrakciu textu. Jeho výkonnosť a presnosť ho predurčujú ako dôležitý nástroj pre mnoho aplikácií automatizácie a spracovania hovoreného slova. Je ľahko integrovateľný do rôznych programovacích jazykov, vrátane Pythonu, C++, Java a pod., no vie byť tiež dostupný ako webová služba, ktorá umožňuje využívať jeho funkcie pomocou API [8].

Zaujímavou možnosťou je to, že VOSK vie fungovať v režime offline sady STT nástrojov s množstvom externých modelov aj v rôznych svetových aj minoritných jazykoch. V závislosti od veľkosti a komplexnosti jazykového modelu, je možné dosiahnuť väčšiu presnosť detekcie slov. Veľkosť modelu ale rovnako vplyva na rýchlosť spracovania a požadovaný výkon. Z praktických testov sa ukázalo, že modely malej veľkosti bežia rýchlo aj na pomalých zariadeniach no nevýhodou je menšia presnosť. O kvalite prevodu reči do textu teda vo výsledku rozhoduje použitý model, ktorého narastajúca komplexnosť ide výrazne na úkor dlhšieho spracovania textu. Komplexnejšie modely môžu zaberat' aj viac ako 1GB a vyžadujú výkonnejšie zariadenie.

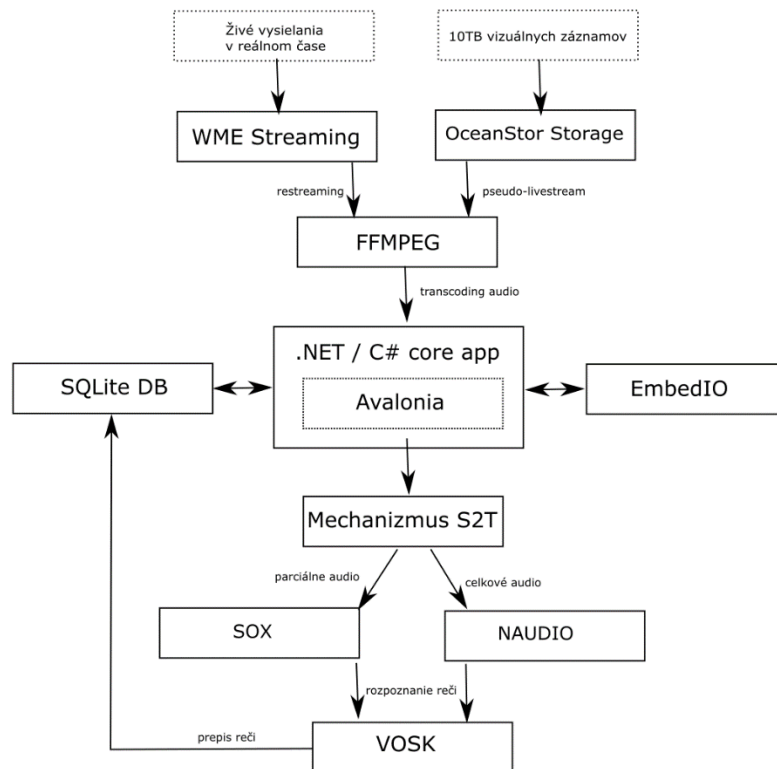
Ideálny zvuk pre spracovanie hovorenej reči nástrojom VOSK by mal byť v jednom audio kanáli o veľkosti vzorky 16000Hz. Pri klasickom režime sa nástroju poskytne extrahovaná nahrávka a ten danú nahrávku analyzuje a vykoná príkazy. V druhom režime sa aktívne pristupuje k vstupnému audio zariadeniu (alebo simulovanému audio vstupu) a získané dáta sa posúvajú na spracovanie, kde ich VOSK neustále čiastočne analyzuje.

C. Multimediálny nástroj FFmpeg

FFmpeg je veľmi mohutný multimediálny rámec, ktorý dokáže veľké množstvo funkcionalít ako kódovanie, dekódovanie, streamovanie a prehrávanie takmer každého audiovizuálneho formátu. Ďalšou výhodou je jednoznačne to, že je vysoko prenosný a je ho možné spustiť na operačných systémoch Windows, Linux, Mac OS X, BSD, Solaris a ďalšie. FFmpeg ponúka množstvo funkcionalít, ktoré budú využiteľné práve v navrhovanom riešení pre re-streaming vysielania, presudostreaming existujúceho záznamu alebo oddelenie zvukovej stopy pre spracovanie. Dokáže pracovať s živým streamom v rôznych protokoloch a ľubovoľnom kodeku. Ďalšou funkcionalitou, ktorú FFmpeg dokáže, je konvertovať jeden typ multimediálneho súboru na druhý, čím je možné dosiahnuť efektívne kompresné metódy a taktiež vie nahrávať živo streamované videá do požadovaného súborového kontajnera a kodeku. Výhodou FFmpeg [9] je to, že podporuje veľké množstvo prenosových protokolov, formátov kontajnerov médií a štandardov kódovania a vďaka tomu je s týmto nástrojom rýchle a efektívne analyzovať stream obrazu a zvuku.

III. NÁVRH NASADENIA STT V ROZSAHLOM MULTIMEDIÁLNO M ARCHÍVE S PREPOJENÍM NA NAHRÁVANIE NOVÝCH ZÁZNAMOV VYSIELANÍ

Existujúce prepojenie s rozsiahlym multimediálnym archívom je zabezpečené prostredníctvom realizácia pseudo-vysielania vybraného záznamu zo zdieľaného úložiska OceanStor cez nástroj FFmpeg, ktorý okrem živého streamovania MP4 súboru zabezpečí aj oddelenie audio stopy. V prípade novovytváraných záznamov vstupujúcich do archívu je možné prostredníctvom FFmpeg realizovať re-streaming živého vysielania z Wowza Media Engine streamingového servera s oddelením audio stopy pre spracovanie. Kódovaná audio stopa teda vstupuje do jadra aplikácie postavenej na .NET / C# s využitím Avalonia a prepojením na EmbedIO. Prepojenie jednotlivých prvkov navrhnutého riešenia je možné vidieť na Obr. 2.



Obr. 2 Prepojenie prvkov riešenia pre rozpoznávanie textu vo videu z rozsiahlej databázy existujúceho archívu vizuálnych záznamov

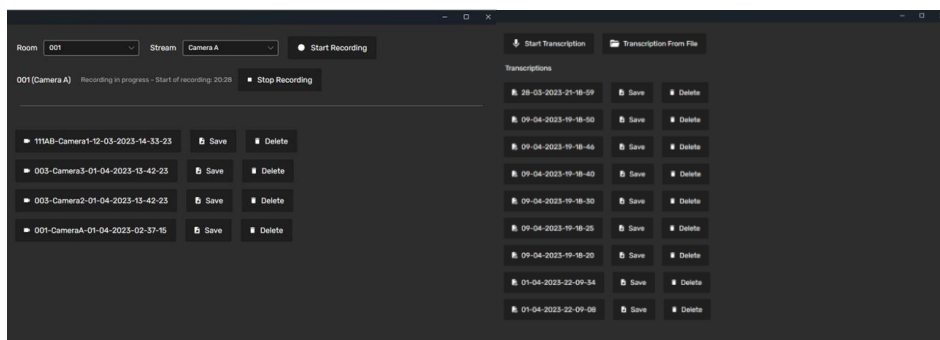
Mechanizmus S2T pozostáva v prvej fáze so samotného spracovania audio signálu, kde sú k dispozícii 2 režimy - klasický a čiastočný režim rozpoznávania reči.

Klasický režim využíva na nahrávanie vstupu zo streamovaného audio toku multiplatformovú aplikáciu SoX. Vďaka tomu je tento režim funkčný na každom operačnom systéme ako Windows, macOS alebo Linux. Tento režim funguje tak, že sa spracováva iba čiastočná sekvencia z videa pre účely extrakcie, napr. úvodných informácií videa a názvu.

Druhý režim prístupu k audio je čiastočný režim rozpoznávania reči. Tento režim funguje iba na operačnom systéme Windows. Využíva knižnicu s otvoreným zdrojovým kódom NAudio s funkcionalitou WaveIn, kde sa stále pristupuje k vstupnému streamovanému audio toku a neustále sa zvuk vyhodnocuje. Tento režim je vykonávaný neustále pokiaľ je zapnutý a je možné využiť aj odoslanie HTTP požiadaviek, teda automaticky rozpoznávané vety sú odosielané na server. Takéto parciálne rozpoznávanie reči má výhodu, že akcia sa vykonáva vždy len keď rozpozná slovo. Táto funkcionalita je taktiež vhodná na automatickú tvorbu tituliek videa v reálnom čase.

Rozpoznávanie reči a extrakcia do textu sa v záverečnej fáze deje pomocou spomenutej knižnice VOSK, ktorá posúva extrahované texty späť do databázy, kde budú slúžiť ako základ pre externé mechanizmy indexácie, vyhľadávania a automatizovaných popisov a tituliek. K dispozícii je množstvo naučených modelov pre rôzne jazyky. Výhodou technológie rozpoznávania reči používanej v aplikácii je to, že vie fungovať aj v režime bez sieťového pripojenia. Všetko sa deje iba v rámci lokálneho počítača. Hovorené slovo sa nikde neodosiela a rozpoznávanie a celá logika za danou funkcionalitou ostáva v rámci aplikácie.

Na Obr. 3 je experimentálne rozhranie sekcie prepisu reči, kde vie používateľ ľahko zapnúť zaznamenávanie reči zo zvoleného audio streamovaného zdroja. Po ukončení sa automaticky nahrávka vyhodnotí a vytvorí sa textový súbor. Prepis reči do textu je rovnako možné ovládať aj pomocou webového servera. Ďalšou možnosťou je manuálne poskytnutie extrahovanej re-streamovanej nahrávky, ktorú program spracuje a vytvorí súbor. Výstupné súbory s prepismi je možné priamo otvoriť, prezrieť text, uložiť do textovej podoby, zmazať alebo naviazať na externé riešenia indexácie a automatizovaných popisov.



Obr. 3 Experimentálne rozhranie pre výber audio streamovaného vstupu na realizáciu STT a extrakciu textu

IV. ZÁVER

Multimediálne archívy videí majú obrovský potenciál na poskytnutie informácií, avšak prehľadávanie a získavanie informácií z týchto archívov môže byť časovo náročné a neefektívne. Ako bolo v tomto článku preukázané, využitie STT technológie v multimediálnych archívoch prináša mnoho výhod a prínosov. Je to dôležitý nástroj pre lepšie a rýchlejšie vyhľadávanie v archívoch videí, ktorý môže byť použitý v rôznych oblastiach pre efektívne a rýchle spracovanie a prácu s rozsiahlymi vizuálnymi záznamami a informáciami. Indexovanie multimediálneho archívu videí pomocou Speech-to-Text umožňuje lepšie a efektívnejšie vyhľadávanie videí. Touto technológiou je možné dosiahnuť automatické pridanie textového popisu do každého videa. Tento popis umožňuje presnejšie vyhľadávanie podľa kľúčových slov a tém. Okrem vyhľadávania môže STT technológia byť využitá aj na ďalšie účely v multimediálnych archívoch, ako napríklad automatické generovanie titulkov pre videá. STT technológia môže byť tiež použitá na zaznamenávanie a transkribovanie rozhovorov alebo prednášok pre ďalšie použitie a archiváciu.

V niektorých prípadoch môže byť konverzia reči na text nepresná, čo môže viesť k chybám pri indexovaní a vyhľadávaní. Tieto chyby je možné eliminovať práve návrhom vhodného riešenia prezentovaného v tomto článku s využitím efektívnych technológií STT ako VOSK a nástrojov zvukového spracovania SoX, NVAudio s prepojením na existujúcu bázu multimediálneho archívu cez FFmpeg. Navrhovaný experimentálny systém je samozrejme možné a aj potrebné naďalej rozširovať, rozvíjať jazykové modely a urobiť ho viac inteligentným vo vzťahu prepojenia na externé systémy.

POĎAKOVANIE

Táto publikácia vznikla vďaka podpore v rámci Operačného programu Integrovaná infraštruktúra pre projekt: „Výskumné centrum pre analýzu a ochranu dát - II. Etapa“, Kód ITMS2014+: 313021W479, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.

LITERATÚRA

- [1] CYMBALAK, D., et al. Adaptation of Multisource Video Streaming in Heterogenous Environment of National Telepresence Infrastructure with Advanced Elements of Processing Visual Recordings to Multimedia Archive. In: 2022 20th International Conference on Emerging eLearning Technologies and Applications (ICETA). IEEE, 2022. p. 117-121.
- [2] JAKAB, F., et al. Technical Challenges and Vision of Slovak National Telepresence Infrastructure during the COVID-19 Pandemic. In: 2020 18th International Conference on Emerging eLearning Technologies and Applications (ICETA). IEEE, 2020. p. 205-210.
- [3] BANSAL, Poonam, et al. Improved hybrid model of HMM/GMM for speech recognition. 2008.
- [4] DENG, Li; HINTON, Geoffrey; KINGSBURY, Brian. New types of deep neural network learning for speech recognition and related applications: An overview. In: 2013 IEEE international conference on acoustics, speech and signal processing. IEEE, 2013. p. 8599-8603.
- [5] GRAVES, Alex; JAITLY, Navdeep; MOHAMED, Abdel-rahman. Hybrid speech recognition with deep bidirectional LSTM. In: 2013 IEEE workshop on automatic speech recognition and understanding. IEEE, 2013. p. 273-278.
- [6] GENG, Mengzhe, et al. Investigation of data augmentation techniques for disordered speech recognition. *arXiv preprint arXiv:2201.05562*, 2022.
- [7] WANG, Hui Hui. Speech Recorder and Translator using Google Cloud Speech-to-Text and Translation. *Journal of IT in Asia*, 2021, 9.1: 11-28.
- [8] HÜBERT, H.; TALIRONAK, V.; YUN, H. S. AI BASED GESTURE AND SPEECH RECOGNITION TOOL FLOW FOR EDUCATORS. In: ICERI2022 Proceedings. IATED, 2022. p. 6395-6400.
- [9] LEI, Xiaohua; JIANG, Xiuhua; WANG, Caihong. Design and Implementation of a Real-Time Video Stream Analysis System Based on FFMPEG. In: 2013 Fourth World Congress on Software Engineering. 2013, s. 212-216. Dostupné z doi: 10.1109/WCSE.2013.38

Video archív s efektívnym odporúčaním obsahu a kategorizácie na báze tagov a tituliek

¹Roman VÁPENÍK, ²Dávid CYMBALÁK, ³Nikola MÓRICOVÁ, ⁴František JAKAB

^{1,2} Spoločné pracovisko UVP Technicom a Centra vedecko-technických informácií SR,
Technická univerzita v Košiciach, Slovenská republika

^{3,4} Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky,
Technická univerzita v Košiciach, Slovenská republika

¹roman.vapenik@cvtisr.sk, ²david.cymbalak@uvptechnicom.sk,
³nikola.moricova@student.tuke.sk, ²frantisek.jakab@tuke.sk

Abstrakt — Článok predstavuje riešenie efektívnej kategorizácie na báze tagov a inteligentného odporúčania relevantného obsahu aplikovateľné do prostredia rozsiahlych video archívov. Pre experimentálne overenie bola použitá dátová množina videozáznamov od organizácie TED. Analyzované boli metódy kategorizácie, vyhľadávania a odporúčania v multimediálnych archívoch. Primárne bol navrhnutý najvhodnejší prístup pre efektívnu kategorizáciu a odporúčanie v takom archíve. Kategorizácia bola navrhnutá a experimentálne realizovaná na základe extrakcie tagov a tituliek a výsledky boli porovnané. Model kategorizácie tagov dosahoval úspešnosť až 90% a bol implementovaný do výsledného archívu. Ďalej bolo realizované odporúčanie na základe kosínusovej podobnosti, ktoré pre každé video odporúča ďalší relevantný video obsah. Tieto prístupy boli následne implementované do celkového experimentálneho multimediálneho archívu a otestované na zvolenej dátovej množine. Výsledkom týchto aktivít je efektívnejší prístup pri správe multimediálneho archívu a jeho využívaní.

Kľúčové slová — multimediálny archív, kategorizácia, odporúčanie obsahu, VOD

I. ÚVOD

Archivácia zabezpečuje bezpečné uchovávanie informácií v rôznych formátoch na dlhú dobu, bez ohľadu na to, či budú tieto údaje v budúcnosti použité alebo nie. Vzhľadom na neustále zvyšovanie množstva takýchto údajov je dôležité ich správne skladovať a riadiť, aby boli ľahko dostupné pre ich určený účel. Napriek tomu môže byť udržiavanie týchto veľkých objemov údajov náročné. Vzhľadom na rastúci multimediálny obsah na internete je klasifikácia a efektívne zobrazovanie obsahu stále dôležitejšie, aby boli tieto údaje ľahšie dostupné a využiteľné pre používateľov. Preto je dôležité vytvárať archívy s lepšou kategorizáciou, odporúčaním, vyhľadávaním a inými podpornými funkcionálnymi vlastnosťami, ktoré uľahčia prístup k týmto údajom. Vzhľadom na náročnosť klasifikácie multimediálnych dát na internete môže návrh efektívneho mechanizmu výrazne pomôcť a uľahčiť celý proces.

Ďalším z dôležitých vylepšení multimediálnych archívov je automatizované odporúčanie obsahu, ktoré pomáha používateľovi rozhodnúť sa alebo navrhnuť, čo by mohlo byť pre neho zaujímavé. Toto odporúčanie zvyšuje aj celkový dojem používateľa z archívu a jeho informovanosť. Napriek tomu môže byť manuálne programovanie systému pri rozsiahlom archíve zdĺhavé a neefektívne. Preto je dôležité vytvoriť systém s efektívnym algoritmom a vhodným dátovým filtrom, ktorý automaticky odporúča vhodný obsah pre konkrétneho používateľa.

II. ANALÝZA TECHNOLÓGIÍ KATEGORIZÁCIE A ODPORÚČANIA OBSAHU VO VIDEO ARCHÍVOCH

Efektívnu kategorizáciu je možné uskutočniť priamo rozpoznávaním obrazu z videa a odporúčaním na základe jeho metadát. Rozpoznávanie obrazu a zaradenie do kategórie je možné realizovať s použitím algoritmov počítačového videnia (napríklad YOLO [1] na báze CNN) a dostupnosťou dátových množín na tréning, akými sú napr. Sports-1M [2], pre

športové videá, a ActivityNet [3] pre ľudské aktivity. Oproti menej náročnej dátovej množine obrázkov ImageNet [4], ktorá obsahuje všeobecné subjekty, sú existujúce video benchmarky obmedzené na akčné a športové triedy. V tejto oblasti sa dátová množina YouTube-8M [5] jednoznačne líši od ostatných video benchmarkov, ktorá nie je obmedzená len na samotné akčné triedy. Pozostáva zo subjektov ktoré pokrývajú aktivity (šport, hry, záľuby), objekty (autá, potraviny, výroby), scény (cestovanie) a udalosti.

Keďže výsledný archív bude orientovaný na vzdelávaciu komunitu, kde človek zvyčajne prezentuje publiku, práca s rozpoznávaním z videa v tomto prípade nie je efektívna. Systém by nemal dostatočné vzorky dát na video, podľa ktorých by sa mal rozhodnúť a priradiť video do kategórie. Ak však video obsahuje titulky, potom je možné využiť práve metódy pracujúce s textom, ktoré budú predmetom týchto výskumných aktivít.

V spojení s inými oblasťami spoločnosti Google, YouTube prešiel zásadným posunom k využívaniu hlbokého učenia ako univerzálneho riešenia pre takmer všetky problémy súvisiace s učením. Systém odporúčania je postavený na Google Brain, ktorý bol sprístupnený ako softvér s otvoreným zdrojovým kódom pod názvom TensorFlow [6]. Poskytuje flexibilný rámec pre experimentovanie s rôznymi architektúrami hlbokéj neurónovej siete s využitím rozsiahleho tréningu. Modely sa učia miliardy parametrov a sú tréňované na obrovskom počte príkladov. Štruktúra systému odporúčania je zložená z dvoch neurónových sietí. Jedna sieť slúži na generovanie kandidátov a druhá na ich zoradenie. Sieť na generovanie kandidátov dostáva ako vstup udalosti z histórie aktivity používateľa na YouTube. Jej výstupom je malá podmnožina videí z ich rozsiahleho súboru. Táto sieť poskytuje iba širokú personalizáciu prostredníctvom kolaboratívneho filtrovania. O tom, ktoré z týchto videí skončia ako odporúčané pre daného používateľa rozhodne sieť určená na ich zoradenie. Jej fungovanie je založené na priradovaní bodov jednotlivým videám, podľa požadovanej objektívnej funkcie, s použitím bohatého súboru vlastností popisujúcich video a samotného používateľa. Videá s najvyšším bodovým ohodnotením sú následne prezentované používateľovi [7].

Autori [8] uvádzajú, že cieľom systému je poskytnúť personalizované odporúčanie, ktoré pomôže používateľom nájsť relevantné videá k ich záujmom. Ako už bolo spomenuté, odporúčanie videí je extrémne náročné z dôvodu prítomnosti rôznych faktorov. Pre udržanie záujmu, je nutné aby sa systém aktualizoval pravidelne a odrážal aktivitu používateľa. Preto obsahuje 2 triedy údajov: samotný video stream, metadáta (titulky) a aktivity používateľa.

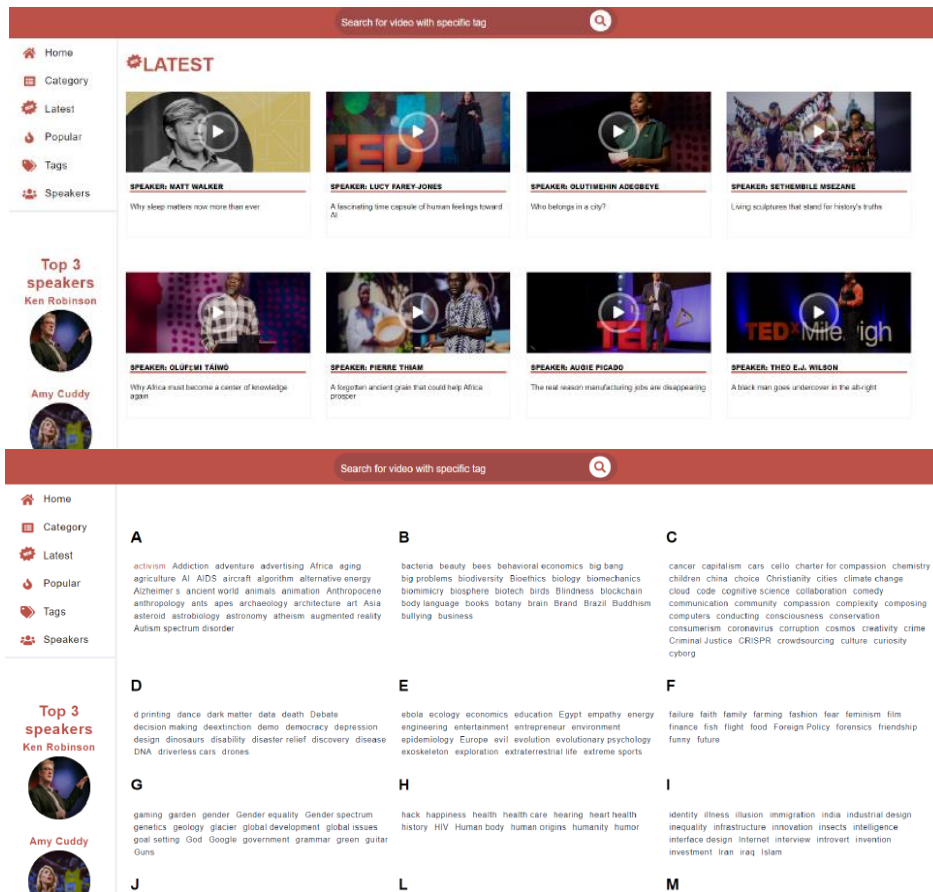
III. NÁVRH A IMPLEMENTÁCIA RIEŠENIA PRE ROZPOZNÁVANIE V ROZSIAHLOM ARCHÍVE VIZUÁLNYCH ZÁZNAMOV

Pre návrh efektívnej kategorizácie a odporúčania v multimediálnom archíve je potrebné zvoliť vhodnú dátovú množinu, nad ktorou budú ďalej realizované operácie. Archív bude orientovaný na videá, kde ľudia zvyčajne prezentujú publiku. Keďže nemáme dostatočnú vzorku vlastných dát, je potrebné nájsť čo najviac podobnú množinu dát s videami. Videá na tieto potreby sa nachádzajú napríklad na stránke TED Talks. Dátová množina s videami z tejto komunity poskytuje údaje o videách ako sú titulky, názvy, meno autora, webové adresy na video, hodnotenie a podobne. Keďže táto komunita je všeobecne známa, je nutné mať povolenie a licenciu aj na nekomerčné účely. Pre túto prácu bolo schválené použitie ich materiálov priamo od organizácie TED s licenciou TED's Creative Commons license - Creative Commons BY-NC-ND 3.0 na nekomerčné účely. Príklad časti takejto dátovej množiny na Obr. 1 obsahuje okolo 2500 vzoriek videí. Keďže dátová množina obsahuje slová v anglickom jazyku, aj multimediálny archív bude zostavený v tomto jazyku, a teda dáta uložené v databáze aj informácie zobrazené na stránke budú v anglickom jazyku.

komentár: popis	trvanie: akcia	dátum natočenia jazyk	Návrny prednášajúci meno	počet prednášajúcich dátum publikovania	hodnotenie
4554: Sen Ken Robinson makes an entertaining a...	1164 TED2006	1140825600	60 Ken Robinson	Ken Robinson: Do s	1 1151367060 [{"id": 7, "name": "Funny", "c
265 With the same humor and humanity he ex...	977 TED2006	1140825600	43 Al Gore	Al Gore: Averting th	1 1151367060 [{"id": 7, "name": "Funny", "c
124 New York Times columnist David Pogue ti...	1286 TED2006	1140739200	26 David Pogue	David Pogue: Simpl	1 1151367060 [{"id": 7, "name": "Funny", "c
200 In an emotionally charged talk, MacArthur	1116 TED2006	1140912000	35 Majora Carter	Majora Carter: Gre	1 1151367060 [{"id": 3, "name": "Courage
593 You've never seen data presented like thi...	1190 TED2006	1140566400	48 Hans Rosling	Hans Rosling: The b	1 1151440680 [{"id": 9, "name": "Ingeniou
672 Tony Robbins discusses the "Invisible forc...	1305 TED2006	1138838400	36 Tony Robbins	Tony Robbins: Why	1 1151440680 [{"id": 7, "name": "Funny", "c
919 When two young Mormon missionaries ki...	992 TED2006	1140739200	31 Julia Sweeney	Julia Sweeney: Lett	1 1152490260 [{"id": 3, "name": "Courage
46 Architect Joshua Prince-Ramus takes the	1198 TED2006	1140652800	19 Joshua Prince-Ram	Joshua Prince-Ram	1 1152490260 [{"id": 9, "name": "Ingeniou
852 Philosopher Dan Dennett calls for religio...	1485 TED2006	1138838400	32 Dan Dennett	Dan Dennett: Let's	1 1153181460 [{"id": 3, "name": "Courage
900 Pastor Rick Warren, author of "The Purpo...	1262 TED2006	1140925600	31 Rick Warren	Rick Warren: A life	1 1153181460 [{"id": 21, "name": "Beautiful
79 Accepting his 2006 TED Prize, Cameron Sii...	1414 TED2006	1140912000	27 Cameron Sinclair	Cameron Sinclair: A	1 1153786260 [{"id": 3, "name": "Courage
55 Jehane Noujaim unveils her 2006 TED Priz...	1538 TED2006	1140912000	20 Jehane Noujaim	Jehane Noujaim: M	1 1153786260 [{"id": 1, "name": "Beautiful
71 Accepting the 2006 TED Prize, Dr. Larry Br...	1550 TED2006	1140652800	24 Larry Brilliant	Larry Brilliant: My v	1 1153786260 [{"id": 8, "name": "Informat
242 Jeff Han shows off a cheap, scalable mult...	527 TED2006	1139184000	27 Jeff Han	Jeff Han: The radiu	1 1154391060 [{"id": 9, "name": "Ingeniou
91 Nicholas Negroponte, founder of the MIT	1057 TED2006	1140652800	25 Nicholas Negropont	Nicholas Negropont	1 1154391060 [{"id": 3, "name": "Courage
325 Violinist Sirena Huang gives a technically l...	1481 TED2006	1140652800	31 Sirena Huang	Sirena Huang: An I	1 1154995860 [{"id": 1, "name": "Beautiful
305 Pianist and composer Jennifer Lin gives a...	1445 TED2004	107753600	32 Jennifer Lin	Jennifer Lin: Impr	1 1154995860 [{"id": 1, "name": "Beautiful
88 Fumes from indoor cooking fires kill more...	906 TED2006	1140739200	27 Amy Smith	Amy Smith: Simple	1 1155600660 [{"id": 9, "name": "Ingeniou
163 Designer Ross Lovegrove expounds his ph...	1170 TED2005	1109289600	22 Ross Lovegrove	Ross Lovegrove: Or	1 1155600660 [{"id": 1, "name": "Beautiful
84 Jimmy Wales recalls how he assembled "c...	1201 TEDGlobal 20	1121299200	32 Jimmy Wales	Jimmy Wales: The I	1 1156119060 [{"id": 1, "name": "Beautiful
108 In 2006, open-learning visionary Richard E...	1114 TED2006	1140652800	27 Richard Baraniuk	Richard Baraniuk: T	1 1156119060 [{"id": 9, "name": "Ingeniou
185 Performer and sub toy maker Ze Frank de...	1136 TED2004	107753600	26 Ze Frank	Ze Frank: Nerdcore	1 1156464660 [{"id": 7, "name": "Funny", "c
106 The founding mother of the blog revoluti...	1006 TED2006	1140652800	20 Mena Trott	Mena Trott: Meet t	1 1156464660 [{"id": 8, "name": "Informat
556 Anthropologist Helen Fisher takes on a tri...	1407 TED2006	1140739200	33 Helen Fisher	Helen Fisher: Why s	1 1157501460 [{"id": 7, "name": "Funny", "c

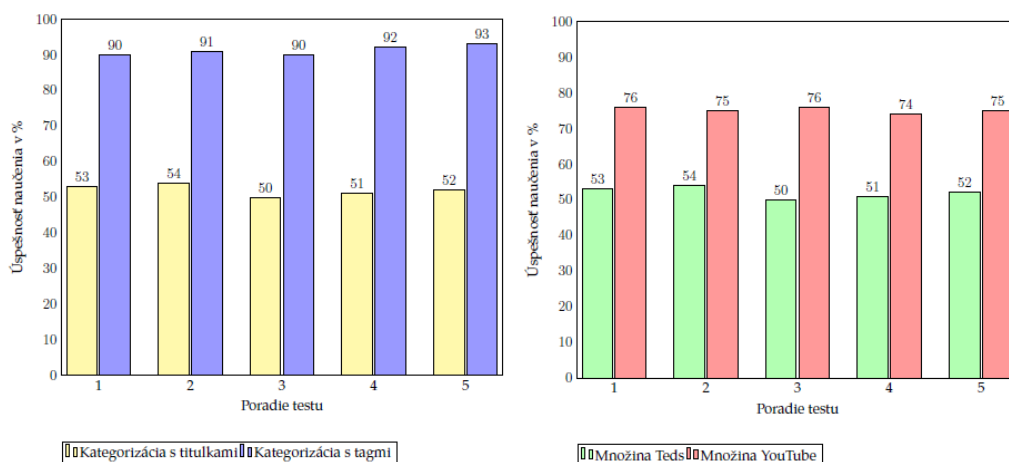
Obr. 1 Časť dátovej množiny TED Talks

Multimediálny archív je rozdelený na dve časti. Všetko to, čo sa deje v pozadí používateľa je implementované v programovacom jazyku Python. Učenie kategórií bolo najprv implementované z tituliek aj tagov v Jupyteri, následne riešenie s najvyššou úspešnosťou na vzorke dát bolo implementované do aplikácie. Vyhľadávanie v tagoch bolo implementované priamo do aplikácie za využitia knižnice FuzzyWuzzy. Databáza bola využitá typu PostgreSQL a aplikácia sa na ňu pripája pomocou technológie psycopg2. Zobrazenie dát z databázy vo webovom prehliadači cez aplikáciu zabezpečuje REST API, v programovacom jazyku Python to zabezpečuje technológia Flask. Táto technológia umožňuje zobrazenie získaných dát vo webovom prehliadači na konkrétnej vytvorenej adrese. Efektívne odporúčanie je uskutočnené pomocou knižníc sklearn, pandas, numpy, ktoré sú podporované v programovacom jazyku Python. Taktiež bola vytvorená autentifikácia pre prihlasovanie do administrátorskej časti pomocou technológie jwt. Prvotným plánom bolo zefektívniť zaužívaný webový archív, čo si vyžaduje aj vytvorenie nového webového riešenia, pri ktorého tvorbe bol použitý rámec React.js (Obr. 2).



Obr. 2 Ukážka používateľského rozhrania podstránky s tagmi

Pri implementácii tohto riešenia boli testované dva spôsoby efektívneho učenia kategórií na zvolenej množine dát. Najprv bol vytvorený model učenia z tituliek a následne z tagov. Porovnanie výsledkov kategorizácie v piatich testoch sú znázornené v grafe na Obr. 3, teda v testovaní bol vždy takmer o polovicu úspešnejší model kategorizácie z tagov. Keďže sa predpokladalo, že dôvodom nízkej úspešnosti v prípade tituliek môže byť aj malá vzorka dát, kategorizácia z tituliek bola otestovaná aj na väčšej dátovej vzorke. Množina obsahovala 13 040 videí s prepisom tituliek do textu a kategórií. Porovnanie úspešnosti kategorizácie z tituliek na menšej množine dát Teds a väčšej z YouTube v piatich testoch je znázornené v grafe na Obr. 3. Graf ukazuje, že klasifikácia na nových dátach má vyššiu úspešnosť ako v predošlom testovaní, čím sa dosahuje to, že algoritmus je použiteľný aj s titulkami. Problémom je teda množstvo dát a na presnejšie odhadnutie kategórie z prepisu tituliek je potrebné mať väčšiu vzorku. V takomto prípade, tým že je dátová množina niekoľko násobne väčšia, tak aj jednotlivé kategórie obsahujú väčší počet rôznorodých prepisov tituliek a model má vyššiu úspešnosť. Vo výslednom multimediálnom archíve bol využitý model učenia kategórií pomocou tagov, pretože bolo nutné využiť vo výslednom multimediálnom archíve prvú dátovú množinu a na nej bola úspešnejšia práve táto kategorizácia.



Obr. 3 Graf porovnania kategorizácie na zvolených množinách na základe titulek a tagov

IV. ZÁVER

Navrhnuté a experimentálne implementované riešenie bolo overené porovnaním výsledkov z tréningu kategórií s titulkami a tagmi na zvolenej dátovej množine, kde bola úspešnejšia kategorizácia z tagov. Následne bol otestovaný model z titulek aj na väčšej množine dát, ktorý dosiahol výrazné zlepšenie a bola dosiahnutá jeho použiteľnosť na väčšej množine dát. Ďalej bol otestovaný vybraný model kategorizácie z tagov, ktorý po tréningu dosahoval v testoch úspešnosť v priemere 90%. Ku vloženým videám do archívu boli otestované odporúčané a hlavná myšlienka každého z nich bola podobná. Ďalej bolo testované vyhľadávanie pri zadaní presného tagu, neúplného a viacerých tagov a zhodnotený výsledok. Posledným krokom bolo otestovanie webovej aplikácie používateľmi a vyhodnotenie výsledkov.

Medzi prínosy práce je možné zaradiť efektívnu kategorizáciu, kedy systém podľa zadaných tagov sám navrhuje kategóriu pre video, prípadne ak kategória k danému videu nesedí je možné do databázy vložiť vlastnú a pretrénuvať model. Taktiež tu patrí efektívne odporúčanie, ktoré porovná novo pridané video s ostatnými v databáze a následne k nemu vyberie 14 najlepšie zhodujúcich sa a taktiež ho priradí k ostatným, ku ktorým sa hodí. Okrem toho prínosom je aj vyhľadávanie v tagoch podľa zadaného textu, ktoré vráti video s rovnakým alebo podobným tagom, na ktorý mohol používateľ myslieť. Prácu by bolo možné ďalej vyvíjať a otestovať na novej vlastnej vytvorenej dátovej množine relevantnej pre implementáciu riešenia do produkcie. Ďalej by bolo možné rozšíriť model o viac kategórií, ktoré sa v ňom momentálne nenachádzajú. Vyhľadávanie by bolo možné taktiež rozšíriť, a okrem vyhľadávania v tagoch implementovať aj vyhľadávanie pre názvy, prípadne titulky a zoradovať ich od najväčšej zhody pre používateľa.

POĎAKOVANIE

Táto publikácia vznikla vďaka podpore v rámci Operačného programu Integrovaná infraštruktúra pre projekt: „Výskumné centrum pre analýzu a ochranu dát - II. Etapa“, Kód ITMS2014+: 313021W479, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja.

LITERATÚRA

- [1] Peiyuan Jiang, et al. A Review of Yolo algorithm developments. *Procedia Computer Science*, 2022, 199: 1066–1073.
- [2] Andrej Karpathy et al. “Large-scale video classification with convolutional neural networks”. In: *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. 2014, s. 1725–1732.
- [3] Fabian Caba Heilbron et al. “Activitynet: A large-scale video benchmark for human activity understanding”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015, s. 961–970.
- [4] Olga Russakovsky et al. “Imagenet large scale visual recognition challenge”. In: *International journal of computer vision* 115.3 (2015), s. 211–252.
- [5] Sami Abu-El-Haija et al. “Youtube-8m: A large-scale video classification benchmark”. In: *arXiv preprint arXiv:1609.08675* (2016).
- [6] Martin Abadi et al. “TensorFlow: Large-scale machine learning on heterogeneous systems, 2015”. In: *Software available from tensorflow.org* 1.2 (2015).
- [7] Paul Covington, Jay Adams a Emre Sargin. “Deep neural networks for youtube recommendations”. In: *Proceedings of the 10th ACM conference on recommender systems*. ACM. 2016, s. 191–198.
- [8] James Davidson et al. “The YouTube video recommendation system”. In: *Proceedings of the fourth ACM conference on Recommender systems*. ACM. 2010, s. 293–296.

Aktívne učenie v analýze sentimentu

¹Kristína MACHOVÁ, ²Mário HORVÁTH

^{1,2} Katedra Kybernetiky and Umelej Inteligencie, Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská Republika

¹kristina.machova@tuke.sk, ²mario.horvath@student.tuke.sk

Abstrakt — Článok sa zameriava na aktívne učenie v analýze sentimentu v prúde dát. Zaoberá sa spracovaním prirodzeného jazyka a pracuje s textovými dátami. Cieľom je preveriť či aktívne učenie môže spresniť a uľahčiť prácu s textovými dátami. Práca sa zameriava na predspracovanie a analýzu textu v slovenskom jazyku. Základnou metódou sú Bayesové siete. Práca porovnáva efektivitu klasického strojového učenia s aktívnym učením.

Kľúčové slová — aktívne učenie, dolovanie v dátach, analýza sentimentu, strojové učenie, Bayesové siete

I. ÚVOD

Príspevok skúma možnosti používania algoritmov strojového učenia pre analýzu sentimentu v prúde dát v slovenskom jazyku. Analýza sentimentu sa bude zameriavať na dolovanie názoru, teda analýzu pozitivity alebo negativity názoru na určitú problematiku. Klasifikácia sentimentu sa v dnešnej dobe s úspechom používa pri analýze textu s rôznymi cieľmi, ako detekcia antisociálneho chovania, klamlivých správ, nenávisťného obsahu, trollingu a pod. na sociálnych sieťach. Naším cieľom je odhad pravdivostnej hodnoty pozitívneho a negatívneho sentimentu. Keďže sa pracuje s dátovými množinami veľkého objemu, ktoré nie niekedy problém manuálne anotovať, v poslednom čase sa na to používa automatická metóda založená na aktívnom učení. Analýza sentimentu [1] je jednou z metód analýzy textu, ktorá nám slúži k tomu aby sme mohli analyzovať ľudský sentiment vo forme názoru.

Podľa amerického prieskumu [2] len za rok 2013 bežný Američan spotreboval 8,75 zettabajtov čo je 8,75 milióna gigabajtov. Tieto obrovské objemy dát môžu poskytnúť veľmi užitočné informácie. Využiť to môžeme na predvídanie volebných výsledkov, získanie sumarizovaného názoru na kvalitu produktu, čo sa oplatí a naopak čo sa neoplatí obchodníkom ponúkať na trhu.

II. ANALÝZA SENTIMENTU

Analýza sentimentu sa zameriava na rozpoznávanie polarity textu. Rozlišujeme dva hlavné druhy sentimentu, ktoré môžeme analyzovať. Prvý sa zameriava na to či je prirodzený jazyk objektívny alebo subjektívny. Teda, či sa jedná o fakt alebo je text názorovo zafarbený a teda či obsahuje sentiment. Objektívny neobsahuje žiadny sentiment a naopak subjektívny vyjadruje svoj osobný názor. Druhý typ sa označuje ako detekcia polarity, ktorá zisťuje či text obsahuje negatívny alebo kladný sentiment [3].

Základné druhy sentimentu sú: pozitívny, negatívny, neutrálny a bipolárny. Problémom je určovať neutrálny a bipolárny sentiment. Istým zjednodušením je považovať neutrálny a bipolárny názor za objektívny. Väčšinou sa sústreďujeme na kladný alebo záporný sentiment na preddefinovanej škále, najčastejšie od -5 do 5. Základné metódy používané na detekciu sentimentu sú:

- metódy založené na lexikóne
- metódy založené na korpuse
- metódy strojového učenia
- metódy založené na znalostiach
- hybridné metódy.

Taktiež rozlišujeme rôzne úrovne analýzy sentimentu:

- úroveň dokumentu
- úroveň vety
- úroveň slova
- úroveň aspektu a entity.

Aspektová analýza sentimentu zvlášť analyzuje polaritu hodnotenia každého aspektu hodnoteného objektu.

III. POUŽITÉ METÓDY

A. Pravdepodobnostný prístup ku strojovému učeniu

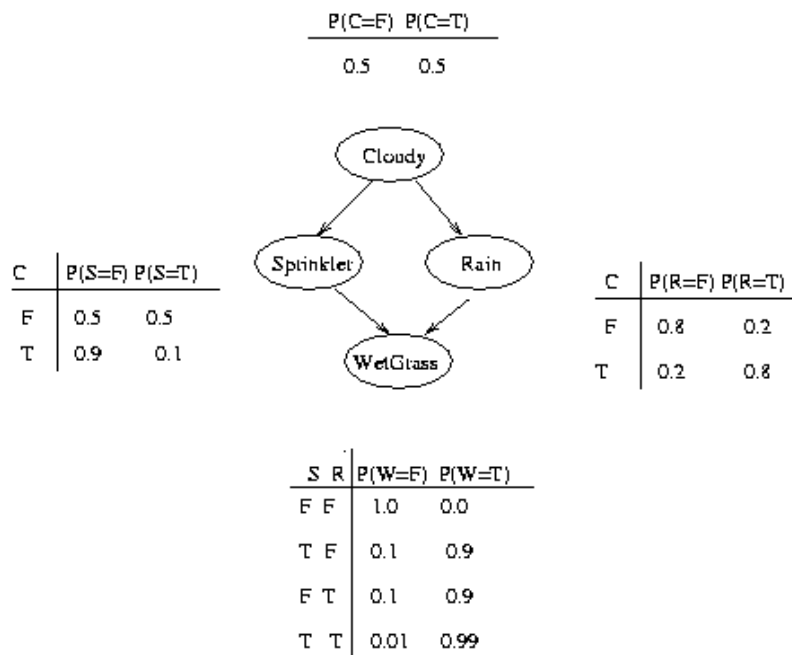
Využívané sú techniky, ktoré klasifikujú stanovené triedy pomocou ich pravdepodobnosti pre analyzované texty. Každá trieda, nájdená znalosť má určenú pravdepodobnosť. Cieľom je nájsť takého modelu, ktorý bude opisovať premenné, ktoré v istých znakoch korelujú s triedami. Tento model má dve úrovne a to štruktúrálnu a kvantitatívnu úroveň. Štruktúrálna popisuje lokálne závislosti premenných na rozdiel od kvantitatívnej, ktorá berie do úvahy silu jednotlivých závislostí. Kvantitatívne riešenia sú väčšinou určované pomocou pravdepodobnosti alebo korelácie [4]. Príkladom pravdepodobnostného prístupu sú Baesové siete.

Thomas Bayes (1702 - 1761) bol anglický štatistik, matematik, ktorý ako prvý použil pravdepodobnosť indukívne a vytvoril pre ňu matematický základ. Jeho teória predstavuje matematický prístup, ktorý vysvetľuje akým spôsobom je potrebné modifikovať domnienky o nejakej hypotéze vo svetle nových poznatkov. Pomocou neho môžeme kombinovať nové údaje s už existujúcimi vedomosťami. Bayesova teória poskytuje praktické učenie modelov vo forme algoritmov, ako napríklad, Naivný Bayes, Bayesove siete, alebo kombinácia apriórnych pravdepodobností s pozorovanými dátami v priebehu učenia. Bayesove metódy predstavujú zlatý štandard pre hodnotenie ostatných metód strojového učenia.

B. Bayesová sieť

Bayesová sieť je definovaná ako trojica premenných V, A, P . Kde N je Bayesova sieť, V je množina premenných a A je množina orientovaných hrán, ktorá s kombináciou množiny premenných vytvára acyklický graf $G = (V/A)$. Definujeme P ako: $P = \{P(v/\pi_v): v \in V\}$.

Potom P je množina podmienených pravdepodobností všetkých ostatných premenných [5]. Obrázok 1 ilustruje takúto Bayesovu sieť. Tento obrázok znázorňuje príklad s trávnikom, ktorý môže byť mokrá, lebo na neho naprší alebo ho polejeme. Počasie vonku bude buď zamračené alebo jasné. V zamračenom dni bude vyššia pravdepodobnosť zrážok ako v jasnom. Logické je aj to, že ak neprší tak trávnik polejeme my. Potrebujeme podmienenú pravdepodobnosť polievania trávniku, ale aj toho, že je zamračené. Nech je pravdepodobnosť zamračeného dňa 50%. Zo zadania nám vyplývajú štyri logické premenné: C-oblačno, S- zalievanie, R – dážď a G – tráva mokrá. Tráva mokrá môže byť true alebo false. Na obrázku 1 je zobrazená tabuľka, ktorá obsahuje udalosti, ktoré mohli nastať. Jednou z možností je pravdepodobnosť, že tráva bude mokrá, za predpokladu, že sme polievali a nepršalo. To znamená, že pravdepodobnosť $P(M = 1 | S = 1, R = 0) = 0.9$.



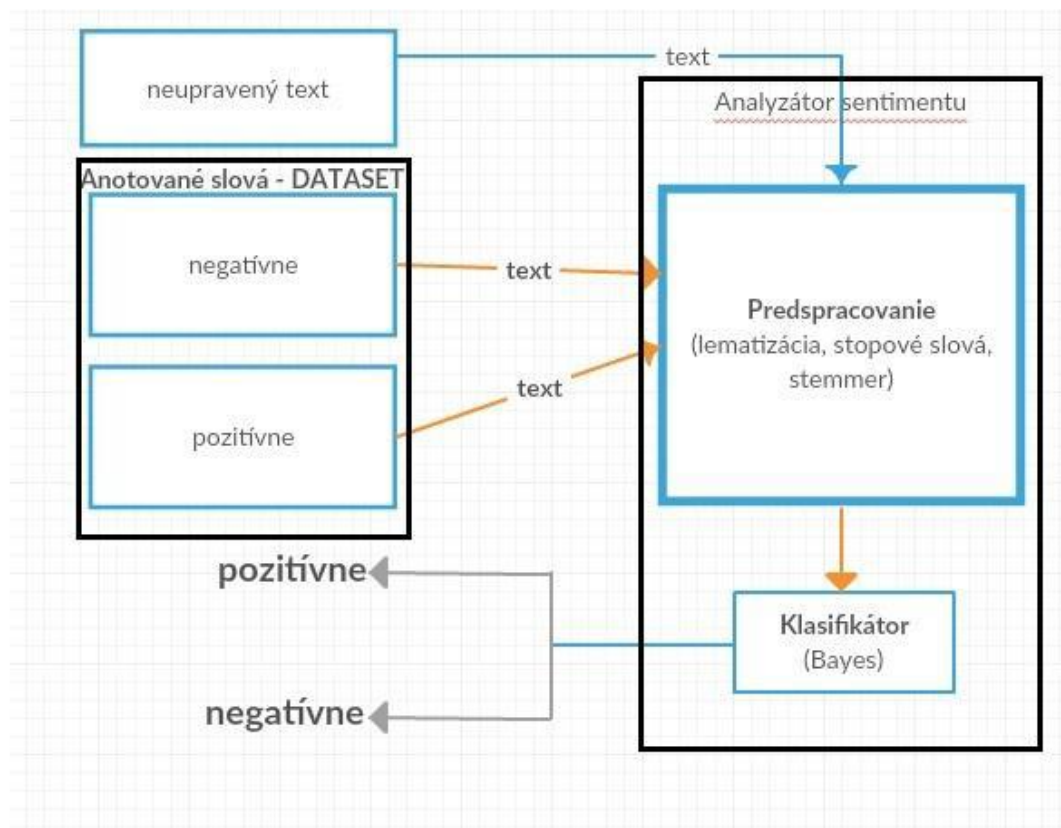
Obr. 1 Ilustrácia Bayesovej siete.

C. Aktívne učenie

Hlavným cieľom strojového učenia je určiť všeobecné vzory z obmedzeného množstva údajov. Strojové učenie môže byť kontrolované (dozorované) alebo nekontrolované. V aktívnom učení sa tieto dva prístupy kombinujú. Kontrolované učenie predpokladá anotované dáta, keď ku každému objektu je priradená trieda. Tento spôsob sa často využíva hlavne pri textových údajoch, ktoré majú veľký rozsah a je problém ich celé anotovať. V takom prípade anotujeme iba časť tréningovej množiny a na nej naučíme model strojovým učení, ktorý potom použijeme na anotovanie zvyšnej časti dátovej množiny. Čím väčšia je časť množiny, z ktorej učíme anotačný model, tým je tento model presnejší [6].

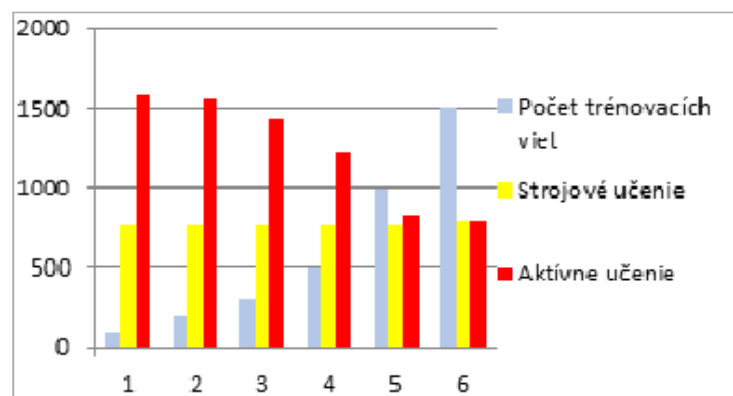
IV. NÁVRH RIEŠENIA

Metodológia nášho prístupu je ilustrovaná na Obrázku 2. Dátové množiny sme najprv predspracovali tak, že sme odstránili stopové slová a interpunkčné znamienka, vykonali sme lematizáciu, upravili tokeny na základný tvar použitím stemera a redukovali lexikálny profil. Potom sme naučili Baesovú sieť.



Obr. 2 Metodológia prístupu ku analýze sentimentu pomocou aktívneho učenia a Bayesovej siete.

Naše riešenie založené na aktívnom učení sme porovnali s klasickým strojovým učení a výsledok je ilustrovaný na Obrázku 3.



Obr. 3 Porovnanie klasického strojového učenia a aktívnym učení.

Ďalší Obrázok 4, znázorňuje použitie našej aplikácie pri analýze vety „*som spokojná, šikovný do ruky, mám ho pár dní a nemám mu čo vytknúť*“. Tento príspevok je správne klasifikovaný ako pozitívny.



Obr. 4 Ilustrácia fungovania aplikácie založenej na Bayesovej sieti.

V. ZÁVER

V práci je navrhnutý prístup ku analýze sentimentu s kladnou a zápornou polaritou, založený na Baesových sieťach a aktívnom učení. Príspevok hodnotí úspešnosť aktívneho učenia v porovnaní s klasickým strojovým učením. Je uvedený aj príklad použitia navrhnutej aplikácie.

Aktívne učenie má určite veľký význam pre uľahčenie spracovania veľkých dát a textových súborov, ktoré v dnešnej dobe pribúdajú astronomickou rýchlosťou. Jeho využiteľnosť pre firmy a marketing má veľký význam. Sledovaním takýchto dát môžu napredovať a úspešne zisťovať aktuálne trendy a potrebu trhu.

ACKNOWLEDGMENT

Táto práca vznikla s podporou Slovenskej grantovej agentúry Ministerstva školstva a SAV Slovenskej republiky v rámci projektu VEGA 1/0685/21.

REFERENCES

- [1] Pang, B., Lee, L., „Opinion mining and sentiment analysis. Foundations and trends in information retrieval,“ vol. 2, no. 2, 2008, 1–35.
- [2] How much Media? Report on American Consumers, Institute for Communications Technology Management, Stránka dostupná na internete [12.05.2023] na : http://ucsdnews.ucsd.edu/pressrelease/u.s._media_consumption_to_rise_to_15.5_hours_a_day_per_person_by_2015.
- [3] Veselovská, K., „Sentence-level sentiment analysis in czech,“ In Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics, ACM, 2012.
- [4] Heres, J., Stijl, R., and Reinders, F., „Machine Learning Methods for the Health-Indexing and Ranking of Underground Distribution Cables and Joints B. Smith, “CIRED 2016, Helsinki, Fínsko, DOI: 10.1049/cp.2016.0751.
- [5] Scutari, M., Denis J.B. „Bayesian Networks,“ 2nd Edition, Routledge publ., 2021, 274 ps., ISBN 978-0367366513.
- [6] Riccardi, G., Hakkani-Tur, D. „Active learning: Theory and applications to automatic speech recognition,“ IEEE Transactions on Speech and Audio Processing, vol. 13, no. 4, 2005, 504-511, DOI:10.1109/TSA.2005.848882.

Numerická optimalizácia inšpirovaná vlčími svorkami

¹Ján SKYBJAK, ²Marián MACH

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

¹jan.skybjak@student.tuke.sk, ²marian.mach@tuke.sk

Abstrakt — Článok sa zameriava na problematiku riešenia úloh numerickej optimalizácie pomocou metaheuristických algoritmov, ktoré sú inšpirované prírodnými fenoménmi. Pozornosť je zameraná na chovanie svorky vlkov počas lovu koristi. Článok sa zaoberá modelovaním tohto lovu v podobe algoritmu GWO, pričom okrem základnej verzie algoritmu uvádza aj dve jeho modifikácie. V rámci experimentovania bolo realizovaných niekoľko experimentov s uvedenými variantmi tohto algoritmu v rámci optimalizácie množiny zvolených testovacích funkcií.

Kľúčové slová — prírodne inšpirovaný algoritmus, GWO, metaheuristika, numerická optimalizácia, priestor prehľadávania

I. ÚVOD

Mnohé prírodné fenomény (napr. prirodzený výber a prežitie, rôzne vzory chovania, spôsoby komunikácie organizmov) sa stali hlavnou ideou v pozadí viacerých metaheuristických algoritmov, ktoré súhrnne označujeme ako prírodne inšpirované algoritmy. Tieto algoritmy sa inšpirujú vynálezmi prírody pre účely hľadania riešení optimalizačných úloh pomocou prehľadávania veľkého množstva kandidátov na riešenie optimalizačného problému, ktorí spolu vytvárajú priestor prehľadávania. Jedným z takýchto fenoménov je chovanie svorky vlkov pri love koristi, založené na stratégii obkľúčovania koristi. Reflexia tohto princípu je základom algoritmu GWO (Grey Wolf Optimizer) [1].

Tento článok sa zameriava na algoritmus GWO. Jeho štruktúra je nasledovná – v druhej sekcii sa venuje popisu metódy, pričom pozornosť je venovaná základnej verzii algoritmu a vybraným rozšíreniam algoritmu. V tretej sekcii je uvedená séria experimentov s danými verziami algoritmu na úlohe hľadania riešenia globálneho minima skupiny vybraných optimalizačných funkcií.

II. ALGORITMUS VLČEJ SVORKY

A. Prírodná inšpirácia

Vlk sivý (alebo vlk dravý či vlk obyčajný, lat. *Canis lupus*) je rozšíreným druhom cicavca z čeľade psovitých žijúcim v Eurázii ale aj v Severnej Amerike. Žije vo svorkách majúcich bežne 5 až 12 jedincov. Inšpiratívnymi sú dva prejavy života vlkov – hierarchia jedincov vo svorke a spôsob lovu koristi [2].

Vo vlčej svorke je sociálna hierarchia – od alfy po omegu. Najvyššie postavenie má alfa (α) pár (samec a samica) ktorý vedie svorku, má mláďatá a žerie prvý. Pod nimi je beta (β) úroveň (môže byť samec alebo samica), ktorý pomáha alfa páru v organizácii aktivít svorky. Nasledujú tzv. delta (δ) jedince, medzi ktoré sa radia prieskumníci, strážcovia, lovci alebo staré jedince – patria tu všetky jedince, ktoré nie sú alfou, betou ani omegou. Posledným typom vlka je omega (ω). Ostatné vlky sú mu dominantné, jeho úlohou je pomáhať avšak vzhľadom na svoje postavenie žerie posledný.

Ďalším zaujímavým spôsobom chovania vlkov je spôsob ich lovu. Nelovia individuálne ale spoločne, pričom postupujú koordinovane. Ich spôsob lovu pozostáva z troch navzájom odlišných fáz:

- Stopovanie, naháňanie a približovanie sa ku koristi
- Prenasledovanie, obkľúčenie a trápenie koristi až pokiaľ ju nezastavia
- Útok na korisť

B. Základný algoritmus vlčej svorky

V algoritme [2] vlci budú reprezentovať kandidátov z priestoru prehľadávania, pričom sa zachovávajú všetky štyri hierarchické úrovne avšak počet vlkov v rámci jednotlivých úrovní bude odlišný od prírodnej predlohy – najlepší kandidát bude α jedincom, druhý najlepší bude zase β jedincom, δ bude tretím najlepším kandidátom a všetci ostatní kandidáti budú reprezentovať ω jedincom. Každý vlk bude reprezentovaný ako vektor d reálnych čísel (bod v d -rozmernom priestore, kde d je počet nezávislých premenných optimalizovanej funkcie). Samotná optimalizácia bude vedená kandidátmi α , β a δ , zatiaľ čo ostatní kandidáti sa im budú prispôsobovať.

Pri obkľučovaní koristi vlk mení svoju pozíciu v okolí koristi, pričom sa presúva z jednej pozície na inú pozíciu. Ak vektor $X_p(t)$ bude reprezentovať pozíciu koristi v čase t a vektor $X_w(t)$ zase pozíciu vlka v danom čase, tak presun vlka na inú pozíciu v okolí koristi je možné vyjadriť vzťahom

$$X_w(t+1) = X_p(t) - A \blacksquare D \quad (1)$$

kde D je vektor vzdialenosti medzi vlkom a korisťou, A je vektor konštánt a operácia \blacksquare reprezentuje násobenie zodpovedajúcich zložiek oboch vektorov. Pre získanie variability možných pozícií vlka má vektor konštánt A náhodný charakter

$$A = 2aR_1 - a \quad (2)$$

kde R_1 je náhodne generovaný vektor so zložkami z intervalu $[0,1]$ a a je hodnota z intervalu $[0,2]$. Vlka takto môže získať náhodnú polohu v oblasti okolo koristi, čo je ilustrované na obr. 1, kde vlk je reprezentovaný červeným bodom zatiaľ čo pozícia koristi žltým bodom (sivé body reprezentujú možné pozície vlka po presune pri obkľučovaní koristi).



Obr. 1 Možné nové pozície vlka v okolí koristi.

Každá zložka vektora A je náhodne vybraná z intervalu $[-a,a]$. V prípade, že pre jeho zložku A_i platí $A_i > 1$, vlk sa môže v danej dimenzii ku koristi približovať ale aj vzdiaľovať (obrázok vľavo, $a=1.3$). Naopak, v prípade $A_i < 1$ sa ku koristi iba približuje (obrázok vpravo, $a=0.8$). Pre lepšie reflektovanie správania vlka pri love (postupne znižuje priestor okolo koristi, v ktorom sa pohybuje, až po záverečný útok na korisť) je hodnota a lineárne znižovaná z hodnoty 2 na začiatku až na nulu podľa

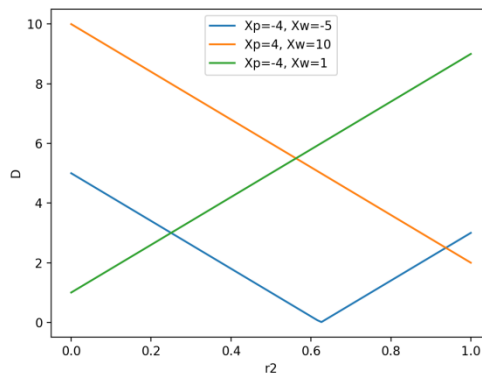
$$a = 2 - 2 \frac{t}{max} \quad (3)$$

kde max je maximálne dosiahnuteľný čas, keď algoritmus končí svoje hľadanie. Polovicu času vyhradeného hľadaniu sa vlk môže aj vzdiaľovať od koristi (s postupujúcim časom stále menej a menej) a polovicu času sa môže iba približovať. Svoj pohyb záverečne končí presne na pozícii koristi.

Navrhnutý mechanizmus pohybu vlka umožňuje exploráciu priestoru v okolí koristi, avšak vďaka uniforme oblasti možných pozícií má algoritmus tendenciu stagnácie v lokálnom okolí, z ktorého sa nevie vymaniť. Pre zlepšenie schopností preskúmať okolie bol výpočet vzdialenosti D vlka a koristi randomizovaný – výpočet presnej vzdialenosti bol nahradený vzťahom

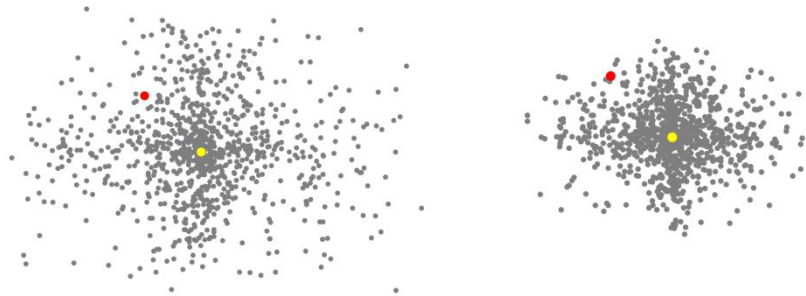
$$D = | C \blacksquare X_p - X_w | \quad (4)$$

kde $C = 2R_2$ je vektor náhodných konštánt, pričom R_2 je náhodne generovaný vektor so zložkami z intervalu $[0,1]$. V prípade, že zložka C_i je rovná jednej, tak pre danú dimenziu sa uvažuje skutočná vzdialenosť medzi vlkom a korisťou. V ostatných prípadoch uvažovaná vzdialenosť sa líši od skutočnej vzdialenosti, môže byť väčšia ale aj menšia – závisí to nielen od generovaného náhodného čísla, ale aj od vzájomnej polohy vlka a koristi. Ilustrácia zmeny vzdialeností je na obr. 2.



Obr. 2 Zmeny skutočnej vzdialenosti medzi vlkom a korisťou.

Skutočné vzdialenosti sú na obrázku zobrazené pre hodnotu $r_2=0.5$. Je zrejmé, že v niektorých prípadoch zmenšenie alebo zväčšenie vzdialenosti je rovnako pravdepodobné (oranžová a zelená priamka) avšak v iných zase nie (v prípade modrej krivky častejšie bude dochádzať k jej zväčšovaniu než k zmenšovaniu). Vlk takto môže získať náhodnú polohu v oblasti okolo koristi, ktorá je väčšia a menej uniformná ako v prípade na obr. 1. To je ilustrované na obr. 3, kde vlk je opäť reprezentovaný červeným bodom zatiaľ čo pozícia koristi žltým bodom a navyše sa vlk aj korisť nachádzajú na rovnakých pozíciách ako pred tým.



Obr. 3 Možné nové pozície vlka v okolí koristi ($a=1.3$ vľavo, $a=0.8$ vpravo).

Vlk teda prenasleduje korisť, približuje sa k nej, obkľučuje ju a záverečne zaútočí. Problémom algoritmu je to, že korisť reprezentuje hľadané riešenie, ktoré nie je známe – a teda poloha koristi je tiež neznáma a je potrebné ju niečím nahradiť. Vlci α , β a δ sú najkúsenejšie vo svorke – predpokladá sa, že majú najlepšie znalosti o výskyte koristi a preto poloha koristi je nahrádzaná polohami týchto troch vlkov. Ostatné vlky menia svoju polohu vzhľadom na polohu týchto troch vlkov. Vlk ω teda bude meniť svoju polohu podľa vzťahu

$$X_w(t+1) = [(X_\alpha(t) - A_\alpha \blacksquare D_\alpha) + (X_\beta(t) - A_\beta \blacksquare D_\beta) + (X_\delta(t) - A_\delta \blacksquare D_\delta)] / 3 \quad (5)$$

teda určí svoju novú polohu voči každému z týchto troch vedúcich vlkov a jeho výsledná nová poloha bude priemerom troch určených nových polôh.

C. Modifikácie algoritmu

V počiatočnej fáze prehľadávania je algoritmus povzbudzovaný pokryť čo najväčšiu časť priestoru prehľadávania namiesto zhlukovania okolo lokálnych extrémov. Naopak v poslednej fáze by mal využívať zhromaždené informácie aby sa priblížil ku globálnemu extrému. Pôvodný algoritmus GWO to rieši lineárnou zmenou parametra a . V niektorých prípadoch takéto riešenie stále neposkytuje dostatočné množstvo náhodnosti. Preto modifikácia MGWO (Modified Grey Wolf Optimizer) [3] pre pokles hodnoty a používa kvadratickú závislosť

$$a = 2 - 2 \frac{t^2}{max^2} \quad (6)$$

kde pomer medzi exploračiou priestoru a využívaním získanej informácie viac preferuje exploračiu.

Modifikácia MVGWO (Modified Variant of Grey Wolf Optimizer) [4] zvyšuje počet hierarchických úrovní vo svorke a pridáva jedného gamma (γ) vlka – vlci ω sa orientujú v každom čase nie podľa troch ale podľa štyroch najlepších známych kandidátov.

III. VYBRANÉ EXPERIMENTY

A. Testovacie problémy

Pre experimentovanie s algoritmom [5] bolo použitých niekoľko pomerne známych funkcií, ktoré sú často využívané pre testovanie a porovnávanie optimalizačných algoritmov [6, 7].

1. Ackleyho funkcia je definovaná vzťahom

$$f(x) = -a \exp\left(-b \sqrt{\frac{1}{d} \sum_{i=1}^d x_i^2}\right) - \exp\left(\frac{1}{d} \sum_{i=1}^d \cos(cx_i)\right) + a + \exp(1) \quad (7)$$

kde pre jednotlivé konštanty boli použité typicky používané hodnoty $a = 20$, $b = 0.2$ a $c = 2\pi$. Globálny extrém s hodnotou $f_{min} = 0$ sa nachádza v počiatku súradnicového systému. Pre testovanie bolo použité $d = 2$, definičný obor každej dimenzie bol $[-32, 32]$.

2. Braninova dvojdimenziálna funkcia je definovaná vzťahom

$$f(x) = \left(x_2 - \frac{5.1}{4\pi^2} x_1^2 + \frac{5}{\pi} x_1 - 6\right)^2 + 10 \left(1 - \frac{1}{8\pi}\right) \cos x_1 + 10 \quad (8)$$

kde globálny extrém má hodnotu $f_{min} = 0.397887$. Pre testovanie bol pre každú dimenziu použitý definičný obor $[-5, 5]$.

3. Hartmanova trojdimenziálna funkcia je definovaná vzťahom

$$f(x) = \sum_{i=1}^4 c_i \exp\left(-\sum_{j=1}^3 a_{ij} (x_j - p_{ij})^2\right) \quad (9)$$

kde koeficienty a , c , p sú zadané tabuľkovým spôsobom. Globálny extrém tejto funkcie má hodnotu $f_{min} = -3.86278$. Pre testovanie bol pre každú dimenziu použitý definičný obor $[1, 3]$.

4. Styblinski-Tangova funkcia je definovaná vzťahom

$$f(x) = \frac{1}{2} \sum_{i=1}^d (x_i^4 - 16x_i^2 + 5x_i) \quad (10)$$

kde globálny extrém tejto funkcie má hodnotu $f_{min} = -39.16599d$. Pre testovanie bola použitá dvojdimenziálna verzia, pre každú dimenziu bol použitý definičný obor $[-5, 5]$.

B. Zistenie vlastností základného algoritmu GWO

Mnohé parametre algoritmu sú generované náhodne. Manuálne je potrebné nastaviť iba niekoľko málo hodnôt – vlastne iba počet jedincov svorky. V tab. 1 sú výsledky pre tri rôzne hodnoty – 4, 12 a 50, pričom v každom prípade algoritmus realizoval 30 iterácií.

Tabuľka 1
 Porovnanie výkonu GWO pre rôzne počty vlkov vo svorke (w - počet vlkov)

Funkcia	$w = 4$		$w = 12$		$w = 50$	
	Priemer	Sd. odchýlka	Priemer	Sd. odchýlka	Priemer	Sd. odchýlka
Ackley	12.67051	0.0452189	1.684387	6.597e-08	1.684387	1.129e-15
Branin	1149.917	9.880124	4.907793	5.936143	3.705573	5.968796
Hartman	-1.169834	0.2003105	-1.899647	6.752e-06	-1.899650	1.177e-06
Styblinski-Tang	0.128266	6.777867	-30.21290	6.928862	-30.21291	6.928866

Výsledky potvrdzujú, že hodnota $w = 4$ je extrémne nízka a nepostačuje pre zabezpečenie spoľahlivého hľadania. Najlepšie výsledky dosahuje prípad $w = 50$, ktorý je vlastne bežne používanou hodnotou pri populačných prehľadávacích algoritmoch. Avšak menšia hodnota $w = 12$, ktorá predstavuje menšiu výpočtovú záťaž a rýchlejšie hľadanie, sa ukazuje pri troch funkciách ako plne použiteľná – ale neplatí to pre každú funkciu, pri Braninovej funkcii jej zaostávanie je očividné.

Pomer medzi exploračiou a využívaním dostupnej informácie je ovplyvňovaný aj počiatkovou hodnotou parametra a (v čase $t = 0$ vo vzťahu (3)). Štandardne má hodnotu 2, pri jej zvýšení nová poloha vlka bude z väčšej oblasti (podpora exploračie) zatiaľ čo pri jej zmenšení nová poloha vlka bude z menšej oblasti (potlačenie exploračie). Výsledky sú v tab. 2.

Výsledky jasne ukazujú, že nie je možné nastaviť taký stupeň exploračie, aby vyhovoval všetkým funkciám. Celkovo ako najlepšie nastavenie sa javí používané $a = 2$, ktoré z tabuľky možno považovať za najúspešnejšiu hodnotu. Potlačenie exploračie pri tretej funkcii síce prinieslo zlepšenie, ale iba nepatrné, takže hodnotu 2 nemá veľmi zmysel znižovať. Avšak zvýšenie exploračie môže priniesť podstatnejšie zlepšenie (ostatne, z toho vychádza aj modifikácia MGWO).

Tabuľka 2
 Porovnanie výkonu GWO pri rôznom dôraze na exploráciu.

Funkcia	a = 1		a = 2		a = 3	
	Priemer	Sd. odchýlka	Priemer	Sd. odchýlka	Priemer	Sd. odchýlka
Ackley	1.695212	0.0171111	1.684387	6.597e-08	1.684387	6.460e-09
Branin	7.253428	7.287249	4.907793	5.936143	4.914059	6.247078
Hartman	-1.899650	1.024e-06	-1.899647	6.752e-06	-1.899642	1.469e-05
Styblinski-Tang	-29.7417	6.778033	-30.21290	6.928862	-34.45390	6.778039

C. Porovnanie algoritmov GWO, MGWO a MVGWO

Pre vzájomné porovnanie základného a modifikačných variantov algoritmu bol počet vlkov nastavený $w = 12$ a počiatková hodnota rozhodujúca o explorácii $a = 2$. Dosiagnuté výsledky sú uvedené v tab. 3.

Tabuľka 3
 Porovnanie algoritmov GWO, MGWO a MVGWO.

Funkcia	GWO		MGWO		MVGWO	
	Priemer	Sd. odchýlka	Priemer	Sd. odchýlka	Priemer	Sd. odchýlka
Ackley	1.684387	1.129e-15	1.684387	1.129e-15	1.684387	1.129e-15
Branin	3.118588	5.302819	3.118073	5.302997	3.269191	5.295810
Hartman	-1.899651	5.079e-09	-1.899651	3.021e-09	-1.899651	9.547e-10
Styblinski-Tang	-30.0211291	6.928866	-30.21011	6.928866	-29.74169	6.778038

V tabuľke možno nájsť potvrdenie, že je vhodnejšie používať modifikovanú verziu algoritmu než verziu pôvodnú, ktorá ani v jednom prípade nebola lepšia než ostatné verzie. V prípade Hartmanovej funkcie priemerné hodnoty boli rovnaké vo všetkých prípadoch, avšak smerodajná odchýlka (a tým aj dosahovaná stabilita riešenia) preferuje použitie modifikovaných verzií algoritmu.

Z hľadiska použitých modifikácií je zrejmé, že prispôbovanie stupňa explorácie má väčší vplyv než zvyšovanie počtu vodiacich vlkov, ktorých ostatné vlky nasledujú – zdá sa, že tri vodiace vlky sú dostatočným počtom pre odhad pohybu svorky smerom ku globálnemu extrému, zvyšovanie ich počtu neprineslo významné zlepšenie prehľadávania.

IV. ZÁVER

Vykonalé experimenty potvrdili, že algoritmus založený na chovaní svorky vlkov počas lovu môže byť použitý ako optimalizačný nástroj v prípade riešenia úloh numerickej optimalizácie. Výhodou je malý počet nastaviteľných parametrov, čo značne zjednodušuje praktické využitie tohto algoritmu. Bonusom je vhodné chovanie algoritmu aj pri menšom počte vlkov, čo znižuje výpočtovú záťaž a urýchľuje činnosť algoritmu. Použitie modifikácií algoritmu malo pozitívny vplyv na jeho chovanie vďaka vytváraniu rovnováhy medzi potrebným stupňom explorácie priestoru prehľadávania a využívaním získaných informácií. Algoritmus má jednoznačne potenciál v oblasti numerickej optimalizácie.

LITERATÚRA

- [1] X.S. Yang, *Nature – Inspired Optimization Algorithms*, Elsevier, 2014.
- [2] S. Mirjalili et al., Grey Wolf Optimizer, *Advances in Engineering Software*, 2014, roč. 69, 46-61.
- [3] N. Mittal, U. Singh a B.S. Sohi, Modified Grey Wolf Optimizer for Global Engineering Optimization, *Applied Computational Intelligence and Soft Computing*, 2016, 16 pages.
- [4] N. Sigh, A modified variant of grey wolf optimizer, *Scientia Iranica D*, 2020, roč. 27, č. 3, 1450-1466.
- [5] J.Skybjak, *Použitie metód inšpirovaných vlčmi svorkami pre riešenie úloh numerickej optimalizácie*, Bakalárska práca, Fakulta elektrotechniky a informatiky Technická univerzita v Košiciach, 2022.
- [6] Virtual Library of Simulation Experiments: Test Functions and Datasets. <https://www.sfu.ca/~ssurjano/optimization.html> [prístupné 13.5.2023]
- [7] Test Functions Index, http://infinity77.net/global_optimization/test_functions.html [prístupné 13.5.2023]

Cieľové kontradiktórne útoky na klasifikačné hlboké modely

¹Dmytro LAHUNOV, ²Marián MACH

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

¹dmytro.lahunov@student.tuke.sk, ²marian.mach@tuke.sk

Abstrakt — Článok sa zameriava na problematiku kontradiktórnych útokov na hlboké modely, keď pomerne malou zmenou spracovávaných vstupných dát je možné dosiahnuť zmenu rozhodnutia hlbokého modelu oproti rozhodnutiu na nemodifikovaných dátach. Článok sa zaoberá cieľovými kontradiktórnymi útokmi, pričom sa hlbšie venuje metóde JSMA (Jacobian Saliency Map Attack). V rámci experimentovania bolo realizovaných niekoľko experimentov s touto metódou pri použití konvolučných modelov pre klasifikáciu obrazov v rámci dvoch dátových množín – jednoduchšej MNIST domény a zložitejšej CIFAR-10 domény.

Kľúčové slová — hlboký model, kontradiktórna perturbácia, kontradiktórny cieľový útok, Jacobian Saliency Map Attack

I. ÚVOD

Kontradiktórne útoky sú založené na modifikácii vstupných dát – pridávaní malých perturbácií, ktoré tieto dáta zmenia určitým spôsobom. Ich cieľom je snaha prinútiť naučený model produkovať nevalidné rozhodnutia [1]. Príkladom sú zmeny jasových úrovní jednotlivých pixelov, z ktorých je vytváraný vstupný obraz. Pre človeka sú tieto zmeny nepodstatné (nerozpoznateľné alebo automaticky kompenzované), modifikovaný obraz pre neho zobrazuje tú istú informáciu ako pôvodný nemodifikovaný obraz. Pre naučený model však diferencia medzi modifikovaným a pôvodným obrazom môže byť významná a dôsledkom dokonca môže byť zmena rozhodnutia modelu odlišná od jeho predpokladaného chovania.

Algoritmy kontradiktórnych útokov sa delia na dve hlavné kategórie – cieľové a necieľové kontradiktórne útoky [2]. Necieľové útoky majú snahu „pokaziť“ rozhodnutia modelu aby produkoval odlišné výsledky pre kontradiktórne príklady voči výsledkom pre pôvodné príklady [3]. Cieľové útoky sa snažia ovplyvniť model riadeným spôsobom – nielen na získanie odlišných výsledkov pre kontradiktórne príklady ale aj na určenie, aké by tieto odlišné výsledky mali byť.

Tento článok sa zameriava na oblasť cieľových kontradiktórnych útokov v oblasti klasifikácie obrazov. Jeho štruktúra je nasledovná – v druhej kapitole sa venuje vybranej metóde vytvárania cieľových útokov a v tretej kapitole je uvedená séria experimentov s danou metódou na dvoch dátových množinách.

II. CIEĽOVÉ KONTRADIKTÓRNE ÚTOKY

Hlboký klasifikačný model sa typicky učí takým spôsobom, že učiaci proces sa snaží minimalizovať chybovú funkciu, ktorá vyjadruje rozdiel medzi skutočným výsledkom klasifikácie (triedou do ktorej by mal byť vstupný obraz klasifikovaný), a výsledkom predikovaným daným modelom. Naučený model reprezentuje funkciu F , ktorej vstup je klasifikovaný obraz a výstupom je pravdepodobnostný vektor, ktorého zložky reprezentujú pravdepodobnosti jednotlivých tried. Vstupný obraz X je klasifikovaný do triedy Y_{pred} podľa

$$Y_{pred} = \operatorname{argmax}_j F_j(X) \quad (1)$$

kde F_j reprezentuje pravdepodobnosť, že obraz X patrí do j -tej klasifikačnej triedy.

Cieľový kontradiktórny útok mení obraz X na X_{adv} takým spôsobom, aby sa predikovanou triedou Y_{pred} stala vopred zvolená cieľová trieda t namiesto skutočnej triedy (teda nie ľubovoľná trieda okrem skutočnej, o čo sa snaží necieľový kontradiktórny útok). Zmena X na X_{adv} sa označuje ako kontradiktórna perturbácia.

A. Kontradiktórna mapa významnosti

Mapy významnosti sa používajú ako vizualizačný nástroj v oblasti vysvetľovania činnosti hlbokých modelov [4]. Je však možné tieto mapy použiť aj pre konštrukciu kontradiktórnej perturbácie [5].

Uvažujme dosiahnutie $Y_{pred} = t$ takým spôsobom, že vo vstupnom obraze sa zvýšia hodnoty tých prvkov obrazu, ktorých zvýšenie podporí nárast pravdepodobnosti $F_j(X)$. Pre identifikáciu týchto prvkov v obraze je potrebné vytvoriť mapu významnosti podľa vzťahu

$$S(X, t)[i] = \begin{cases} 0 & \text{ak } \frac{\partial F_t(X)}{\partial X_i} < 0 \text{ alebo } \sum_{j \neq t} \frac{\partial F_j(X)}{\partial X_i} > 0 \\ \left(\frac{\partial F_t(X)}{\partial X_i} \right) \left| \sum_{j \neq t} \frac{\partial F_j(X)}{\partial X_i} \right| & \text{inak} \end{cases} \quad (2)$$

kde i reprezentuje i -ty prvok obrazu X .

Z uvažovania jednotlivých prvkov obrazu niektoré prvky vylúčime – budú pre nás nevýznamné a preto v mape významnosti nadobúdajú nulovú hodnotu (prvý riadok vyššie uvedeného vzťahu). Pre toto vylúčenie platí:

- Ak zvýšenie hodnoty i -teho prvku X_i má zvýšiť hodnotu F_t , tak parciálna derivácia podľa daného prvku nemôže byť záporná (prvá podmienka v prvom riadku).
- Zvýšenie hodnoty i -teho prvku X_i nemôže zvyšovať pravdepodobnosť necieľových tried – celkové pôsobenie na ostatné triedy má byť negatívne alebo neutrálne (druhá podmienka v prvom riadku).

Tie prvky, ktoré nebudú vylúčené, získajú kladné ohodnotenie dané ako súčin ich vplyvu na zvyšovanie pravdepodobnosti cieľovej triedy a celkové znižovanie pravdepodobnosti necieľových tried. Povedané inak, vysoké ohodnotenie bude mať prvok, ktorý silne zvyšuje pravdepodobnosť cieľovej triedy, silne znižuje pravdepodobnosti ostatných tried alebo má obidva účinky súčasne.

Komplementárnym prístupom je dosiahnutie $Y_{pred} = t$ takým spôsobom, že vo vstupnom obraze sa znížia hodnoty tých prvkov obrazu, ktorých zníženie podporí nárast pravdepodobnosti $F_j(X)$. V tomto prípade je možné vytvoriť mapu významnosti podľa vzťahu

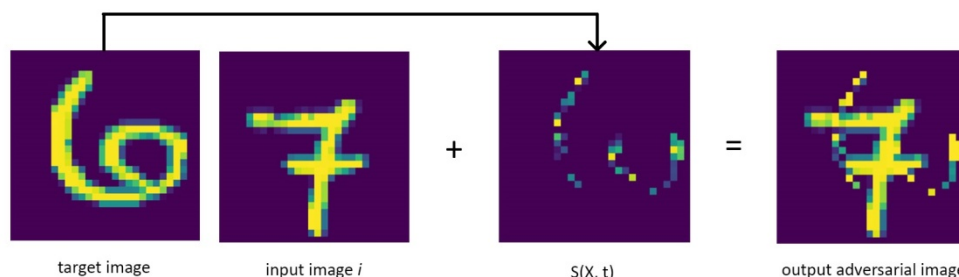
$$S(X, t)[i] = \begin{cases} 0 & \text{ak } \frac{\partial F_t(X)}{\partial X_i} > 0 \text{ alebo } \sum_{j \neq t} \frac{\partial F_j(X)}{\partial X_i} < 0 \\ \left| \frac{\partial F_t(X)}{\partial X_i} \right| \left(\sum_{j \neq t} \frac{\partial F_j(X)}{\partial X_i} \right) & \text{inak} \end{cases} \quad (3)$$

ktorý je komplementárnym k predchádzajúcemu vzťahu.

B. Metóda JSMA (Jacobian Saliency Map Attack)

Po identifikácii prvkov obrazu, ktoré sú významné z hľadiska pomýlenia klasifikátora pre klasifikovanie obrazu do cieľovej triedy, je možné realizovať kontradiktórnu perturbáciu. Iteračným spôsobom sa modifikujú vybrané prvky obrazu (jeden na iteráciu), pričom hodnota týchto prvkov sa znižuje alebo zvyšuje (v závislosti na type vytvorenej kontradiktórnej mapy významnosti). Veľkosť zvýšenia/zníženia môže byť riadená hodnotou ϵ – koeficientom perturbácie.

Počet iterácií je ohraničený (parameter algoritmu), teda nie je nutné perturbovať všetky identifikované významné prvky obrazu – zvyčajne sa perturbujú iba tie, ktoré v mape významnosti dosiahli najvyššie hodnoty. Obmedzenie počtu iterácií sleduje úmysel dosiahnuť čo najmenšie (ale dostatočné) skreslenie obrazu – príliš veľké skreslenie je ľahšie identifikovateľné človekom.



Obr. 1 Činnosť algoritmu JSMA

Činnosť algoritmu JSMA je ilustrovaná ukážkou na obr. 1. Cieľová trieda (číslica 6) je reprezentovaná obrázkom vľavo. Mapa významnosti bola očividne vytvorená prvým vyššie uvedeným spôsobom s cieľom zvyšovať hodnoty vybraných prvkov v obraze. Klasifikovaný

obraz (patriaci do triedy 7) bol perturbovaný zvyšovaním hodnoty vybraných prvkov, čím vznikol výsledný kontradiktórny obraz (vpravo).

III. VYBRANÉ EXPERIMENTY S KONTRADIKTÓRNymi ÚTOKMI

A. Testovací problém

Pre experimentovanie s metódou cieľových kontradiktórnych útokov boli zvolené dve dátové množiny obrazov:

- MNIST – množina šedotónových obrazov rozmeru 28x28 pixelov. Obrazy reprezentujú ručne písané číslice 0 až 9 (každý obraz obsahuje iba jednu z číslic), pričom sú použité rôzne formy písma a sklon číslic.
- CIFAR-10 – farebné obrazy kódované pomocou RGB farebného modelu, obrazy sú rozmeru 32x32 pixelov. Obrazy reprezentujú desať rôznych tried: kôň, pes, mačka, auto, nákladné auto, loď, lietadlo, vták, žaba, jeleň.

V oboch prípadoch sa jednalo o klasifikáciu obrazov do jednej z možných klasifikačných tried, použitých v danej dátovej množine.

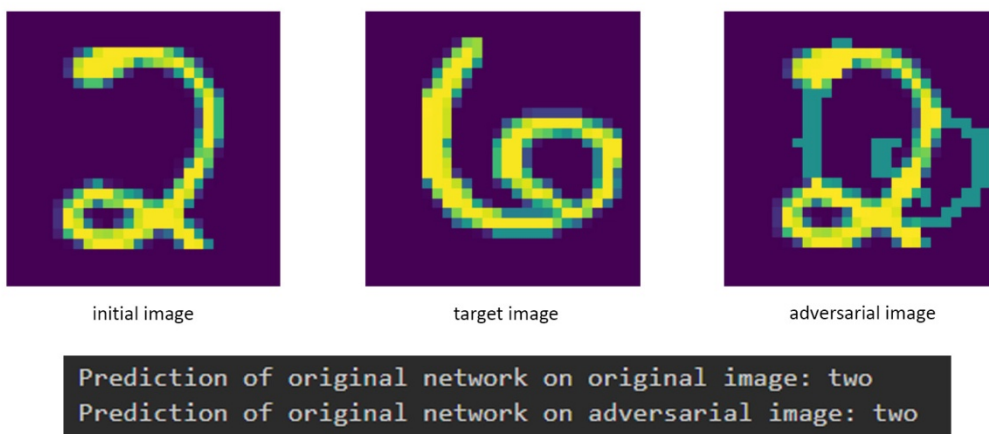
B. Hlboké modely pre klasifikáciu obrazov

Pre účely klasifikácie obrazov použitých dátových množín boli použité konvolučné siete. Pretože obtiažnosť klasifikácie je rôzna pre použitie uvažovaných dátových množín, pre každú z nich bola použitá iná štruktúra konvolučnej siete (jednoduchšia pre MNIST a zložitejšia pre CIFAR-10). Tieto štruktúry boli:

- Pre MNIST – tri konvolučné vrstvy (s 32, 64 a 64 výstupnými kanálmi), poolingová vrstva (max, jadro 2x2), tri plne prepojené vrstvy (s 32, 32 a 10 výstupnými neurónmi) doplnené dropout vrstvami pre zníženie nebezpečenstva preučenia.
- Pre CIFAR-10 – šesť konvolučných vrstiev (s 32, 32, 64, 64, 128 a 128 výstupnými kanálmi), dve poolingové vrstvy (obe max s jadrom 2x2), tri plne prepojené vrstvy (so 128, 128 a 10 výstupnými neurónmi) doplnené dropout vrstvami a vrstvami pre dávkovú normalizáciu.

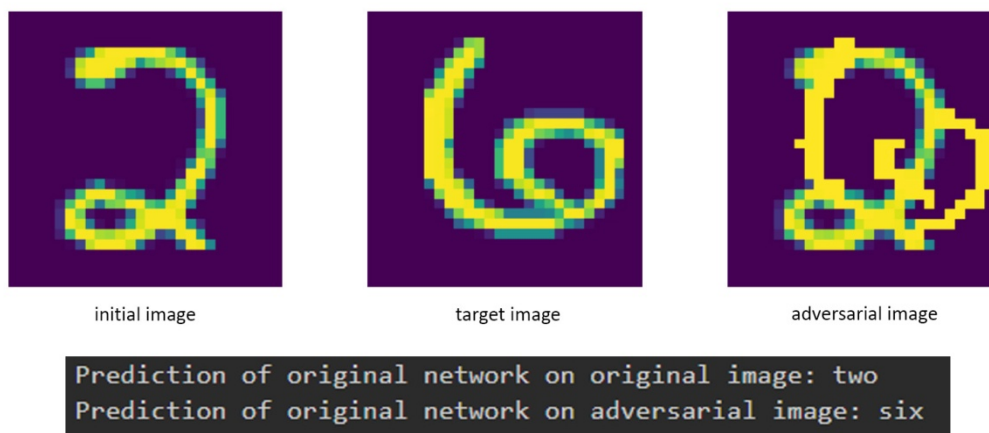
C. Experiment s algoritmom JSMA

Metóda má iba dva parametre - koeficient perturbácií ϵ , na ktorom závisí intenzita pridávaných perturbácií, a maximálny počet iterácií, počas ktorých je uskutočňované pridávanie perturbácií. Teda podstata experimentov je v skúmaní vplyvu veľkosti týchto parametrov na výsledok klasifikácie modelu [6]. Začneme s databázou MNIST. Pre $\epsilon=0.5$ a maximálny počet iterácií 90 dostaneme



Obr. 2 MNIST, $\epsilon=0.5$, $\text{iter}_{\max}=90$

Na obr. 1 vidno pôvodný, cieľový a kontradiktórny obrázok spolu s ich klasifikáciou naučeným modelom. Vidno, že model je do určitej miery robustný, pre danú hodnotu perturbácií sa nedal zmiast'. Zvýšime hodnotu na $\epsilon=1.0$ a maximálny počet iterácií na 105 a skúsime zaútočiť znova.



Obr. 3 MNIST, $\epsilon=1.0$, $\text{iter}_{\max}=105$

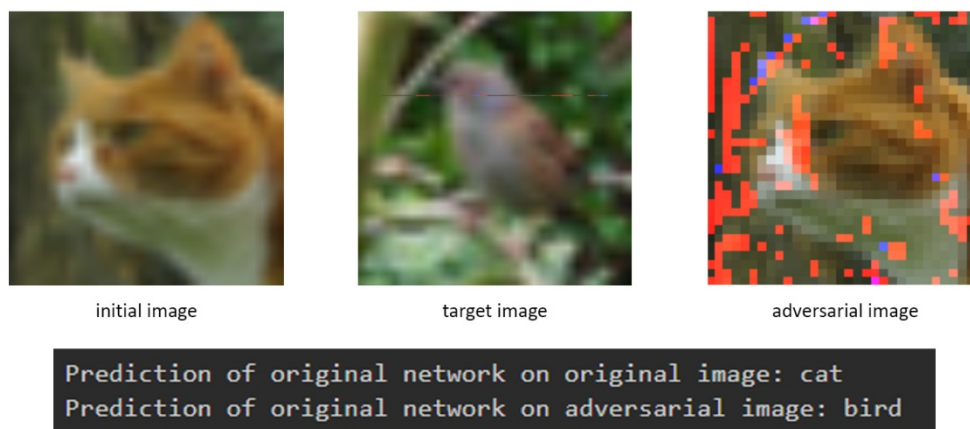
V tomto prípade sa predikcia zmenila. Intenzita a množstvo perturbácií mali vplyv na výslednú klasifikáciu – v kontradiktórnom príklade možno jasne rozoznať obrysy číslice 6. Vidno, že na rozdiel od necieľových útokov teraz nebol perturbovaný celý vstupný obraz ale perturbácie sa obmedzili iba na pridanie obrysu číslice 6.

Skúsme aplikovať daný algoritmus na farebné obrázky z datasetu CIFAR-10. Na základe predošlej skúsenosti nastavme hodnotu $\epsilon=1.0$ a maximálny počet iterácií na 100.



Obr. 4 CIFAR-10, $\epsilon=1.0$, $\text{iter}_{\max}=100$

V prípade farebných obrázkov sa model ukázal viac robustný – napriek zvýšenému koeficientu perturbácie sa nedal oklamať a klasifikoval obraz správne. Keďže intenzita perturbácií je už dostatočne vysoká, zvýšme maximálny počet iterácií na 150.



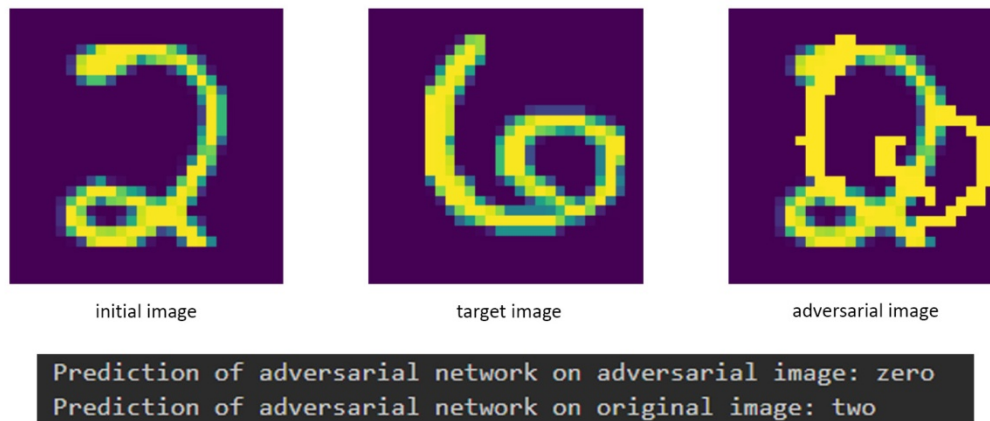
Obr. 5 CIFAR-10, $\epsilon=1.0$, $\text{iter}_{\max}=150$

Pri väčšom počte iterácií (a teda zmien v obraze) model už nepredikuje správnu originálnu triedu ale triedu cieľovú.

D. Experiment s algoritmom a chránenou sieťou

Pre účely experimentu sme pri učení siete použili metódu proaktívnej ochrany proti kontradiktórnym útokom. Prístup spočíva v tréningu siete nielen na pôvodných obrazoch ale aj na kontradiktórnych dátach (avšak klasifikovaných do správnych tried). Pomocou necieľového algoritmu FGSM [7] bola vytvorená kontradiktórna množina, v ktorej bolo 25 000 kontradiktórnych príkladov. Následne bol model pretrénovaný na pôvodných aj doplnených príkladoch.

Na pretrénovaný model sme vyskúšali zopakovať predchádzajúce experimenty. Experiment číslo 2 (MNIST, $\epsilon=1.0$, $\text{iter}_{\max}=105$) je na obr. 6.



Obr. 6 MNIST, $\epsilon=1.0$, $\text{iter}_{\max}=105$, chránená sieť

Podľa obr. 6 vidno, že použitý spôsob chránenia siete nebol úspešný. V tomto prípade síce model neklasifikuje obraz do žiadanej cieľovej triedy, avšak napriek tomu sa podarilo model pomýliť natoľko, že výsledná klasifikácia je opäť chybná.

IV. ZÁVER

Vykonané experimenty potvrdili citlivosť hlbokých klasifikačných modelov na malé zmeny vo vstupných dátach – je ich možné zmiasť malou zmenou klasifikovaných obrazov, ktorá by nerobila problém ľuďom. Cieľové útoky nemodifikujú celé obrazy ale iba ich vybrané časti a to takým spôsobom, že dokážu model zmiasť s výsledným efektom, ktorý je požadovaný útočníkom. Proti takýmto typom útokov zlyhávajú aj metódy proaktívnej ochrany modelov, čím sa cieľové kontradiktórne útoky stávajú skutočnou hrozbou. Je zrejmé, že citlivosť hlbokých modelov na kontradiktórne útoky môže byť v praxi problémom a preto je potrebné sa venovať problematike zvyšovania robustnosti týchto modelov a tým pádom aj ich ochrane pred kontradiktórnymi útokmi.

LITERATÚRA

- [1] J. Shaw, "AI and Adversarial Attacks", *Harvard Magazine*, January-February 2019.
- [2] A. Chakraborty et al., "A survey on adversarial attacks and defences", *CAAI Trans. on Intelligence Technology*, roč. 6, 2021, 25-45.
- [3] D. Lahunov, I.Čík a M. Mach, "Necieľové kontradiktórne útoky na klasifikačné hlboké modely", *Electrical Engineering and Informatics* 13, 2022, 276-281.
- [4] T.N. Mundhenk, B.Y.Chem a G. Friedland, Efficient Saliency Maps for Explainable AI. Dostupné: <https://arxiv.org/abs/1911.11293>
- [5] N. Papernot et al., The Limitation of Deep Learning in Adversarial Settings, IEEE European Symposium on Security and Privacy, Saarbruecken Germany, 2016, 372-387.
- [6] D. Lahunov, *Štúdium algoritmov vysvetliteľnej umelej inteligencie: kontradiktórne útoky*, Bakalárska práca, Fakulta elektrotechniky a informatiky Technická univerzita v Košiciach, 2021.
- [7] M. Andriushchenko a N. Flammarion, "Understanding and Improving Fast Adversarial Training", 34th Conference on Neural Information processing NeurIPS, 2020, Vancouver, Canada.

Analýza webových trendov

¹Kristína MACHOVÁ, ²Michal JANOK

^{1,2} Katedra Kybernetiky and Umelej Inteligencie, Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská Republika

¹kristina.machova@tuke.sk, ²michal.janok@student.tuke.sk

Abstrakt — Článok sa zameriava na návrh prístupu k detekcii a analýze webových trendov v čase pandémie COVID-19. Prvá časť popisuje spôsob identifikácie trendových tém v čase pandémie COVID-19. Identifikácia prebieha za pomoci techník tematického modelovania nad dátami zo siete Twitter. V ďalšej časti pomocou aplikácie aspektovo-orientovanej analýzy sentimentu, je vyhodnocovaná dynamika trendov pre zvolené témy. Sledované sú trendové pohyby v týchto témach, a následne aj vzťahy a medzi témami navzájom. Práca popisuje navrhnutý prístup k analýze trendov z online príspevkov na sociálnych sieťach, ktorý úspešne koreluje so skutočnými udalosťami. Práca obsahuje výsledky testov, zhodnotenie prístupu a jeho užitočnosti pre použitie v nových doménach. Metódy použité v práci sa ukázali byť vhodné a aj v dnešnej dobe relevantne pre úlohy analýzy trendov.

Kľúčové slová — dolovanie znalostí z textu, analýza webových trendov, modelovanie tém, aspektovo-orientovaná analýza sentimentu

I. ÚVOD

Dnes žijeme v „informačnom veku“, ktorý je definovaný ako obdobie prechodu od tradičného priemyslu k ekonomike založenej na masívnej informatizácii. Spravodajstvo, knihy, aj obchody sú dostupné online. Objemy dát však astronomicky narastajú a je pre nás stále ťažšie s nimi manipulovať a porozumieť im. Rozvíjajúce sa oblasti počítačovej vedy a spracovania prirodzeného jazyka však dokážu na tento problém nájsť odpoveď. Tieto techniky sú použiteľné aj pri analýze globálnej pandémie COVID-19. Pri každom boji s krízou je dôležité ju pochopiť, hľadať odpovede a možné východiská. Pandémia COVID-19 spôsobená novoobjaveným typom koronavírusu podľa dostupných informácií vypukla v meste Wuchan v Číne na prelome rokov 2019/2020. Situácia začala byť vážna po diagnostikovaní pozitívnych prípadov v súvislosti s trhom s morskými plodmi Huanan (Aljazeera,2020). Svetová zdravotnícka organizácia WHO vyhlásila 11.3.2020 stav globálnej pandémie. Začala sa kríza, ktorá sa dotkla každej krajiny sveta. Kríza v ktorej neprebíhal len boj o zdravie, ale aj o ekonomickú stabilitu krajín. Rastúca svetová populácia, urbanizácia a široké sociálne a obchodne aktivity medzi krajinami, zohrali dôležitú úlohu pri rýchlom šírení vírusu COVID-19 [1].

Tento článok sa zameriava na analýzu textových dát vo forme príspevkov o COVID-19, ktoré boli zozbierané zo sociálnej siete Twitter. Sociálne siete dnes plnia už aj spravodajské úlohy a poskytujú vhodný priestor na monitorovanie postojov nielen k zdravotníckym témam. Sociálne siete dnes môžeme vnímať ako informačne platformy v globálne prepojenom svete. V prezentovanom výskume sme identifikovali trendové témy, ktoré boli v sledovanom období najviac preberané v sociálnych médiách. Aplikovali sme analýzu sentimentu na sledovanie dynamického vývoja postojov populácie k týmto trendovým témam v priebehu času.

II. DETEKCIA TRENDOV

Detekcia trendov predstavuje sledovanie dlhodobých trendových pohybov, cyklických pohybov a ich variácií, sezónnych pohybov a ich variácií ako aj nepravidelných alebo náhodných pohybov. Uvažovať o trende sa dá aj z perspektívy udalostí. Ide o situáciu, ktorá sa môže ale nemusí vyvíjať rýchlo, ako napríklad pri pozorovaní reakcií na erupciu sopky alebo akt terorizmu [2]. Niekedy môže vzniknúť aj pri slede viacerých udalostí, ktoré zmenia dlhodobý trend, napríklad finančná kríza, ktorá začala postupne v roku 2008 a nebola spojená iba s jednou konkrétnou udalosťou. Išlo skôr o sériu udalostí, ktoré so sebou priniesli aj sériu následkov. V našom prípade sme predpokladali, že práve na základe rozšírenia vírusu COVID-19 mohli vo webovom prostredí vzniknúť nové trendové témy. Zamerali sme sa na ich identifikáciu a analýzu.

Detekcia vznikajúcich trendov (ETD - Emerging Trend Detection) patrí do oblasti dolovania znalostí z textov. Ide o detegovanie tematických oblastí voči ktorým rastie záujem, alebo stúpa ich užitočnosť v priebehu času. Vstupom je kolekcia textových dát a výstupom sú identifikované

témy, ktoré sú v rámci korpusu nové alebo mimoriadne dôležité. Jednou zo základných úloh pre ETD je hľadanie novo-objavujúcich sa výskumných trendov v množine vedeckých článkov. Užitočnosť pre výskumníkov je v nájdení tém, ktoré sú nové a relevantné pre ich výskum. Fyzická kontrola všetkých dostupných článkov v doméne by bola časovo náročná, alebo dokonca prakticky nemožná. Súčasná aplikácia v ETD sú rozdeľované do dvoch kategórií na plne automatické a polo-automatické systémy. Výsledným produktom práce týchto systémov sú zoznamy vznikajúcich tém spolu vizualizáciou vývoja témy v určitom období.

III. MODELOVANIE TÉM

V rámci spracovania textových dokumentov, pod modelovaním tém rozumieme úlohu, ktorej cieľom je odhaliť témy, ktorými sa dokument zaoberá. Deje sa to analýzou slovníka textu, teda hľadaním kľúčových slov textu alebo kľúčových slovných spojení. Napríklad, dokument s vysokým výskytom slov „mačky“ a „psy“, sa pravdepodobne zaoberá témou „zvieratá“, kým na druhej strane, pri výskyte slov „kôň“ a „jazdecký“ je dokument z určitej časti o „zvieratách“, ale v určitej miere je aj o „športe“.

Od doby objavenia prvých techník pre modelovanie tém v dolovaní znalosti sa modelovanie tém využíva v rôznych prístupoch v spracovaní textu, vrátane klasifikácie, kategorizácie, sumarizácie a segmentácie dokumentov. Medzi zriedkavejšie aplikácie patri oblasť počítačového videnia, populačnej genetiky a sociálnych sietí. Modelovanie tém tiež našlo uplatnenie v spoločenských vedách alebo v bio informatike [3]. Medzi najpopulárnejšie techniky modelovania tém patria Latentná sémantická indexácia (LSI) a Latentná Dirichletova alokácia (LDA). V našom výskume sme pracovali s LDA [4].

LDA funguje na základe pravdepodobnostného modelu Bayesovho typu, ktorý je generovaný nad korpusom dokumentov. Patri medzi najčastejšie používané metódy NLP na modelovanie tém. Pri získavaní informácií a z textu, LDA vychádza z predpokladu, že každý dokument vzniká generatívnym procesom. Znamená to, že dokument je zmesou tém a každá téma je mixom slov. Predpokladáme, že dokumenty pozostávajú zo slov, ktoré sa dajú použiť na identifikáciu tém. Príslušnosť k témam je udávaná ako pravdepodobnosť výskytu slov spojených s danou témou v danom dokumente. Každá téma je definovaná ako multinominalna distribúcia pravdepodobnosti nad slovníkom, ktorá je založená na Dirichletovom rozdelení [5].

IV. ASPEKTOVO-ORIENTOVANÁ ANALÝZA SENTIMENTU

Analýza sentimentu sa zameriava na rozpoznávanie polarít textu. Rozlišujeme dva hlavné druhy sentimentu, ktoré môžeme analyzovať. Prvý sa zameriava na to či je prirodzený jazyk objektívny alebo subjektívny. Teda, či sa jedná o fakt alebo je text názorovo zafarbený a teda či obsahuje sentiment. Objektívny neobsahuje žiadny sentiment a naopak subjektívny vyjadruje svoj osobný názor. Druhý typ sa označuje ako detekcia polarít, ktorá zisťuje či text obsahuje negatívny alebo kladný sentiment [6].

Základné druhy sentimentu sú: pozitívny, negatívny, neutrálny a bipolárny, v ktorom je zastúpený aj pozitívny aj negatívny sentiment. Problémom je určovať neutrálny a bipolárny sentiment. Istým zjednodušením je považovať neutrálny a bipolárny názor za objektívny. Väčšinou sa sústreďujeme na kladný alebo záporný sentiment na preddefinovanej škále, najčastejšie od -5 do 5. Základné metódy používané na detekciu sentimentu sú:

- metódy založené na lexikóne
- metódy založené na korpuse
- metódy strojového učenia
- metódy založené na znalostiach
- hybridné metódy.

My sme použili metódy založené na lexikóne, respektíve slovníkový prístup k analýze sentimentu, konkrétne aspektovej analýze sentimentu. Rozlišujeme totiž rôzne úrovne analýzy sentimentu:

- úroveň dokumentu
- úroveň vety
- úroveň slova
- úroveň aspektu a entity.

Aspektovo orientovaná analýza sentimentu (AOAS) neurčuje iba polaritu celého textu, ale zvlášť hodnotí polaritu hodnotenia každého aspektu, respektíve vlastnosti hodnoteného objektu. Samotná aspektovo - orientovaná analýza sentimentu sa využíva napríklad pri recenziách, kedy podniku nestačí zistiť len celkový postoj zákazníkov napríklad k hotelu alebo reštaurácii.

Zaujímajú ich aj ďalšie detaily ako je spokojnosť s obsluhou, spokojnosť s vybavením, spokojnosť s cenou, lokalitou a podobne. AOAS pomáha podnikom sledovať, ako sa mení postoj verejnosti v súvislosti so špecifickým atribútom, službou alebo produktom. Typická aspektovo orientovaná analýza potrebuje označenie kategórie aspektu aj s hodnotením sentimentu pre každý text [30]. Tento problém je možné riešiť aj nekontrolovanými metódami učenia, v prípade, že nemáme dáta anotované. Podobne sme postupovali aj v tejto práci.

V. NÁVRH PRÍSTUPU K DETEKCII TRENDOV

Pri riešení problému analýzy webových trendov sme použili kombináciu prístupov modelovania tém v množine dokumentov a následnej analýze sentimentu príspevkov so zameraním na jednotlivé aspekty, tvorené vygenerovanými témami. Po dôslednom predspracovaní dát sme porovnali užitočnosť tak Latetnej sémantickej indexácie ako aj Latentnej Dirichletovej alokácie z hľadiska užitočnosti pre nájdenie webových trendov a vybrali sme LDA. Pred použitím LDA je potrebné určiť niekoľko parametrov. Dôležitá je voľba optimálneho počtu tém, ktorú sme vykonali na základe výpočtu koherenčného skóre. Potom bolo uskutočnené tréningovanie modelu s konečným počtom tém. To prebiehalo za pomoci knižnice Gensim [4]. Na vizualizáciu tém bol použitý interaktívny balík PyLDAvis [4]. Na jednotlivé témy, ktoré sme považovali za aspekty trendov, sme aplikovali aspektovo - orientovanú analýzu sentimentu. Aspekty boli vygenerované témy v spojitosti s pandemiou COVID-19, pričom zdroj príspevkov bola sociálna sieť Twitter.

A. Prvá fáza návrhu riešenia

Prvá fáza návrhu prístupu ku analýze trendov je znázornená na Obrázku 1. V prvej fáze riešenia prebehlo niekoľko krokov, ktoré viedli od dát ku konkrétnym témam, ktoré boli predmetom online diskusií v sledovanom období. Kľúčové je v tejto fáze predspracovanie dát.



Obr. 1 Ilustrácia prvej fázy návrhu riešenia.

Dáta obsahovali záznamy vo forme príspevkov zo sociálnej siete Twitter v rovnakom období a boli dostupné na internetovom portáli Kaggle. Zber dát bol uskutočnený pomocou Twitter API. Zozbierané dáta obsahovali posty, ktoré majú súvis s pandemiou nového koronavírusu. Ide o časové rozmedzie od marca 2020 do konca apríla 2020. Práve to je obdobie, kedy sme sa nachádzali za vrcholom počtu s pandemiou spojených príspevkov na Twitteri, no pandémia bola stále hlavnou témou diskusií. Diskusie na internete sa už skôr zaoberali reakciami na vzniknutú krízu a jej následkami aj ekonomickými.

Dáta boli pred spracované nasledovne. Najprv boli vybraté iba príspevky v Angličtine. Ďalej boli vykonané:

- Tokenizácia a odstránenie špeciálnych znakov
- Odstránenie stopových slov
- Stemming a lematizácia
- Transformácia textovej reprezentácie na maticu Dokument-Term maticu.

Potom boli tréňované modely na generovanie tém, tak LSI ako aj LDA a LDA sa viac osvedčil. Ten sme použili pre detekciu trendov. Výsledné vygenerované témy ako aj ich interpretácia sú uvedené na Obrázku 2.

B. Druhá fáza návrhu riešenia

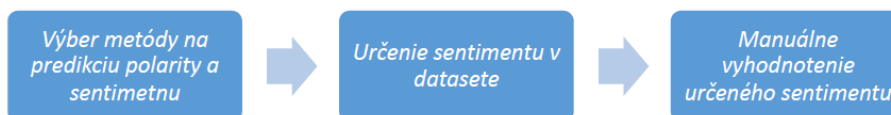
Druhá fáza návrhu prístupu ku analýze trendov je znázornená na Obrázku 3. V tejto fáze sme zvolili nástroj na určenie polaritu príspevkov a to Textblob, ktorý sme použili na určovanie sentimentu v online príspevkoch. Ďalej sme vyhodnocovali získané výsledky monitoringu sentimentu v sledovanom období tak, že sme ich porovnali s reálnymi udalosťami A, B a C. To je ilustrované na Obrázku 4.

- A. Celková miera pozitivity príspevkov začala klesať. Boris Johnson, premiér spojeného kráľovstva, bol prevezený na jednotku intenzívnej starostlivosti. Negatívny sentiment sa mohol objavovať v súvislosti k jeho postojom k pandémie do toho okamihu. Zároveň sa ukázalo, že s ochorením bojujú aj mocní predstavitelia krajín, čo práve mohlo podčiarknuť pocit zraniteľnosti. (7. apríl 2020, BBC News)

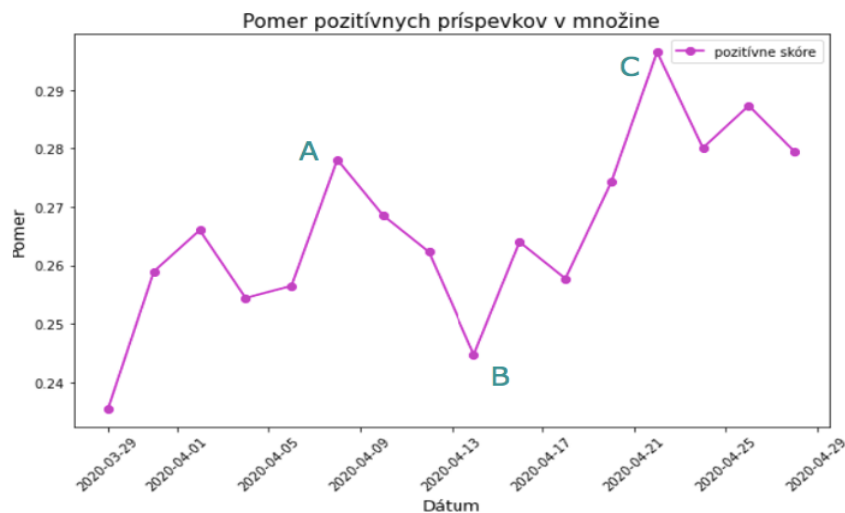
- B. Celková miera pozitivity začala stúpať. Americká vláda cez víkend potvrdila to, o čom sa už dlhšie hovorilo. Oficiálne oznámili, že príspevky vo forme stimulov v hodnote 1200 dolárov, ktoré mali dostať milióny Američanov na prekvenutie neľahkého obdobia, začnú prichádzať v nasledujúcom týždni na účty. (13.apríl 2020, CNBC)
- C. Celková miera pozitivity začala klesať. Šéf Svetovej zdravotníckej organizácia WHO, ktorá sídli vo švajčiarskej Ženeve, vo svojom vyhlásení 21. apríla varuje, že pandémia zatiaľ neukázala svoju pravú silu a najhoršie nás ešte len čaká. Bezpochyby takéto vyhlásenie šéfa renomovanej inštitúcie nemohli viesť k lepšej nálade v spoločnosti. (21.apríl 2020, The Guardian)

<u>Interpretácia</u>	<u>Kľúčové slová</u>
Ekonómika a globálna kríza	price - oil - market - crisis - economy - fall - industry - global - impact - financial
Bezpečnosť pracovníkov	grocery - work - employer - protect - worker - risk - safe - staff - thank - line
Zásoby jedla	food - need - panic - stock - demand - supply - shelf - item - chain - shortage
Výskum dopadov	consum - update - china - impact - sector - data - research - read - latest - affect
Zmena spotreb. správania (online nakupovanie)	online - spend - order - amazon - digital - change - product - delivery - wait-home
Lockdown	retail - store - close - update - future - quarantine - restaurant - open - university - stayathome
Ochrana a hygiena	sanitizer - hand - paper - mask - toilet - wash - alcohol - soap - glove - roll
Hrozby	scam - website - warn - generation - scammer - report - execute - death - insurance - offer

Obr. 2 Pomocou LDA vygenerované témy a ich interpretácia.



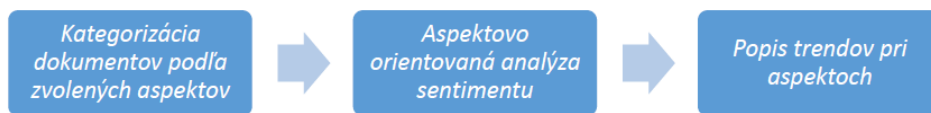
Obr. 3 Ilustrácia druhej fázy návrhu riešenia.



Obr. 4 Monitoring vývoja pozitivity online príspevkov a paralela tohto vývoja s reálnymi udalosťami A, B a C.

C. Tretia fáza návrhu riešenia

V tejto fáze sme príspevky klasifikovali podľa jednotlivých aspektov a aplikovali sme na tieto aspekty analýzu sentimentu. Podľa nej toho sme potom analyzovali vývoj trendov v konkrétnych aspektoch. To je ilustrované na Obrázku 5.

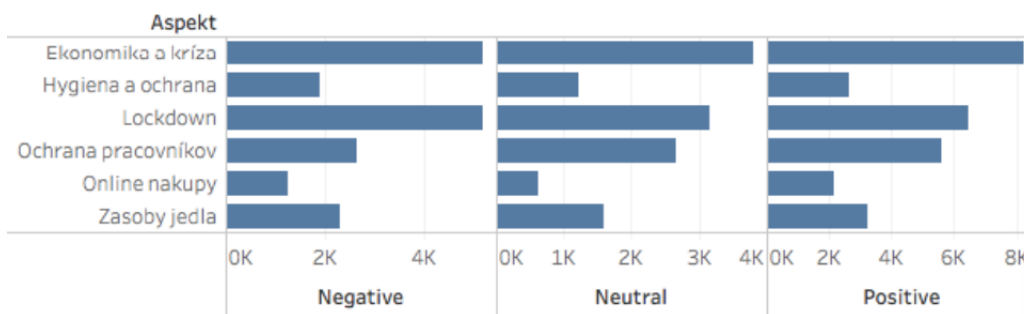


Obr. 5 Ilustrácia tretej fázy návrhu riešenia.

Jednotlivé témy z prvej fázy návrhu boli redukované na nasledovných šesť tém a tieto témy boli v ďalšom uvažované ako hlavné aspekty ktoré boli analyzované z hľadiska polarít názoru:

- Ekonomika a kríza
- Hygiena a ochrana
- Lockdown
- Ochrana pracovníkov
- Online nákupy
- Zásoby jedla

Výsledok aspektovo – orientovanej analýzy sentimentu pre týchto šesť aspektov (tém) sú uvedené na Obrázku 6.



Obr. 6 Výsledná analýza sentimentu v jednotlivých vybraných aspektoch.

VI. ZÁVER

Bol navrhnutý prístup k analýze trendov z online príspevkov na sociálnych sieťach, ktorý funguje úspešne. Metódy použité v práci sú dlhšie známe, no aj v dnešnej dobe sú užitočne a relevantne. V budúcnosti by bolo vhodné experimentovať s neurónovými sieťami, napríklad s BERT, ktorý bol vytvorený Googlom na pochopenie prirodzeného jazyka a kontextu, z textových dát. Taktiež by bolo dobré preskúmať využitie metód strojového učenia na aspektovú analýzu sentimentu.

POĎAKOVANIE

Táto práca vznikla s podporou Slovenskej grantovej agentúry Ministerstva školstva a SAV Slovenskej republiky v rámci projektu VEGA 1/0685/21.

LITERATÚRA

- [1] World Health Organization: WHO Timeline-COVID-19, [Online, April 2020]. Dostupné na: <https://www.who.int/news/item/27-04-2020-who-timeline---covid-19>.
- [2] Buntain, C.; Golbeck, J.; Liu, B.; LaFree, G. Evaluating public response to the Boston marathon bombing and other acts of terrorism through Twitter, In Proceedings of tenth AAAI Conf. Web Social Media(ICVWM), 2016.
- [3] Liu, L.; Tang, L.; Dong, W. Shaowen Yao and Wei Zhou. An overview of topic modelling and its current applications in bioinformatics. Springer Plus, 2016.
- [4] Dostupné na : [Extract Word Saliency from Gensim LDA or pyLDAvis - Stack Overflow](#), [Online, Júl 2023].
- [5] Řehůřek, R.; Sojka, P. Software Framework for Topic Modelling with Large Corpora. In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, 2010.
- [6] Veselovská, K., „Sentence-level sentiment analysis in czech,“ In Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics, ACM, 2012.

Scolopendra: Softvérové rozšírenie hexapodnej robotickej platformy

¹Samuel TITKO, ²Ján VAŠČÁK

^{1,2}Katedra Kybernetiky a Umelej Inteligencie, Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská Republika

¹samuel.titko@student.tuke.sk, ²jan.vascak@tuke.sk

Abstrakt – Cieľom tejto práce je rozšíriť softvér už existujúceho hexapodného robota Scolopendra. Zameriava sa na vytvorenie virtuálnych modelov vo vizualizácii a simulácii, grafického užívateľského rozhrania pre riadenie robota, softvérového balíka z celého projektu a architektúr prepájateľných aplikácií. Práca je rozdelená na teoretickú analýzu, popis aplikovaných technológií a praktickú implementáciu všetkých technológií do vlastného softvéru. V závere sú zhodnotenú dosiahnuté ciele celej práce.

Kľúčové slová – grafické užívateľské rozhranie, simulácia, virtuálny model, vizualizácia, Webots

I. ÚVOD

V práci [1] bola predstavená vlastná konštrukcia hexapodného robota Scolopendra a návrh jeho kinematického modelu, pomocou ktorého bolo možné zrealizovať základné typy pohybov hexapoda. Avšak pre zakomponovanie tohto typu robota do zložitejších prostredí a riešenie takých úloh, ako je napr. navigácia v prostredí s prekážkami, je potrebné navrhnuť ďalšie softvérové prostriedky. Ich zloženie a vzájomné prepojenie je predmetom tejto práce.

Softvérové rozšírenie robota Scolopendra bude zamerané na vytvorenie usporiadanej štruktúry adresárov a súborov, ktoré majú vzájomný logický súvis a budú tak vytvárať vhodné prostredie pre vývoj aplikácií. Ďalej bude rozšírená o virtuálne modely v prostredí vizualizácie a simulácie. Virtuálne modely uľahčia prácu s modelom robota každému programátorovi, ktorý je neskúsený v oblasti elektroniky, 3D tlače, obsluhy robota a jeho servisu. Vývoj a testovanie softvéru tak bude možný aj bez jeho aplikovania na fyzický model. Pri vytváraní komplexného softvéru zloženého z viacerých aplikácií, ktoré musia navzájom komunikovať, bude potrebné implementovať štandardné komunikačné rozhranie. Posledným rozšírením softvéru bude vytvorenie aplikácie užívateľského rozhrania pre ovládanie ľubovoľného modelu robota, ktorého obsluha bude intuitívna a prijateľná aj pre ľudí bez programátorských znalostí.

II. PREHL'AD IMPLEMENTOVANÝCH TECHNOLOGIÍ

V súčasnosti najpopulárnejšou technológiou na vytváranie prostredia pre vývoj robotického softvéru je middleware ROS (Robot Operating System). Má široké uplatnenie v praxi no jeho najväčšou nevýhodou je však viazanosť na Linuxové operačné systémy. Preto cieľom tejto kapitoly bude vybrať také technológie, ktoré budú nahrádzať softvér poskytovaný ROS-om ako sú RViz alebo Gazebo. Zároveň bude možné ich používať aj pod ostatnými operačnými systémami. Primárne však boli vybrané tie technológie, ktoré majú podporu operačného Windows 10 a Windows 11, ktoré nepodporujú prácu v ROS-e. Okrem toho v teoretickej analýze boli popísané grafické rozhrania, s čím súvisí aj výber vhodnej technológie na realizáciu konkrétnej aplikácie grafického ovládača. Poslednou požiadavkou na implementáciu bude zvoliť také komunikačné rozhranie, ktoré umožní komunikáciu medzi viacerými procesmi. Implementovanými technológiami sú knižnica Open3D, simulátor Webots, softvérový rámec WPF pre tvorbu grafických rozhraní a rozhranie socket.

A. Knižnica Open3D

Robotika a počítačová 3D grafika potrebovala pre vývoj knižnicu, ktorá by bola schopná rýchlo spracovať dáta formátov 3D súborov napríklad súbory *.stl, *.obj, *.pcd a manipulovať s nimi za použitia grafického procesora (GPU) v počítači. Knižnica Open3D bola vytvorená na riešenie tohto problému. Podľa [2], dovtedy neexistovala žiadna tzv. *open-source* knižnica, ktorá by bola rýchla,

jednoduchá na použitie, schopná efektívne načítať a vizualizovať 3D dáta pomocou niekoľkých riadkov kódu v súlade s modernými postupmi softvérového inžinierstva. Open3D podporuje rýchly vývoj softvéru pracujúceho s 3D dátami. Kľúčovou vlastnosťou pre zvolenie tejto knižnice je implementácia v jazyku Python a nezávislosť na operačnom systéme. Backend je naprogramovaný v jazyku C++, zatiaľ čo dátové štruktúry a funkcie v jazyku Python predstavujú frontendové rozhranie. Vývojárom aplikácie postačuje použiť jazyk Python na načítanie 3D objektov do vizualizácie a manipulovať tak s vloženými 3D objektami prostredníctvom API [2]. Samotná knižnica Open3D nedisponuje možnosťou priameho vytvárania virtuálnych modelov robotov, ale je dobrým základom na realizáciu interaktívneho vizualizačného nástroja. Tento nástroj nahradí funkcionality RVizu, ktorý je súčasťou ROS-u. Jeho úlohou bude zobrazovať virtuálny model robota a poskytovať možnosť ovládať ho prostredníctvom požiadaviek z komunikačného rozhrania.

B. Simulátor Webots

Webots je open source a multiplatformová desktopová aplikácia určená na simuláciu robotov vo virtuálnom prostredí. Simulátor je navrhnutý pre profesionálne použitie v modelovaní a simuláciách v priemysle, výskume a vzdelávaní. Prostredie simulátora poskytuje prostriedky na vytvorenie sveta a modelov robotov podľa vlastných potrieb. Okrem týchto prostriedkov vlastnej tvorby dáva možnosť vyskúšať existujúce simulácie modelov robotov, celé svety s modelmi robotov a ich predprogramovanými vzorovými ovládačmi. Zdrojový kód ovládača virtuálneho robota môže byť napísaný v jazykoch C, C++, Python, Java, MATLAB [3].

V rámci riešenia softvérového rozšírenia bude vytvorený virtuálny model robota Scolopendra spolu so základným ovládačom a modulmi pre podporu jeho programovania. Webots tak nahradí funkcionality simulátora Gazebo, ktorý je rovnako ako RViz súčasťou ROS-u.

C. Softvérový rámec WPF

Windows Presentation Foundation (WPF) je softvérový rámec užívateľského rozhrania (z angl. User Interface – UI) určený na tvorbu GUI v podobe klientskych aplikácií pre operačný systém Windows. WPF je navyše rozšírený o značkovací jazyk XAML (eXtensible Application Markup Language). XAML je postavený na značkovacom štandarde jazyka XML podobne ako značkovací jazyk HTML (HyperText Markup Language), ktorý je určený na štrukturalizáciu webových stránok. Použitie XAML na vývoj používateľských rozhraní oddeľuje grafickú stránku aplikácie od aplikačnej logiky - frontend od backendu. XAML poskytuje deklaratívny model pre tvorbu dizajnu aplikácie rozhrania [4], [5].

Softvérový rámec WPF spolu so značkovacím jazykom XAML bude použitý pri vytváraní grafického ovládača modelov robota Scolopendra. Navrhnutú a vytvorenú aplikáciu ovládača bude možné spúšťať iba pod operačným systémom Windows. Ovládač bude klientskou aplikáciou v komunikácii, a tým ju bude možné hocikedy nahradiť inou podobnou aplikáciou prispôbenou pre iný operačný systém alebo použiť takú technológiu, ktorá umožní spúšťanie grafického rozhrania pod všetkými operačnými systémami.

D. Sieťový socket

Sieťový socket je dátová štruktúra sieťového uzla – koncového zariadenia počítačovej siete. Koncovému zariadeniu slúži ako softvérový bod odosielania a prijímania dát v sieti. Štruktúru a vlastnosti socketu definuje API sieťovej architektúry. Sockety sa vytvárajú, iba ak je proces aplikácie spustený na zariadení. Termínom socket sú označované aj body komunikácie interných procesov (IPC) vrámci jedného zariadenia, ktoré často používajú rovnaké API ako sieťové sockety [6].

Sockety sa podľa prednastaveného transportného protokolu rozdeľujú na:

- 1) datagramové sockety,
- 2) streamové sockety.

Serverové aplikácie virtuálnych modelov spolu s grafickým rozhraním budú spúšťané na jednom zariadení a serverová aplikácia fyzického modelu robota bude spustená priamo na robotovi. Zároveň by bolo vhodné, aby aplikácia grafického rozhrania informovala užívateľa o tom aká aplikácia je práve spustená. Preto v realizácii komunikačného rozhrania medzi aplikáciou grafického rozhrania a serverovými aplikáciami budú implementované práve TCP sockety. Výhodou použitia socketov narozdiel od technológií vyšších internetových protokolov je možnosť spustenia v neblokovanom režime (režim kedy aplikácia neblokuje hlavný cyklus kvôli čakaniu na prichádzajúci paket), rýchlejší prenos dát a možnosť definovať si štruktúru prenášaných dát podľa potrieb riešenia.

III. NÁVRH KNIŽNICE CORE

Pre prácu so softvérovými balíkmi kĺbových robotov bola vytvorená knižnica *Core* so zdrojovými kódmi napísanými v jazyku Python.

A. Modul *Communication.py*

Prvý modul *Communication.py* je zameraný na realizáciu komunikácie medzi dvoma spustenými aplikáciami prostredníctvom rozhrania socket využívajúceho TCP protokol transportnej vrstvy. Streamové sockety boli zvolené kvôli ich spoľahlivému transportu paketov, schopnosti vzájomnej notifikácie aplikácií pri ukončení prepojenia a architektúre klient-server. Všetky dáta prenášané medzi klientskými a serverovými aplikáciami sú v štandardnom formáte JSON. Sockety sú komponentom dvoch tried vystupujúcich v komunikácii:

- 1) *JsonServer* – server,
- 2) *JsonClient* – klient.

B. Modul *Kinematics.py*

Druhým modulom je *Kinematics.py*, ktorý je optimalizovanou verziou modulu kinematiky navrhnutého v [1]. Štruktúra kinematického modelu robota je rovnako načítaná prostredníctvom súboru formátu XML, avšak s menšími zmenami názvov značiek a ich atribútov. Optimalizácia kódu bola nevyhnutná z dôvodu pomalého výpočtu inverznej kinematiky. Pomalý výpočet by znižoval obnovovaciu frekvenciu vizualizácie a simulácie a znemožňoval použitie v reálnom čase na zariadeniach s menej výkonným hardvérom [1].

C. Modul *Visualization.py*

Tretí modul je zameraný na vizualizáciu modelu robota. Využíva prvky knižnice Open3D, ktorá umožňuje načítavať súbory formátov 3D objektov, vykresliť ich do okna vizualizácie a manipulovať s nimi. Konštrukcia kĺbového robota sa skladá z veľkého množstva 3D objektov vzájomne viazaných kĺbmi alebo pevnými spojmi, ktoré sú definované štruktúrou jeho kinematického modelu. Túto funkcionality však knižnica Open3D nemá zahrnutú.

D. Modul *Package.py*

Modul *Package.py* je jadrom celej knižnice, ktorý obsahuje iba jednu triedu *Package*. Funkciou inštancie triedy *Package* je generovanie dátových štruktúr z modulov knižnice *Core* inicializovaných obsahom súborov softvérového balíka. Týmto spôsobom sú všetky dátové štruktúry knižnice spravované z jedného centrálného miesta. Zníži sa tak množstvo importovaných modulov v zdrojovom kóde vytvárajúcej aplikácie, ktorá bude inštanciu implementovať.

IV. NÁVRH SOFTVÉROVÉHO BALÍKA SCOLOPENDRA

Prostredníctvom knižnice *Core* bol vytvorený softvérový balík pre robota Scolopendra. Rozsiahle programové projekty si vyžadujú vytvorenie usporiadanej štruktúry adresárov a súborov. Softvérový balík je skupina na seba nadväzujúcich súborov a programov, ktoré majú spoločný logický súvis. Pokiaľ je balík určený pre vývoj alebo individuálne prispôsobenie, môže navyše obsahovať zdrojové kódy, vzorové súbory, spustiteľné aplikácie, dokumentáciu k zdrojovým kódom alebo fungovaniu súčastí balíka a manuály k ovládaniu aplikácií. Všetky komponenty softvérového balíka môžu byť určené na výrazne odlišné veci, ale pri spúšťaní alebo kompilovaní aplikácie sú komponenty spojené do jedného funkčného celku [7].

Balík nemusí obsahovať úplne všetky zdroje potrebné na spustenie jeho obsahu, pokiaľ existuje spôsob, ako nezahrnuté zdroje externe dodať používateľovi, poprípade sú súčasťou operačného systému (externé knižnice, externé programy, kompilátory, interpretery atď.). Tým je zabezpečená kompaktnosť – nižšia náročnosť na úložný priestor a rýchlejšia distribúcia balíka zo zariadenia na zariadenie. Na druhú stranu programátor by mal pri publikácii balíka dodať potrebnú dokumentáciu, ktorá obsahuje zoznam všetkých použitých externých zdrojov a ich verzií pri spúšťaní jeho obsahu [8].

V. KOMBINÁCIE PREPÁJATEĽNÝCH APLIKÁCIÍ

Spustené aplikácie softvérového balíka robota Scolopendra je možné prepájať v rôznych kombináciách podľa potrieb užívateľa alebo podľa výkonu hardvéru. Každá aplikácia ľubovoľného modelu, či už virtuálneho alebo fyzického robota Scolopendra, má svoj vlastný nezávislý výpočet kinematického modelu a trajektórií. Rôznym prepojením vytvorených aplikácií softvérového balíka robota Scolopendra je možné dosiahnuť 10 možných kombinácií prepojení:

- 1) Scolopendra UI a vizualizácia Open3D bez multicastu,
- 2) Scolopendra UI a simulácia Webots bez multicastu,
- 3) Scolopendra UI a fyzický robot bez multicastu,
- 4) Scolopendra UI, vizualizácia Open3D a multicast.
- 5) Scolopendra UI, simulácia Webots a multicast.
- 6) Scolopendra UI, fyzický robot a multicast.
- 7) Scolopendra UI, vizualizácia Open3D, simulácia Webots a multicast.
- 8) Scolopendra UI, vizualizácia Open3D, fyzický robot a multicast.
- 9) Scolopendra UI, simulácia Webots, fyzický robot a multicast.
- 10) Scolopendra UI, vizualizácia Open3D, simulácia Webots, fyzický robot a multicast.

A. Scolopendra UI a vizualizácia Open3D

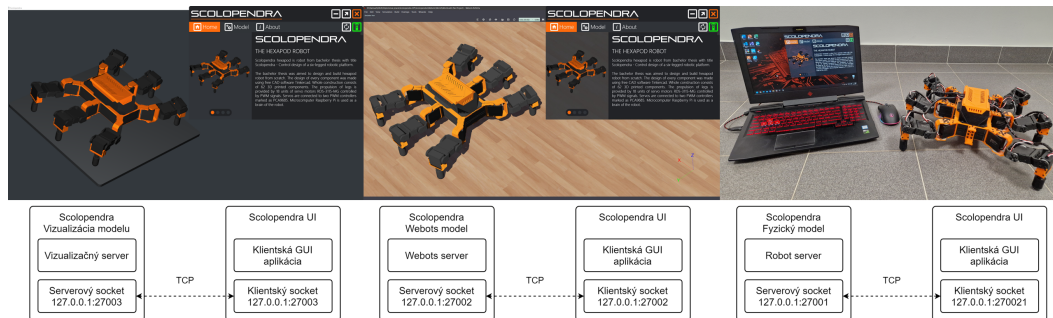
Prvá kombinácia prepojenia je medzi klientskou aplikáciou grafického užívateľského rozhrania a serverovou aplikáciou Open3D vizualizácie virtuálneho modelu robota Scolopendra. Vizualizácia je spustiteľná iba na operačnom systéme Windows. Obr. 1 napravo zobrazuje stav prepojenia týchto dvoch aplikácií.

B. Scolopendra UI a simulácia Webots

Druhá kombinácia prepojenia je medzi klientskou aplikáciou grafického užívateľského rozhrania a serverovou aplikáciou virtuálneho modelu robota Scolopendra v prostredí simulátora Webots. Obr. 1 v strede zobrazuje stav prepojenia týchto dvoch aplikácií.

C. Scolopendra UI a fyzický robot

Tretia kombinácia prepojenia je medzi klientskou aplikáciou grafického užívateľského rozhrania a serverovou aplikáciou fyzického modelu robota Scolopendra. Obr. 1 naľavo zobrazuje stav prepojenia fyzického robota a grafického rozhrania.



Obr. 1 Pohľad na prepojené serverové aplikácie s GUI

D. Scolopendra UI a multicast

Zvyšných sedem kombinácií prepojenia je medzi klientskou aplikáciou grafického užívateľského rozhrania a serverovými aplikáciami prostredníctvom aplikácie multicastu. Aplikácia multicastingu sa správa ako rozdeľovač prijatého TCP paketu od klienta. Prijatý paket je rozposlaný všetkým serverom, na ktorý sa multicasting dokázal pripojiť počas inicializácie. Multicasting odosiela klientovi späť paket od posledného pripojeného servera v poradí. Pre fungovanie tejto architektúry je nevyhnutné mať spustenú aspoň jednu serverovú aplikáciu.

VI. EXPERIMENT MERANIA VÝPOČTOVÉHO ČASU

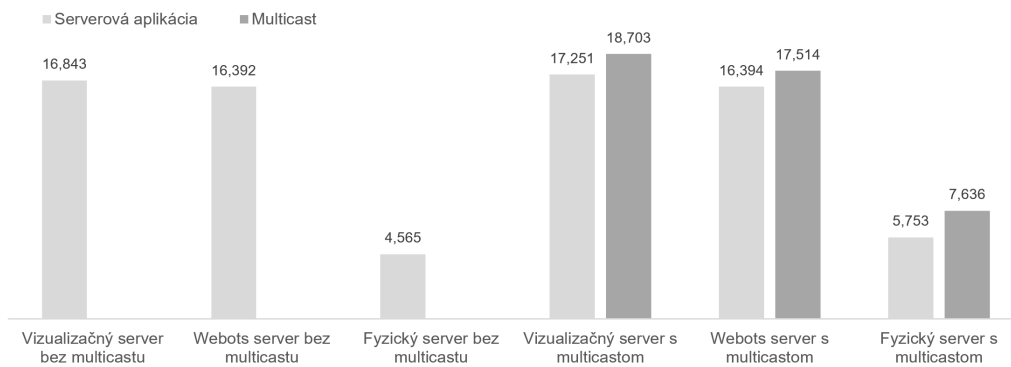
Cieľom experimentu bolo meranie času jedného výpočtu hlavnej slučky programov všetkých serverových aplikácií. Čas bol meraný pri aplikovaní translačného pohybu na model robota prostredníctvom virtuálneho joysticku z grafického rozhrania. Diaľkové ovládanie spúšťa a generovanie trajektórií a ich aplikovanie na kinematický model robota. Virtuálne modely boli spúšťané na desktopovom počítači. Hardvérové parametre testovacieho desktopového zariadenia sú:

- Centrálny procesor – AMD Ryzen 7 5700X
- Grafický procesor – NVIDIA RTX 3070
- Operačná pamäť – 32GB DDR4 3600MHz

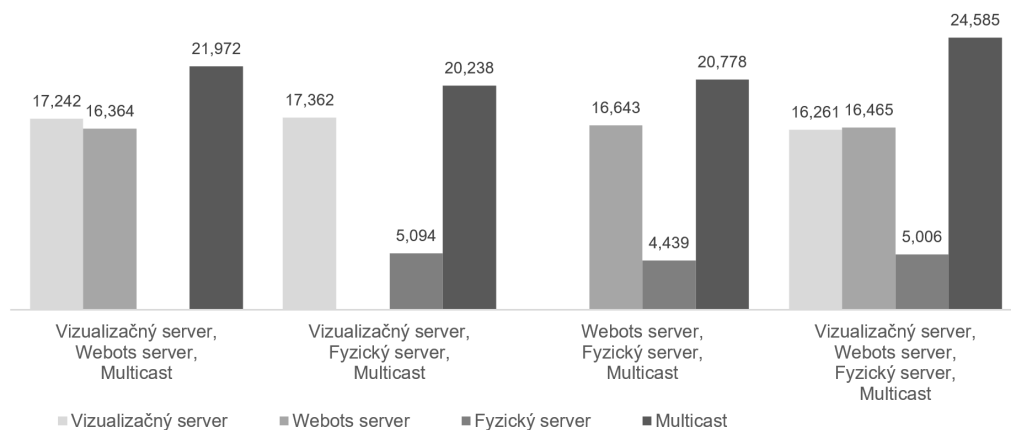
Fyzický model bol riadený jednodoskovým mikropočítačom Raspberry Pi 3B+. Riadiaca jednotka bola pripojená na počítač, ktorý mal spustenú funkčnosť Wi-Fi prístupového bodu. Meraním času jedného výpočtu sa overovalo, či generovanie trajektórií a výpočet priamej a inverznej kinematiky nadmerne nezaťažuje výpočtové jednotky. Ďalej bol meraný aj výpočtový čas multicasu. Z nameraných hodnôt bolo možné zistiť činnosť aplikácií v reálnom čase a vplyv multicasu na výpočtový čas serverových aplikácií. Nadmerná záťaž výpočtovej jednotky by mala za následok pomalú odozvu na užívateľský vstup z grafického rozhrania a nízku obnovovaciu frekvenciu aktualizácie zobrazovaných virtuálnych modelov.

A. Výstupy experimentu

Výstupom experimentu bolo desať meraní, ktoré boli rozdelené na dve skupiny. Ku každému meraniu bol vytvorený stĺpcový graf (Obr. 2 a 3) pre lepšie porovnanie nameraných hodnôt. Prvú skupinu predstavovalo šesť meraní časov jednej spustenej serverovej aplikácie pripojenej na klientskú aplikáciu bez zapnutého multicasu a so zapnutým multicastom. Zvyšné štyri merania prebiehali pri zapnutom multicaste a viacerých spustených serverových aplikáciách. Priemerný čas bol vypočítaný z 1000 nameraných vzoriek spustenej aplikácie v milisekundách.



Obr. 2 Graf časov výpočtov jednej spustenej serverovej aplikácie



Obr. 3 Graf časov výpočtov kombinácií spustených serverových aplikácií

B. Vyhodnotenie experimentu

Z prvej časti experimentu zameraného na jednu serverovú aplikáciu bolo zistené, že najnižší čas výpočtu má aplikácia fyzického modelu robota a najvyšší čas výpočtu má aplikácia vizualizácie. Desktopový počítač je oveľa výkonnejším zariadením ako jednodoskový mikropočítač Raspberry Pi 3B+, no jeho limitujúcim faktorom je neustále grafické vyobrazovanie virtuálneho modelu robota. Raspberry Pi vykonáva namiesto vyobrazovania modelu priame riadenie servo motorov, čo nemá až taký zásadný vplyv na výpočtovú záťaž až procesora. Taktiež je možné si povšimnúť, že vizualizácia Open3D je oproti simulácii Webots výpočtovo náročnejšia. Tento fakt je daný tým, že aplikácia simulátora Webots je naprogramovaná v programovacom jazyku C++ a logika aplikácie vizualizácie s využitím knižnice Open3D je naprogramovaná v jazyku Python. Interpretovaný jazyk Python je mnohonásobne pomalší ako kompilovaný jazyk C++, a tým je aj výsledná aplikácia pomalšia. Pre všetky tri aplikácie však platí, že výpočtový čas so spusteným multicastom bol o trochu vyšší ako bez neho. Dôvodom je dvojúrovňové odosielanie paketov z klientskej aplikácie do multicastu a následne z multicastu na serverovú aplikáciu.

Z druhej časti vykonaného experimentu merania zameraného na kombinácie serverových aplikácií vyplýva, že najviac zaťažujúcou kombináciou pre desktopové zariadenie je trojkombinácia všetkých aplikácií. Prejavilo sa to hlavne na multicaste, ktorý mal najvyšší čas výpočtu. Pri simultánnom ovládaní všetkých troch modelov z grafického rozhrania nebola registrovaná extrémne pomalá reakcia na užívateľský zásah v porovnaní s ostatnými kombináciami. Pri dostatočne výkonnom zariadení dokážu byť paralelne spustené všetky serverové aplikácie s grafickým rozhraním a reagovať na užívateľský vstup takmer okamžite. Pri menej výkonnom zariadení je možné znížiť záťaž tak, že sa zvolí dvojkombinácia alebo iba jedna spustená aplikácia.

VII. ZÁVER

Vytvorená knižnica *Core* v jazyku Python je vhodným základom pre prácu so softvérovými balíkmi mobilných kľbových robotov alebo kľbových robotov s pevne viazanou základňou. Knižnica zahŕňa optimalizovaný zdrojový kód modulu kinematiky z bakalárskej práce a taktiež modul vizualizácie využívajúci prvky knižnice *Open3D*, ktorým je možné vizualizovať virtuálny model kľbového robota zo súboru formátu XML. Vytvorené komunikačné rozhranie s využitím TCP socketov a obsahu balíka formatovaného v štandarde JSON je funkčné, spoľahlivé a rýchle. Pre robota Scolopendra bol prostredníctvom knižnice *Core* vytvorený softvérový balík s názvom *Scolopendra* obsahujúci súbory zdrojových kódov a inicializačné súbory pre rýchle programovanie aplikácií. Súčasťou balíka sú vytvorené serverové aplikácie virtuálnych modelov robota Scolopendra v prostredí vizualizácie a simulácie. Vizualizácia bola realizovaná prvkami vlastného modulu vizualizácie postavenom na dátových štruktúrach knižnice Open3D. Pre simuláciu virtuálneho modelu robota bol použitý voľne dostupný robotický simulátor Webots. K programovaniu virtuálneho modelu robota v simulácii bol vytvorený doplnkový balík, ktorý obsahuje modul ovládača virtuálnych servo motorov a ovládač robota. Balík obsahuje okrem iného aj plne funkčnú klientskú aplikáciu grafického užívateľského rozhrania s názvom Scolopendra UI. V závere práce experimentálne merané časy rôznych kombinácií architektúr vzájomného prepojenia grafického rozhrania a serverových aplikácií.

POĎAKOVANIE

Táto práca je finančne podporená Vedeckou grantovou agentúrou MŠVVaŠ SR a SAV prostredníctvom projektu VEGA 1/0480/22 — EDEN: EDge-Enabled inteligentné systémy.

LITERATÚRA

- [1] S. Titko, "Scolopendra – návrh riadenia šesťnohej robotickéj platformy," in *Technická univerzita v Košiciach, bakalárska práca*, 2021.
- [2] Q.-Y. Zhou, J. Park, and V. Koltun, "Open3D: A modern library for 3D data processing," *arXiv:1801.09847*, 2018.
- [3] Webots, "http://www.cyberbotics.com," open-source Mobile Robot Simulation Software. [Online]. Available: <http://www.cyberbotics.com>
- [4] L. A. MacVittie, *XAML in a Nutshell*. O'Reilly Media, Inc., 2006.
- [5] Microsoft, "What is XAML - WPF .NET," 2023, [online]. cit. 2023-03-04. [Online]. Available: <https://learn.microsoft.com/en-us/dotnet/desktop/wpf/xaml/?view=netdesktop-7.0&redirectedfrom=MSDN&viewFallbackFrom=netdesktop-5.0>
- [6] L. Kalita, "Socket programming," *International Journal of Computer Science and Information Technologies*, vol. 5, no. 3, pp. 4802–4807, 2014.
- [7] S. Schaal, "The sl simulation and real-time control software package," Citeseer, Tech. Rep., 2009. [Online]. Available: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=a729eb21ff7d95d2d247a220283c4490daf06411>
- [8] W. Khalil, A. Vijayalingam, B. Khomutenko, I. Mukhanov, P. Lemoine, and G. Ecorchard, "Opensymoro: An open-source software package for symbolic modelling of robots," in *2014 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 2014, pp. 1206–1211.

Návrh softvérového rámca pre komunikáciu V2X v oblasti robotických rojov

¹Matúš SMOLKO, ²Ján VAŠČÁK

^{1,2}Katedra Kybernetiky a Umelej Inteligencie, Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská Republika

¹matus.smolko@student.tuke.sk, ²jan.vascak@tuke.sk

Abstrakt – Táto práca prináša prehľad problematiky návrhu softvérového rámca pre komunikáciu V2X v oblasti robotických rojov s použitím dronov typu DJI RYZE Tello, taktiež obsahuje prehľad charakteristiky dronov, ich komponentov, princípy pohybu a softvérové nástroje. Súčasťou teoretického návrhu softvérového rámca sú aj overenie funkčnosti, testovanie jedincov roja dronov a experimenty so softvérovým rámcom navrhnutým pre komunikáciu s dronmi.

KLúčové slová – dron, programovací jazyk Python, rojová robotika, V2X komunikácia

I. ÚVOD

V súčasnosti sa robotika radí medzi najrýchlejšie sa rozvíjajúce oblasti v technologickom priemysle. Jedným z najzaujímavejších prístupov v robotike je využitie robotických rojov, ktoré pozostávajú z viacerých samostatných robotov, ktoré spolupracujú na dosiahnutie určitého cieľa. Pre úspešné využitie robotických rojov je kľúčová komunikácia medzi robotmi, a preto je perspektívne využiť technológiu V2X komunikácie (Vehicle-to-Everything). Softvérový rámec pre V2X komunikáciu v robotických rojoch umožňuje ich vzájomnú komunikáciu, ako aj zdieľanie informácií na ich vzájomnú kooperáciu. Táto komunikácia v robotických rojoch je založená na navrhnutom a implementovanom softvérovom rámci V2X komunikácie [1].

Tento navrhnutý softvérový rámec môže pomôcť k rozvoju oblasti V2X komunikácie pre jednotlivé drony v roji tým, že umožní vzájomnú komunikáciu a zdieľanie informácií medzi dronmi, čo môže zlepšiť ich koordináciu a efektivitu pri plnení úloh, taktiež poskytne nový pohľad do komunikácie V2X.

II. KOMUNIKÁCIA V ROBOTIKE

Každá činnosť, ktorú vykonávame, má komunikačný aspekt, ktorý zahŕňa prenos informácií, oznámení, nariadení, spoločné vytváranie a udržiavanie spojenia. S vývojom robotiky vznikajú nové možnosti využitia nových konštrukčných materiálov, komponentov, senzorických a riadiacich systémov. To má pozitívny vplyv na komunikáciu medzi ľuďmi a strojmi, ako aj na komunikáciu medzi robotmi samotnými. Tento trend vedie k vývoju systému Internetu vecí (IoT), ktorý umožňuje autonómu komunikáciu medzi fyzickými objektmi a stáva sa bežnou súčasťou každodenného života. Komunikácia v robotike môže byť synchrónna alebo asynchrónna. Synchronná komunikácia umožňuje účastníkom komunikovať v reálnom čase pomocou technológií, ako sú hlasové, textové, zvukové, alebo video prenosy. Asynchrónna komunikácia sa odohráva v rôznom čase, a reakcie sú časovo posunuté, čo umožňuje účastníkom času premyslieť si svoje reakcie pred odpoveďou. Komunikácia je dôležitou súčasťou fungovania robotov v rôznych aplikáciách a využíva pokročilé informačno-komunikačné technológie [2].

V robotike využívame rôzne druhy komunikácie, ako napríklad riadiaca komunikácia pre prenos príkazov a informácií medzi riadiacim systémom a robotom, senzorická komunikácia pre získavanie informácií o okolí prostredí a komunikácia medzi robotmi pre spoluprácu v tíme. Pre pripojenie robotov k sieti a ich vzdialené ovládanie sa využívajú aj bezdrôtové technológie ako sú Wi-fi, Bluetooth a Zigbee, ktoré sú najvýhodnejšie pre ich vyššiu spoľahlivosť, rozšírené možnosti ovládania a úspory nákladov. V súčasnosti sa technológie sústreďujú na zvyšovanie prenosových rýchlostí a na nahradenie vodičov signálu na bezdrôtový prenos pri V2X komunikácií [2].

III. ROJOVÁ ROBOTIKA

Mobilné roboty sú zaujímavé pre riešenie úloh spolupráce, pretože sú flexibilné a schopné pohybu v rôznych prostrediach. Pre komplikovanejšie úlohy je potrebné použiť robotický systém spolupráce, ktorý zahŕňa viac mobilných robotov. Riešením môže byť použitie rojovej robotiky [2].

Rojová robotika, tiež nazývaná kolektívna robotika, je nový prístup v robotike, ktorý sa zameriava na vývoj robotov, ktorí spolupracujú ako súčasť kolektívu. Cieľom tejto oblasti robotiky je umožniť robotom spolupracovať a komunikovať medzi sebou, aby mohli vykonávať zložité úlohy a riešiť problémy, ktoré by boli pre jednotlivého robota ťažké, alebo neriešiteľné. Rojová robotika sa sústreďuje na skúmanie spôsobov, ako roboti môžu spolupracovať a vytvárať tak zložité kolektívne správanie [2].

V rojovej robotike sa často využíva princíp "samoorganizácie", ktorý umožňuje robotom prirodzene sa organizovať a komunikovať s ostatnými robotmi. Tento princíp im umožňuje prispôbiť sa novým situáciám a prostrediam a tak zlepšovať svoje kolektívne správanie [2].

Pri použití rojovej robotiky sa často využívajú algoritmy, ktoré umožňujú robotom spolupracovať a komunikovať medzi sebou. Tieto algoritmy môžu byť založené na rôznych prístupoch, ako napríklad na umelých neurónových sieťach, genetických algoritmoch, evolučných algoritmoch alebo na algoritmoch strojového učenia. Cieľom použitia týchto algoritmov je dosiahnutie optimálneho riešenia pre danú úlohu. Každý z týchto algoritmov má svoje vlastnosti a výhody, ktoré umožňujú robotom spolupracovať a riešiť rôzne úlohy. Pre dosiahnutie optimálneho riešenia pre danú úlohu je dôležité zvoliť správny algoritmus a prispôbiť ho konkrétnym potrebám [2].

IV. DRON A JEHO ÚLOHY

V súčasnosti majú drony široké uplatnenie v každodennom živote a je veľmi dôležité poznať ich fungovanie, klasifikáciu, zloženie, ovládanie a využitie. V tejto časti sa venujeme základnej teórii týkajúcej sa bezpilotných leteckých systémov, ich komponentom, princípu a dynamike, softvérovým nástrojom na programovanie a opíšeme DJI RYZE Tello dron.

V súčasnosti sa termín "dron" používa na akékoľvek lietajúce, plávajúce alebo pozemné zariadenie, ktoré nie je ovládané človekom. Takéto zariadenia sú označované skratkami UAV (Unmanned Aerial Vehicle - bezpilotné lietajúce zariadenie) alebo UAS (Unmanned Aerial System - bezpilotný lietajúci systém). Existuje mnoho rôznych typov bezpilotných zariadení, ktoré majú rozličný tvar, typ pohonu a funkciu. Pojem "dron" sa v posledných rokoch stal populárnym v širšej verejnosti, aj keď takéto zariadenia boli vyvíjané a používané už dlhé roky. Postupne sa bezpilotné zariadenia a systémy zdokonaľovali a používali sa aj v civilnom sektore, kde zaznamenali úspech. Tieto zariadenia boli modifikované a prispôbované pre rôzne odvetvia a úlohy.

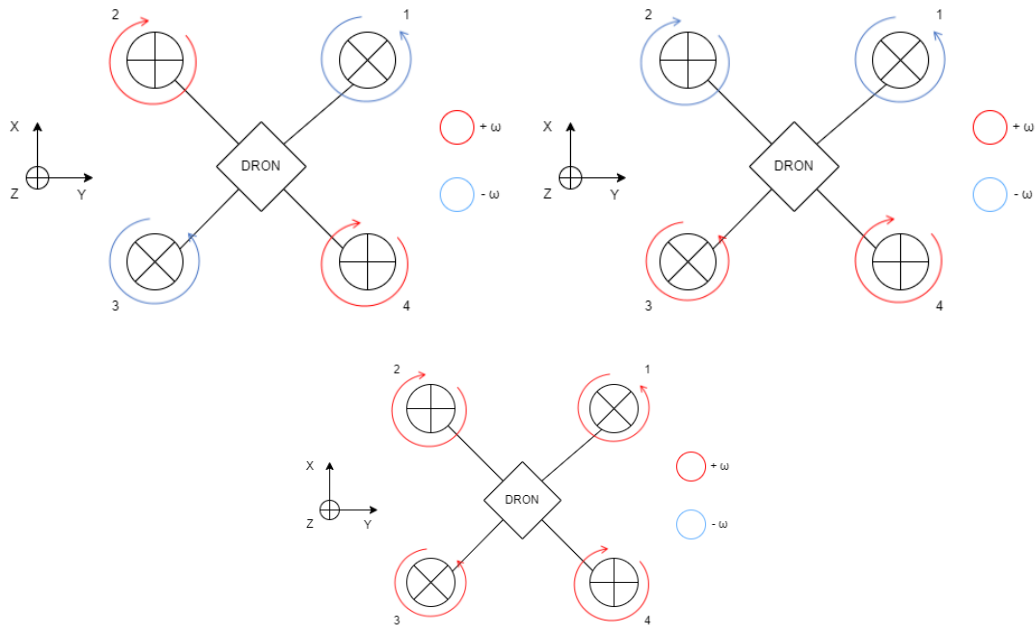
A. Princíp a dynamika pohybu drona

Multikoptéry využívajú rotory a vrtule s vertikálnou osou na vytvorenie dostatočnej vztlakovej sily pre svoj let. Všeobecne platí, že n -koptéra s menším počtom vrtúl býva obratnejšia, zatiaľ čo koptéra s viacerými vrtulami býva spoľahlivejšia. Riadiaca jednotka je ďalším dôležitým komponentom pre pohyb multikoptéry, ktorá reguluje motory, na ktorých sú umiestnené vrtule, v spolupráci so senzormi zakomponovanými v konštrukcii. Názov multikoptéry závisí na počte rotorov, ktoré používa. Následne sa zameráme na pohyb drona DJI RYZE Tello, ktorého vieme zaradiť podľa typu konštrukcie a počtu rotorov do skupiny kvadroptér.

Kvadroptéra má štyri nosné motory, ktoré sú spojené s elektromotormi, ktoré umožňujú každému rotoru poháňať sa inou uhlovou rýchlosťou. Rýchlosť rotorov sa dá regulovať pomocou zadaných príkazov pre pohyb, ako napríklad zvýšenie napätia na elektromotore, čo zvyšuje otáčky, alebo zníženie napätia, čo spôsobuje pokles otáčok. Rýchlosť rotorov ovplyvňuje dĺžku letu kvadroptéry, keďže rýchly let si vyžaduje vyššie napätie a to znižuje výdrž batérie. Kvadroptéra vykonáva priestorový pohyb zložený z kombinácie translačného a sférického pohybu. Pohyb sa popisuje pomocou kolmých osí x , y , z a orientácie, ktorá je definovaná pomocou troch Eulerových uhlov - rotácia - ϕ , nutácia - θ , precesia - ψ . Uhol ϕ vyjadruje rotáciu okolo osi X , uhol θ vyjadruje rotáciu okolo osi Y a uhol ψ vyjadruje rotáciu okolo osi Z . Pre tieto pohyby sa používajú termíny *roll*, *pitch* a *yaw*.

Základné pohyby kvadroptéry v priestore sú (viď Obr. 1):

- Translačné pohyby - zahŕňajú vznášanie, stúpanie a klesanie, ktoré zabezpečujú vertikálny pohyb kvadroptéry.
- Sférické pohyby - zahŕňajú otáčanie okolo svojej osi, čo umožňuje kvadroptére meniť smer letu.
- Kombinované pohyby - zahŕňajú sklony v rôznych smeroch, ktoré umožňujú kvadroptére meniť smer letu a prispôbovať sa okolitej situácii.



Obr. 1 Ukážky schém jednotlivých pohybov kvadroptéry v priestore.

Aplikačné softvéry sú dôležitou súčasťou dronov a zabezpečujú ich plnohodnotnú prevádzku a užitočnosť. Medzi najpoužívanejšie aplikácie pre riadenie patria:

- Tello SDK - je open-source softvér vyvinutý spoločnosťou DJI pre programovanie a ovládanie dronu Tello. Umožňuje vývojárom a nadšencom vytvárať nové funkcie a aplikácie na riadenie drona.
- TelloPy - je ďalší open-source softvér pre programovanie drona Tello, ktorý poskytuje viac funkcionalít ako Tello SDK. Taktiež umožňuje prístup k rozšírenej funkcionalite, ako napríklad aj živé streamovanie videa.
- ROS (Robot Operating System) - je platforma pre robotiku, ktorá zjednodušuje vývoj a riadenie robotov. S jeho pomocou je možné vytvárať softvér pre riadenie dronov a robotov.
- Scratch je vizuálny programovací jazyk, ktorý umožňuje jednoduché programovanie dronov a robotov bez nutnosti znalosti kódu. Poskytuje základné funkcie pre ovládanie dronov, ako napríklad vzlietnutie a pristátie.
- DroneBlocks je softvér pre programovanie dronov pomocou vizuálneho programovania.

B. Dron DJI RYZE Tello

V tejto časti si zhrnieme podrobnosti o drone DJI RYZE Tello, ktorý sme používali pri navrhovaní, implementovaní a testovaní softvérového rámca pre komunikáciu V2X v oblasti robotických rojov.

Dron DJI RYZE Tello je kvadroptéra vyrábaná firmou Ryze Tech v spolupráci s DJI a Intel. Tento model drona je cenovo dostupný a patrí k najľahším miniatúrnym dronom s hmotnosťou iba 80 g. Je programovateľný v jazyku Scratch, ako aj v Python a Swift. Tello dron je vybavený HD kamerou so zorným uhlom 82,6 stupňa, ktorá zachytáva videá s rozlíšením 720/30p a fotografie v rozlíšení 5 Mpx. Stabilizácia obrazu je elektronická a dron neobsahuje internú ani externú pamäť. Na ovládanie drona sa využívajú aplikácie Tello a Tello Hero, ktoré sú kompatibilné s operačnými systémami iOS a Android [3].

Tento dron disponuje množstvom funkcií, ako je Bounce Mode, kde dron automaticky vykonáva pohyby nahor a nadol počas letu. Ďalšou funkciou je Throw&Go, ktorá umožňuje vyhodit' drona z ruky a dron sa automaticky stabilizuje a vznáša sa. Dron dokáže aj rôzne druhy premetov a má možnosť nastavenia rýchleho a pomalého letu. Tento dron taktiež disponuje funkciou Up&Away, pri ktorej dron používateľ sleduje a pritom natáča video. Pri využití týchto funkcií je však dôležité mať nabitú batériu nad polovicu svojej kapacity. Dron Tello má maximálnu vzdialenosť na ovládanie 100 metrov, maximálnu výšku vzletu 10 metrov a na jedno nabitie batérie dokáže lietať až 13 minút. Farebné prevedenie je čierno-biele a existuje aj farebná kombinácia v spolupráci s firmou Marvel. Tento dron je vhodný na vzdelávanie aj zábavu [3].

V. V2X KOMUNIKÁCIA

V2X komunikácia znamená komunikáciu vozidla s inými zariadeniami v jeho okolí, umožňujúcu zdieľanie informácií o rýchlosti, smeru a polohe pre bezpečnejšiu komunikáciu. Táto komunikácia sa bežne využíva pomocou ad hoc sietí, ktoré sú bezdrôtové a spontánne vytvárané medzi zariadeniami v dosahu bez potreby trvalého infraštruktúrneho zázemia. Ad hoc siete sú flexibilné, mobilné a jednoduché na vytvorenie, ale môžu byť menej spoľahlivé a bezpečné ako tradičné pevné siete. Existujú rôzne typy ad hoc sietí, ako MANET, VANET a WANET, ktoré sú určené pre konkrétne použitia a situácie [4]. Komunikáciu V2X vieme rozdeliť do nasledujúcich základných podkategórií.

Vehicle-to-Vehicle (V2V) – komunikácia medzi dvoma vozidlami

Komunikácia V2V je jedna z kľúčových častí V2X, ktorá umožňuje priamu komunikáciu medzi vozidlami. Táto komunikácia umožňuje posielanie informácií o výstražných upozneniach alebo núdzovom brzdení. Prenos dát vyžaduje platnú autorizáciu a povolenie od operátora siete, a správy sú posielané tak, aby sa do sieťového prenosu zmestilo čo najviac informácií. Prenos správy o aplikácii V2V prebieha prevažne pomocou vysielania a môže sa diať priamo medzi vozidlami alebo prostredníctvom infraštruktúry, ktorá podporuje komunikáciu V2X, ak je dosah priamej komunikácie obmedzený[4].

Vehicle-to-Infrastructure (V2I) - komunikácia medzi vozidlom a infraštruktúrou

Vozidlá môžu podporovať aplikáciu V2I a vysielat' správy obsahujúce informácie o infraštruktúre a lokálnom serveri. Infraštruktúra, alebo aplikačný server môžu zasa vysielat' správy obsahujúce informácie alebo príkazy pre vozidlá, ktoré podporujú prijímanie správ V2I. Aplikačný server je priradený k určitej geografickej oblasti a môže existovať viacero takýchto serverov, ktoré sa môžu prekrývať, aby uľahčili a urýchlili komunikáciu[4].

Vehicle-to-Pedestrian (V2P) – komunikácia medzi vozidlom a chodcami

V2P je súčasťou technológie V2X, ktorá umožňuje komunikáciu medzi vozidlami a ich okolím. Táto komunikácia sa zameriava najmä na interakcie medzi vozidlami a chodcami, a to najmä pri prechádzaní cez cestu. Cieľom tejto komunikácie je zvýšiť bezpečnosť chodcov na cestách pomocou hodnotenia ich správania[4].

VI. NÁVRH A IMPLEMENTÁCIA SOFTVÉROVÉHO RÁMCA

Schéma fungovania softvérového rámca je nasledovná:

- Inicializácia: Softvérový rámec sa inicializuje a načíta údaje z konfiguračného súboru. Informácie zo softvérového rámca sa zapisujú do logovacieho súboru.
- Kontrola konfiguračných údajov: Prebehne kontrola načítaných údajov z konfiguračného súboru. Ak sú tieto údaje nesprávne, upozorní sa používateľ na chybu a softvérový rámec sa ukončí. Ak sú údaje správne, softvérový rámec pokračuje ďalej.
- Inicializácia komunikačného rozhrania: Softvérový rámec inicializuje komunikačné rozhranie a zapíše informácie do logovacieho súboru. Následne sa čaká na pripojenie drona.
- Pripojenie drona: Ak sa dron nepripojí po určitom čase, používateľ sa upozorní na problém s dronom a softvérový rámec sa ukončí. Ak sa dron pripojí, pošle sa mu prvotný inicializačný príkaz "command" pre aktiváciu drona, zapisujú sa informácie do logovacieho súboru.
- Vykonávanie príkazov: Softvérový rámec vykonáva príkazy zadané používateľom. Jednotlivé príkazy sa zapisujú do logovacieho súboru a prebieha kontrola správnosti zadaných príkazov. Ak je príkaz chybný, používateľ sa upozorní na chybu a pošle sa príkaz pre pristátie drona. Komunikácia a softvérový rámec sa ukončia.
- Spätná väzba a pokračovanie: Ak je príkaz správny, príde spätná väzba od drona a pokračujú ďalšie príkazy od používateľa. Tento proces sa opakuje až do príkazu pre pristátie drona.
- Ukončenie: Keď je zadaný koncový príkaz, softvérový rámec pošle príkaz na pristátie drona, zapisujú sa informácie do logovacieho súboru a nasleduje proces vypnutia a odpojenia drona od softvérového rámca. Následne sa softvérový rámec ukončí.

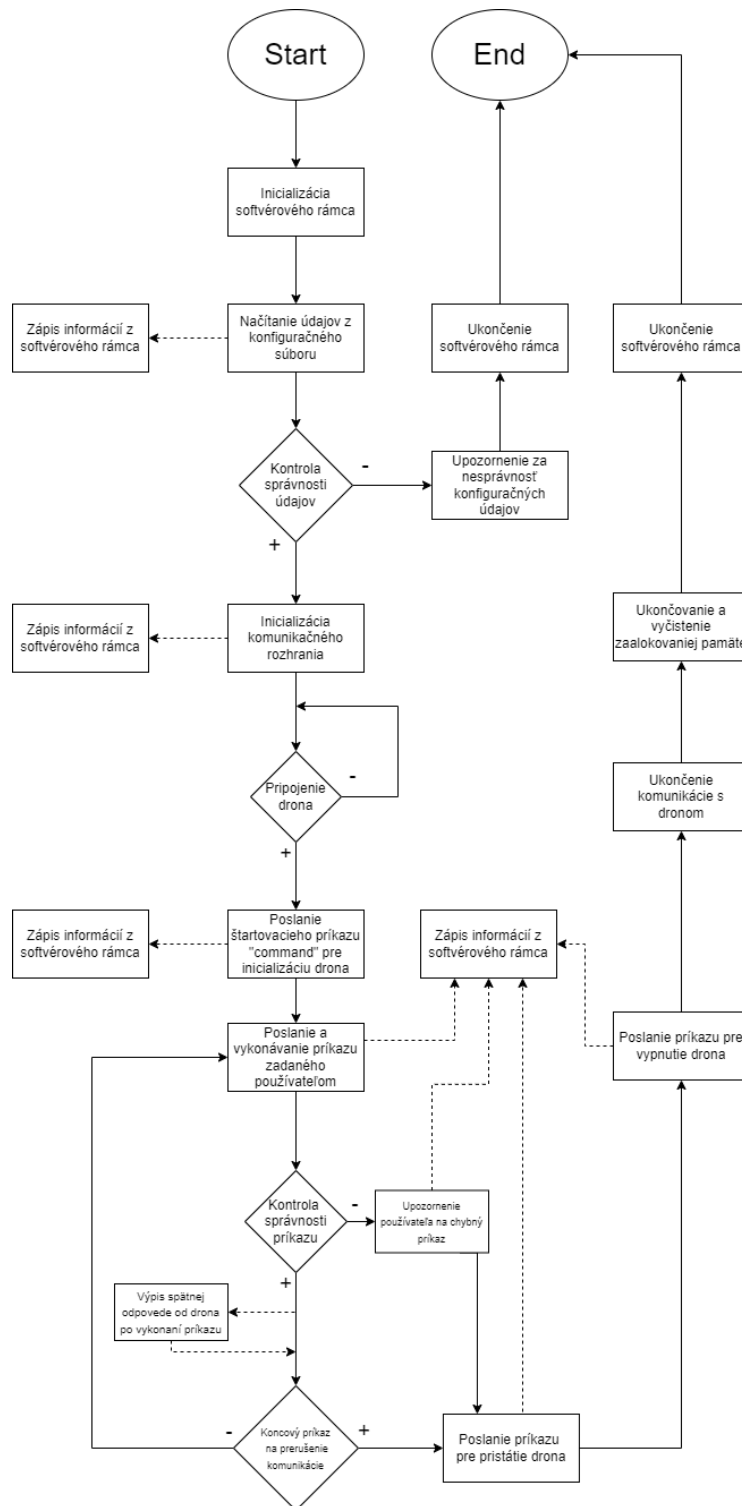
V priebehu celého procesu fungovania softvérového rámca sa taktiež vykonávajú rôzne ďalšie operácie a kontroly. Napríklad, po prijatí príkazu od používateľa sa môže vykonať kontrola, či je príkaz povolený pre aktuálny stav drona. Ak nie je, softvérový rámec môže upozorniť používateľa na neplatný príkaz.

Okrem toho, softvérový rámec môže obsahovať aj funkcie pre monitorovanie stavu drona, napríklad sledovanie jeho polohy, stavu batérie, pripojenia a iných parametrov. Tieto informácie

sa môžu zaznamenávať pomocou zadaných príkazov do logovacieho súboru pre neskoršiu analýzu a diagnostiku.

Pri ukončení komunikácie a odpojení drona od softvérového rámca sa vykonávajú aj rôzne uvoľňovacie operácie, napríklad zatvorenie komunikačného rozhrania, zapisovanie záverečných informácií do logovacieho súboru a ukončenie softvérového rámca.

Celkový proces fungovania softvérového rámca je teda riadený postupne načítanými konfiguračnými údajmi, komunikáciou s dronom, kontrolou zadaných príkazov od používateľa, zapisovaním informácií do logovacieho súboru a vykonávaním rôznych operácií v závislosti od stavu drona a zadaných príkazov, vid' jeho štruktúra na Obr. 2. Podrobnejšie informácie o jeho celej funkcionalite sa nachádzajú v [5].



Obr. 2 Schéma fungovania softvérového rámca s dronom.

VII. EXPERIMENTY A ICH VYHODNOTENIE

Pri validácií a testovaní softvérového rámca sme využívali tieto hardvérové prostriedky:

- Lenovo Legion 5 - gamingový notebook so štvorjadrovým procesorom AMD Ryzen 7, 16 GB RAM a grafickou kartou NVIDIA GeForce RTX 3060.
- Lenovo B590 - notebook určený pre kancelárske použitie s dvojjadrovým procesorom Intel Core i3, 4 GB RAM a integrovanou grafickou kartou Intel HD Graphics 4000.
- HP 250 G7 - notebook určený pre bežné použitie s dvojjadrovým procesorom Intel Celeron, 4 GB RAM a integrovanou grafickou kartou Intel HD Graphics 500.

Tieto zariadenia sme použili na testovanie softvérového rámca pre rojovú robotiku s rôznymi konfiguráciami. Cieľom týchto testov bolo overiť, či softvérový rámec je spoľahlivý, efektívny a dokáže zvládnuť potrebné nároky v reálnych podmienkach. V našich experimentoch sme sa zameriavali na meranie pohybu a presnosti drona s použitím softvérového rámca a rôznych notebookov.

V prvých experimentoch s dronom sme sledovali pohyb drona dopredu a dozadu na vzdialenosť jedného metra a zaznamenávali jeho odchýlky od cieľa. V ďalších dvoch experimentoch sme sledovali vzlietnutie a pristátie drona a jeho odchýlky. Štvrtý experiment zahŕňal otáčanie drona o určitý uhol a meranie jeho odchýlok od cieľa.

Pri meraní pohybu drona dopredu a dozadu sme zistili, že jednotlivé notebooky so softvérovým rámcom nemali žiadny vplyv na presnosť pohybu a odchýlky drona. Pri meraní vzlietnutia a pristátia sme zaznamenali zvýšené odchýlky, ale rozdiely neboli výrazné. V prípade merania otáčania drona sme nedokázali zaznamenať žiadne významné rozdiely v presnosti a odchýlkach pri použití rôznych notebookov s rovnakým softvérovým rámcom.

Okrem toho sme tiež sledovali rýchlosť importu, inicializácie a ukončenia softvérového rámca a zistili sme, že sa tieto procesy robia rýchlo a bez výraznej odchýlky.

Celkovo teda naše experimenty naznačujú, že pri použití softvérového rámca a rôznych notebookov nedochádza k výraznému vplyvu na presnosť a pohyb drona.

VIII. ZÁVER

Navrhli sme a implementovali spoľahlivý softvérový rámec pre V2X komunikáciu, ktorý sa osvedčil na rôznych typoch hardvérových zariadení bez väčších výkonostných rozdielov. Výsledkom našej práce je konkrétny návrh softvérového rámca pre rojových robotov s dôrazom na drony, ako členov roja. Tento rámec rozširuje poznatky o možnostiach komunikácie a koordinácie rojových robotov a môže byť úspešne aplikovaný v reálnom svete bez nutnosti prílišného zohľadňovania hardvérových požiadaviek. Naša práca tak poskytuje základ pre ďalšie rozvojové a optimalizačné kroky, ktoré môžu zvýšiť efektívnosť a spoľahlivosť rojových robotov v konkrétnych aplikáciách.

POĎAKOVANIE

Táto práca je finančne podporená Vedeckou grantovou agentúrou MŠVVaŠ SR a SAV prostredníctvom projektu VEGA 1/0480/22 — EDEN: EDge-Enabled inteligeNtné systémy.

LITERATÚRA

- [1] C. Hammerschmidt, "Applying v2x technology to drones," Oct 2017. [Online]. Available: <https://www.eenewseurope.com/en/applying-v2x-technology-to-drones-2/>
- [2] M. Smolko, "Omicron – návrh swarm robota s rádiovou navigáciou," in *Technická univerzita v Košiciach, bakalárska práca*, 2021.
- [3] "Ryze tello - dron," 2019. [Online]. Available: <https://www.alza.sk/ryze-tello-d5256900.htm#parametry>
- [4] M. Hasan, S. Mohan, T. Shimizu, and H. Lu, "Securing vehicle-to-everything (v2x) communication platforms," *IEEE Transactions on Intelligent Vehicles*, vol. 5, no. 4, pp. 693–713, 2020.
- [5] M. Smolko, "Návrh softvérového rámca pre komunikáciu v2x v oblasti robotických rojov," in *Technická univerzita v Košiciach, diplomová práca*, 2023.

Analýza mikropásikových anténnych polí pre hybridné FSO/RF systémy

¹Luboš OVSENÍK, ²Jakub ORAVEC, ³Samuel ANDREJČÍK,
⁴Martin SCHLACHTA

^{1,2,3,4} Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹lubos.ovsenik@tuke.sk, ²jakub.oravec@tuke.sk, ³samuel.andrejczik@tuke.sk,
⁴martin.schlachta@student.tuke.sk

Abstrakt — Príspevok sa zaoberá problematikou návrhu, simulácie a optimalizácie vhodného mikropásikového anténneho poľa pracujúceho na frekvencii 6 GHz, ktoré by bolo použiteľné v záložnom RF systéme hybridného bezdrôtového FSO/RF prenosového systému. Kládli sme si za cieľ dosiahnutie čo najvyššieho zisku a najnižšej hodnoty VSWR pre anténne pole, aby sa zabezpečila čo najlepšia komunikácia v bezdrôtovom prostredí. Pri návrhu bolo potrebné zohľadniť veľkosť antény, tvar a vzdialenosť jednotlivých mikropásikov.

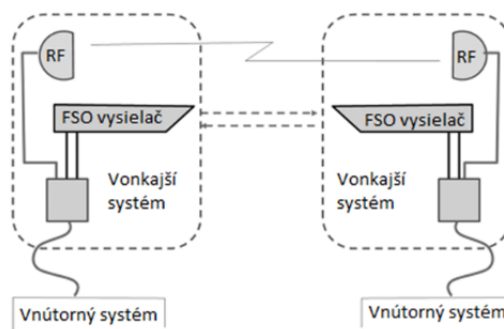
Kľúčové slová — FSO siete, FSO/RF systémy, mikropásikové antény, RF siete

I. ÚVOD

Udalosti posledných rokov a prechod mnohých ľudí na prácu z domu zapríčinili výrazné zvýšenie dopytu po vysokorýchlostnom pripojení. Poskytovatelia telekomunikačných služieb sú tak pre zachovanie konkurencieschopnosti nútení nahradzovať prístupové siete založené na zastaralých technológiách za moderné optické prístupové siete, ktoré môžu byť vláknové a bezdrôtové, známe tiež ako FSO (Free Space Optics). Prenos informácií v FSO systéme prebieha formou optického lúča, ktorý je modulovaný prenášanými dátami a ako prenosové médium využíva voľný priestor. 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. Táto technológia taktiež vykazuje citlivosť na rôzne poveternostné podmienky, ktoré môžu spôsobiť prerušenie spojenia.

II. HYBRIDNÝ FSO/RF SYSTÉM

Hybridná FSO/RF bezdrôtová komunikácia voľným priestorom (FSO) a komunikácia cez rádiové frekvencie (RF, Radio Frequency) 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 navrhnuť a vytvoriť spôsob prepínania medzi primárnou FSO a záložnou RF linkou [1].



Obr. 1 Hybridný FSO/RF prenosový systém.

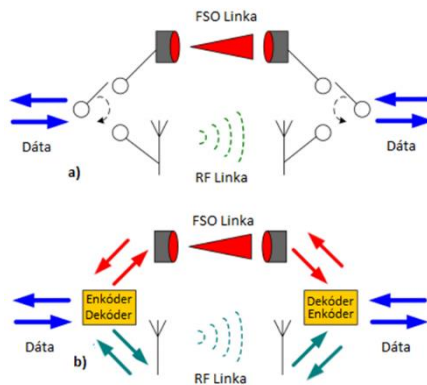
Očakáva sa, že dostupnosť spojenia FSO bude lepšia ako 99,999 %, no v niektorých oblastiach sa častejšie vyskytuje hustá hmla, krupobitie, sneh, dážď, piesočné búrky, atď. Keď sa hustá hmla ustáli v oblasti, na niekoľko hodín, zničí predpísanú očakávanú dostupnosť služieb. Aby služba bola k dispozícii 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; akonáhle 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 [2].

Výkon FSO komunikácie je drasticky ovplyvnený poveternostnými 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 poskytovať vysokú dostupnosť spojenia aj za nepriaznivých poveternostných podmienok [3,4]. 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.

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 hmla a atmosférickej scintilácie, zatiaľ čo RF sú obzvlášť citlivé na dažďový rozptyl. Použitie obidvoch systémov na prekonanie týchto problémov môže viesť k zvýšeniu spoľahlivosti a rýchlosti [4].

Pri konfigurácii tvrdého prepínania FSO linky (Obr. 2a) vysielateľ a prijímač spoločne vyberajú buď FSO alebo RF kanál pre dátový prenos. Vysokorýchlostné spojenie FSO sa volí len vtedy, ak podmienky kanálu umožňujú spoľahlivú komunikáciu, inak sú všetky údaje odosielané cez RF kanál. Vysielateľ 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“.



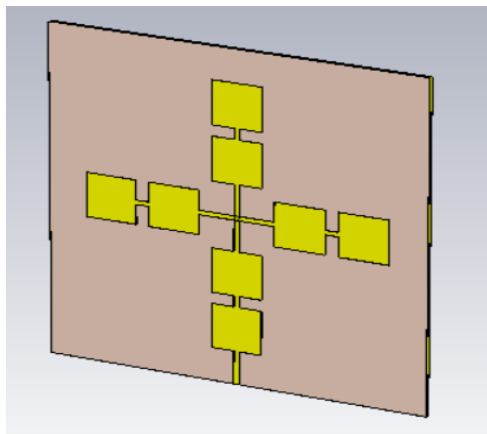
Obr. 2 a) Tvrdé, b) Mäkké prepínanie hybridného FSO/RF systému.

Mäkké prepínanie zobrazené na Obr. 2b, umožňuje použitie oboch liniek súčasne v rovnakom časovom okamihu a využíva krátke pakety zakódované pomocou Raptor kódu na odosielanie dát do oboch liniek. Tieto pakety sú potom prispôbované podmienkam v každej línii a zväčša nie sú obmedzené spätnou väzbou. Implementácia Raptor kódovania v programovateľných hradlových poliach (FPGA) podporuje kódovaciu rýchlosť 1,22 Gbps a dekódovaciu rýchlosť 714 Mbps so spotrebou 97 mW. Pri použití mäkkého prepínania v kombinácii s FSO prenosom o rýchlosti 1 Gbps a RF prenosom o rýchlosti 96 Mbps, bola dosiahnutá priemerná rýchlosť viac ako 472 Mbps, zatiaľ čo s tvrdým prepínaním bola priemerná rýchlosť len 112 Mbps [4,5].

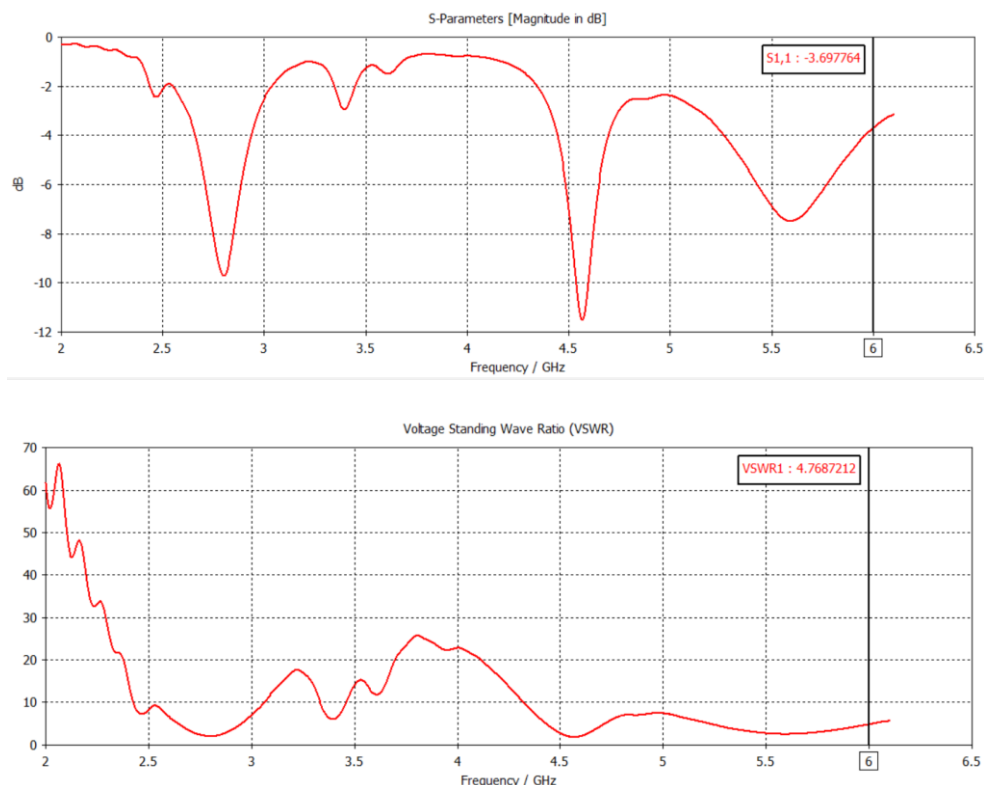
III. NÁVRH A ANALÝZA MIKROPÁSIKOVÉHO ANTÉNEHO POĽA PRE FREKVENCIU 6 GHz

A. Návrh základného mikropásikového anténneho poľa pre frekvenciu 6 GHz

Cieľom tohto návrhu je dosiahnutie čo najvyššieho zisku a najnižšej hodnoty VSWR (Voltage Standing Wave Ratio), aby sa zabezpečila čo najlepšia komunikácia v bezdrôtovom prostredí. Pri návrhu bolo potrebné zohľadniť veľkosť antény, tvar a vzdialenosť jednotlivých mikropásikov. Napájanie je navrhnuté ako priame pripojenie na mikropásikové vedenie, toto napájanie umožňuje väčšiu flexibilitu pri návrhu. Veľkosť substrátu a zemniatej plochy antény malo byť čo najmenšie aby sa zachovala jednoduchá manipulácia a inštalácia, nepresahuje teda 155 milimetrov. Ukážka modelu je zobrazená na Obr. 3.



Obr. 3 Ukážka základného mikropásikového anténneho poľa pre frekvenciu 6 GHz – pred optimalizáciou.

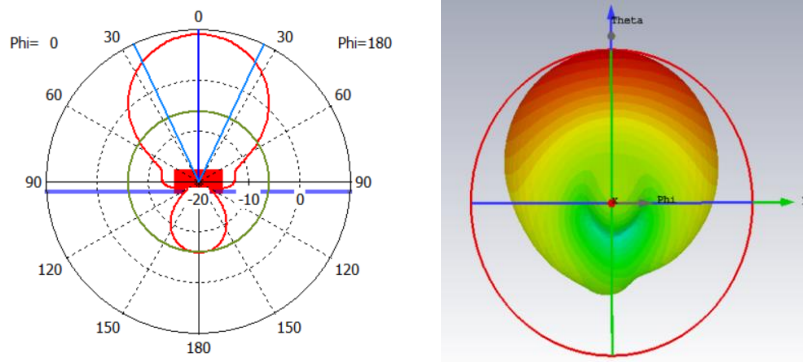


Obr. 4 a) Hodnota S-parametra, b) Hodnota VSWR modelu pre frekvenciu 6 GHz – pred optimalizáciou.

Obr. 4 a Obr. 5 zobrazuje vlastnosti ako S-parameter, VSWR hodnotu a vyžarovacie charakteristiky, ktoré boli namerané pre tento model pred optimalizáciou.

Tabuľka 1
 Hodnota zisku a smerovosti pre základný model s frekvenciou 6 GHz – pred optimalizáciou

Zisk (dBi)	Smerovosť (dBi)
7,836	8,931

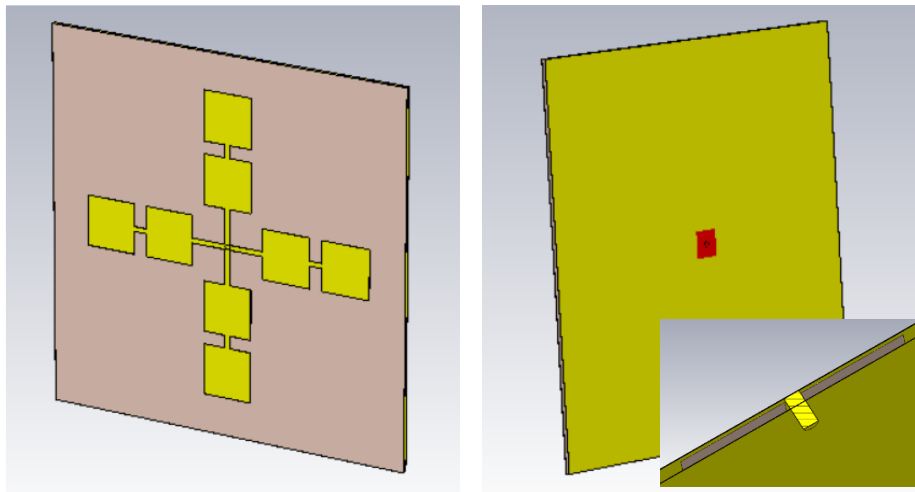


Obr. 5 Vyžarovacie charakteristiky modelu pre frekvenciu 6 GHz – pred optimalizáciou.

B. Návrh optimalizovaného mikropásikového anténneho poľa pre frekvenciu 6 GHz

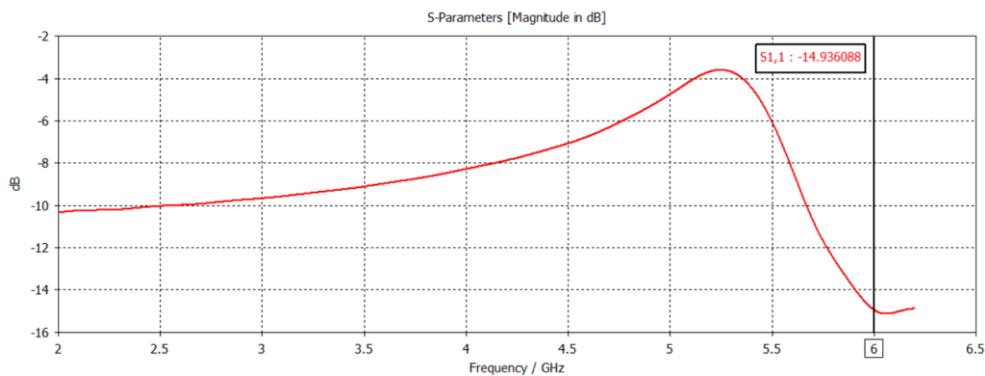
Ďalším krokom bola optimalizácia modelu mikropásikových anténnych polí pre frekvenciu 6GHz. Pri optimalizácii sme sa zamerali na čo najnižšie hodnoty S-parametra a hodnotu VSWR a na čo najlepšie vyžarovacie charakteristiky preto, aby sa dosiahol maximálny zisk a smerovosť antény. Boli vykonané viaceré experimenty, zmenšenie substrátu a zemniatej plochy, zmena materiálu substrátu, zmena veľkosti jednotlivých mikropásikových vedení a plôšok. V neposlednom rade sme taktiež vyskúšali a implementovali napájanie cez koaxiálny kábel aby sa minimalizovalo rušenie, ktoré by mohlo ovplyvniť výkon anténneho poľa.

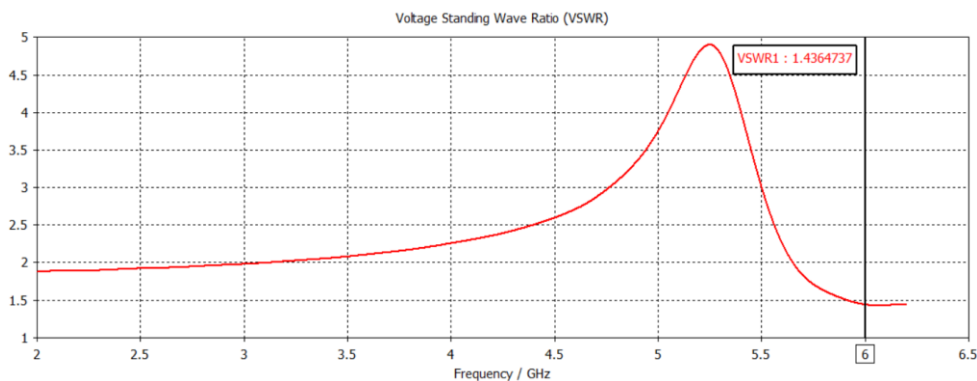
Po všetkých experimentoch sme zistili, že optimálnym modelom je model pôvodných rozmerov, t.j. 155 x 155 mm na substráte FR04 s koaxiálnym napájaním (Obr. 6).



Obr. 6 Ukážka optimalizovaného modelu pre frekvenciu 6 GHz s koaxiálnym napájaním: a) predná a b) zadná strana.

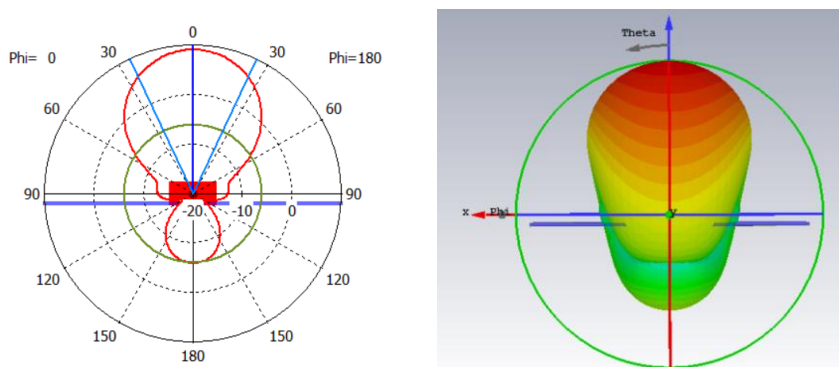
S-parameter vykazuje veľmi dobrú hodnotu a to -14,93 dB (Obr. 7a), a taktiež hodnota VSWR dosiahla výborné výsledky, je to zlepšenie oproti všetkým pokusom v optimalizácii a značí, že v anténe nedochádza k veľkým stratám (Obr. 7b).





Obr. 7 a) Hodnota S-parametra, b) Hodnota VSWR pre optimalizovaný model s frekvenciou 6 GHz.

Obr. 8 predstavuje vyžarovaciu charakteristiku tohto modelu. Hlavný lalok nie je odklonený, čo sa prejavilo aj na hodnotách zisku a smerovosti (Tabuľka 2), čo však postačuje pre danú aplikáciu.



Obr. 8 Vyžarovacie charakteristiky pre optimalizovaný model s frekvenciou 6 GHz.

Tabuľka 2
 Výsledné hodnoty pre model s koaxiálnym napájaním

Model	S-parameter (dB)	VSWR	Celková účinnosť (%)	Zisk (dBi)	Smerovosť (dBi)
S koaxiálnym napájaním	-14,936	1,436	49	14,568	10,741

IV. ZÁVER

Výstupom tejto práce je vytvorený optimalizovaný model mikropásikového anténneho poľa, ktoré je vhodné pre aplikáciu v hybridnom FSO/RF prenosovom systéme krátkého dosahu (do 200 m). Tento model môže byť nasadený ako externý model k wifi smerovaču impedančne prispôbenému 50 Ω. Riadenie toho smerovača je realizované prostredníctvom serveru, na ktorom prebieha strojové učenie schopné s dostatočným časovým predstihom predikovať výpadok spojenia a podrobenie možnosti jeho implementácie v záložnej linke hybridného FSO/RF systému.

POĎAKOVANIE

Táto práca bola podporovaná nasledujúcimi výskumnými projektami: APVV-17-0208 – „Odolné mobilné siete na doručovanie obsahu“ a VEGA 1/0260/23 – „Umelá inteligencia, robustné MANET a viacpreskakové D2D siete integrované do siete 6.generácie“.

LITERATÚRA

- [1] L. B. Stotts, et al., “Progress towards reliable free-space optical networks”, in *Military Communications Conference MILCOM 2011*, IEEE, 2011, pp. 1720-1726.
- [2] D. W. Young, et 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, et al., “Hybrid optical/radio frequency (RF) communications”, in *Advanced Free Space Optics (FSO)*, Springer, New York, NY, 2015, pp. 295-342.
- [4] A. Eslami, et al., “Hybrid channel codes for efficient FSO/RF communication systems”, *IEEE Transactions on Communications*, 58(10), 2010, pp. 2926 - 2938.
- [5] M. I. Jordan, et al., “Machine learning: Trends, perspectives, and prospects”, *Science*, 2015, pp. 255-260.

Návrh modelu systému „Radio over Fiber“

¹Luboš OVSEŇÍK, ²Jakub ORAVEC, ³Samuel ANDREJČÍK, ⁴Ján BUJŇÁK

^{1,2,3,4} Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹lubos.ovsenik@tuke.sk, ²jakub.oravec@tuke.sk, ³samuel.andrejczik@tuke.sk,
⁴jan.bujnak@student.tuke.sk

Abstrakt — Príspevok sa zaoberá návrhom a simuláciou RoF (Radio over Fiber) sietí, čo je integrácia optických vláknových sietí a mikrovln, ktoré predstavujú možné riešenie na zvýšenie kapacity a mobility a zároveň zníženie nákladov v prístupovej sieti.

Kľúčové slová —optické vláknové siete, RF siete, RoF siete

I. ÚVOD

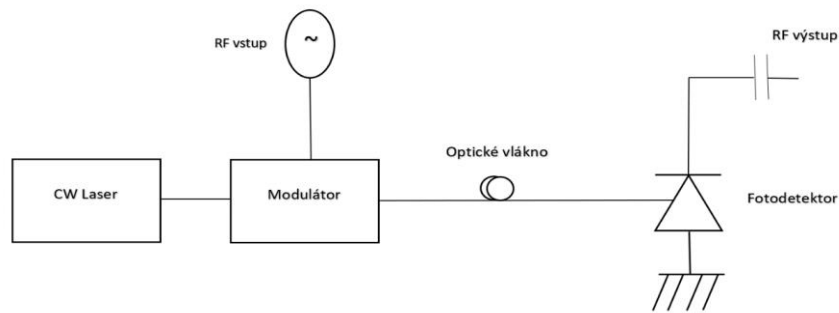
Vyššie prenosové rýchlosti, či prístup k internetovým službám je bežnou a často riešenou otázkou v dnešnej „informačnej“ spoločnosti. Jednou z najdynamickejšie rozvíjajúcich sa oblastí života je oblasť informačných technológií. Je preto vhodné a veľmi potrebné skúmať, objavovať a nadobúdať nové informácie o nových prístupoch k internetu. Získavanie nových informácií a skúmanie rôznych technológií (optických vláknových, rádiových,...), ktoré je možné využívať v počítačových sieťach, môže pomôcť pri rozšírení internetu aj do oblastí, kde je ťažké pripojiť sa k takýmto službám.

História súčasných optických telekomunikácií sa začína v 60. rokoch 20. storočia vynálezom lasera. Optické telekomunikácie boli v počiatku z dôvodu vysokých finančných nákladov a pomerne nízkej šírky pásma využívané len v chrbticových sieťach. Počas posledných desaťročí však technológie v oblasti optických komunikácií výrazne pokročili a dnes je možné v rámci jedného optického vlákna pomocou multiplexovania (TDM, WDM) vyslať na frekvenciách dosahujúcich stovky terahertzov (THz). Významným míľnikom bol vynález optického rozbočovača – splitttra, vďaka ktorému sa optické komunikácie začali aplikovať aj do prístupových sietí. Používaním optických rozbočovačov v prístupových sieťach je možné pripojiť až 128 koncových užívateľov na jedno optické vlákno, čím sa výrazne znižujú finančné náklady a zvyšuje efektívnosť siete. Motiváciou k tvorbe tejto práce boli rýchlo sa rozrastajúce optické prístupové siete FTTH, v ktorých nasadzovaní patrí Slovensko medzi popredné európske krajiny.

Taktiež rádiokomunikácia (RF, Radio Frequency) je nevyhnutnou technológiou pre súčasné prenosové systémy. Zažívame nárast dopytu po širokých pásmach pre prenos dát a zvyšovanie počtu používateľov bezdrôtovej komunikácie. Problémom sa stáva účinná a spoľahlivá prenosová kapacita. V tomto kontexte spojenia rádiokomunikácie a prenosu optickým vláknom sa RoF technológia (Radio over Fiber) javí ako perspektívne riešenie. RoF umožňuje prenos rádiových signálov prostredníctvom optického vlákna, čo umožňuje vysokorýchlostný a spoľahlivý prenos dát na veľké vzdialenosti [1].

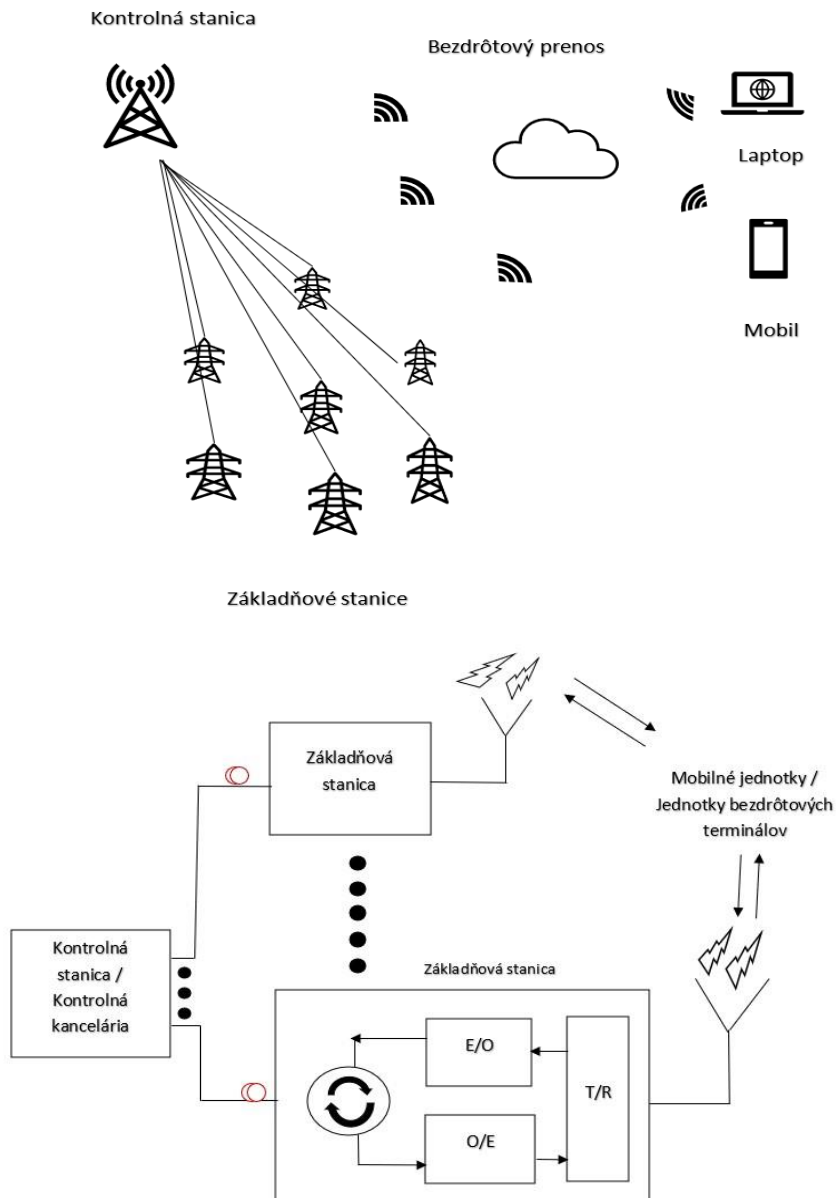
II. KONCEPT PRENOSU RF SIGNÁLU CEZ VLÁKNOVÚ OPTIKU

RoF systémy využívajú proces, pri ktorom je svetlo modulované rádiofrekvenčným (RF) signálom. Tento signál je prenášaný pomocou optického vlákna. Môžeme povedať, že je to systém, ktorý využíva prenos rádiofrekvenčného signálu cez optické vlákna, ako je zobrazené na Obr. 1. RoF vyhovuje veľkému množstvu používateľov, pretože využíva veľkú šírku pásma optických vlákien v porovnaní s obyčajným bezdrôtovým komunikačným systémom. Tento systém je výhodný z ekonomického hľadiska, z hľadiska flexibility, jednoduchej inštalácie a taktiež jeho charakteristickou vlastnosťou je nízky útlm. Ďalšou vlastnosťou RoF systému je odolnosť voči vysokofrekvenčnému rušeniu. Vďaka RoF systému je možné teda prenášať dáta v riedko, ale aj v husto osídlených lokalitách [2, 3, 4].



Obr. 1 Základná schéma RoF systému.

Technológia RoF je integrácia optických sietí a mikrovln, ktoré predstavujú možné riešenia na zvýšenie kapacity, mobility a zníženia nákladov v prístupovej sieti. RoF je vlastne analógové optické spojenie, ktoré prenáša modulované rádiových signály. Tieto signály sú smerované v optickej forme medzi centrálnou (kontrolnou) stanicou a základňovými stanicami pomocou optických vlákien. Základňové stanice prenášajú údaje do mobilných jednotiek, ktoré sú v ich dosahu, vo forme RF signálov (Obr. 2a) [5, 6].



Obr. 2 a) Názorné zobrazenie fungovania RoF systému, b) Základný blokový diagram RoF systému.

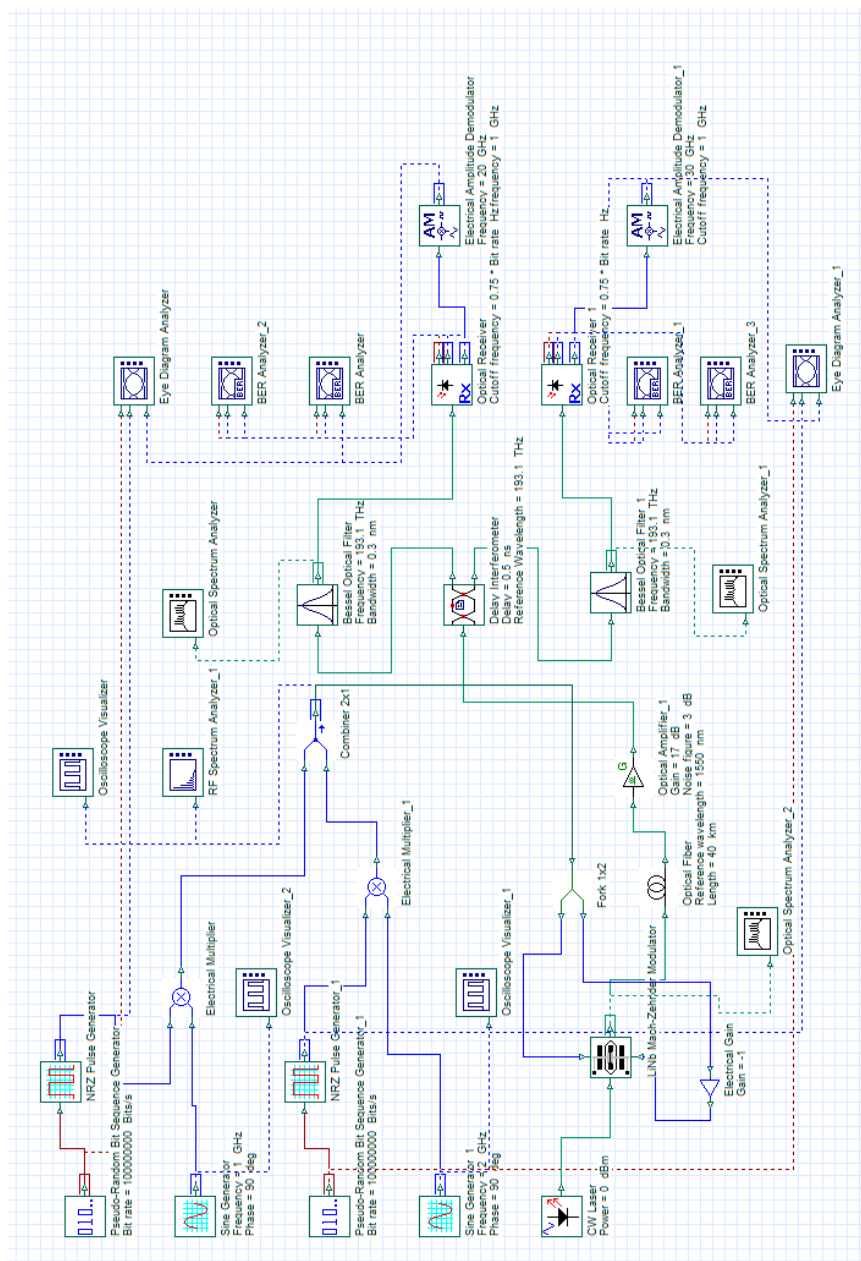
V kontrolnej stanici sa prichádzajúcim RF signálom moduluje svetelný lúč z kontinuálneho lasera (CW Laser). Potom sa takto upravený optický signál prenesie pomocou optického vlákna do príslušnej základňovej stanice, pričom veľkosť útlmu v tomto signáli je zanedbateľná. Na základňovej stanici sa signál prevádza z optickej formy na elektrickú (O/E). Prenos signálu medzi základňovou stanicou a mobilnou jednotkou (jednotkami bezdrôtových terminálov) sa uskutočňuje pomocou antény, ktorá je umiestnená na jednotke základňovej stanice (Obr. 2b). Kontrolná stanica komunikuje s jednotlivými základňovými stanicami pomocou optického signálu prenášaného pomocou optických vlákien, čím sa znižujú náklady na prenosové zariadenia, ako antény a zosilňovače. Môžeme povedať, že z hľadiska riadenia systému je RoF systém efektívnejší ako konvenčné bezdrôtové siete [7, 8, 9].

III. NÁVRH A ANALÝZA ROF SYSTÉMU

Hlavnou časťou tejto práce bolo vytvoriť, nasimulovať a optimalizovať RoF prenosový systém v programovacom balíku OptiSystem. Pracovali sme s rôznymi vlastnosťami a hodnotami, ktoré sme nastavovali pre dosiahnutie čo najlepšej optimalizácie navrhovaného systému.

A. Návrh základného 2 kanálového systému pre bitovú rýchlosť 100 Mb/s

V základnej schéme sme pracovali s dvoma kanálmi. Vzďialenosť sme nastavili na 40 km a bitovú rýchlosť na 100 Mb/s. Vytvorili sme dvojkanálový prenosový RoF systém s frekvenciou 1550 nm. Na Obr. 3 je znázornená základná schéma s vyššie opisovanými vlastnosťami.



Obr. 3 Základná schéma použitého RoF systému.

V nasledujúcej tabuľke (Tabuľka 1) môžeme vidieť použité vstupné parametre prenosového RoF systému v základnej schéme. V Tabuľke 2 sú hodnoty, ktoré sme zistili pri simulácii.

Tabuľka 1
 Parametre základnej schémy prenosového RoF systému

Vlastnosť	Hodnota	Jednotka
Počet kanálov	2	
Vlnová dĺžka použitého signálu	1550	nm
Výkon lasera	1	mW
Typ modulácie	NRZ	
Bitová rýchlosť	100	Mb/s
Dĺžka optického vlákna	40	km
Výkon zosilňovača	10	dBm
Zisk zosilňovača	1	dB

Tabuľka 2
 Hodnoty získané po simulácii prenosového RoF systému

Parameter/Kanál	Prvý	Druhý
Max. Q faktor	9.545	9.772
Min. BER	5.684×10^{-22}	5.988×10^{-23}
Výška oka	1.03×10^{-4}	1.03×10^{-4}

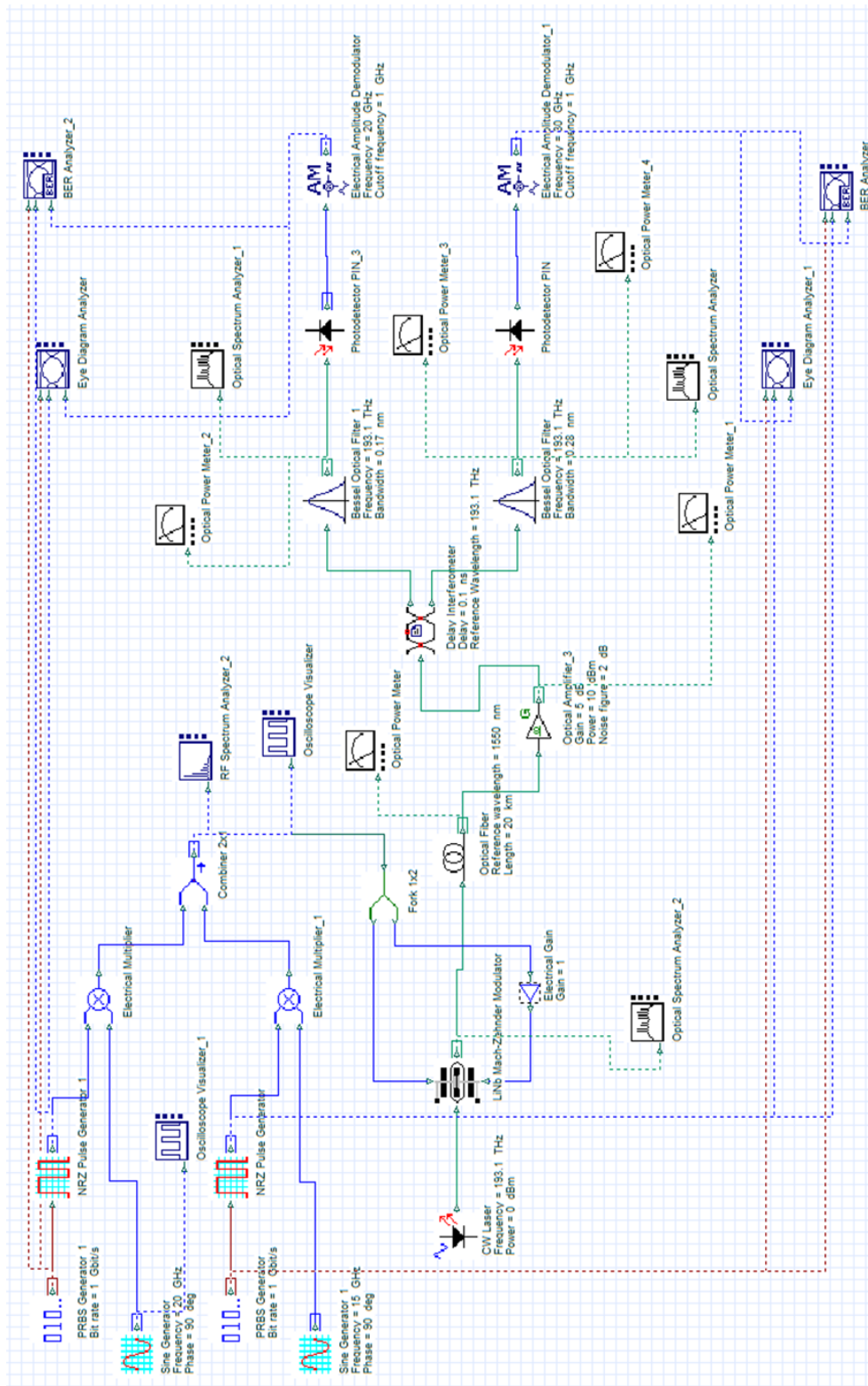
B. Návrh RoF systému s bitovou rýchlosťou 1 Gb/s

Pôvodnú základnú schému s bitovou rýchlosťou 100 Mb/s sme optimalizovali a zvýšili sme bitovú rýchlosť na 1 Gb/s. Nasimulovali a optimalizovali sme systém pre rôzne dĺžky optického vlákna (20 km, 30 km, 40 km, 50 km, 60 km a 100 km).

V Tabuľke 3 sú zhrnuté základné parametre, ktoré sme použili pri optimalizácii schémy. Na Obr. 4 je optimalizovaná schéma RoF prenosového systému pre 1Gb/s režim práce.

Tabuľka 3
 Parametre optimalizovanej schémy prenosového RoF systému pre 1Gb/s

Vlastnosť	Hodnota	Jednotka
Počet kanálov	2	
Bitová rýchlosť	1	Gb/s
Výkon lasera	1	mW
Typ modulácie	NRZ	
Vlnová dĺžka použitého signálu	1550	nm
Dĺžka optického vlákna	20, 30, 40, 50, 60, 100	km
Útlm optického vlákna	0.2	dB/km
Zisk zosilňovača	5, 10	dB
Frekvencia lasera	193.1	THz



Obr. 4 Optimalizovaná schéma RoF prenosového systému pre 1Gb/s režim práce.

Po optimalizovaní RoF systémov na rôzne dĺžky sme dostali výsledky zobrazené v Tabuľke 4. Kvalitné výsledky sme dostali pri simuláciách o rýchlosti 1 Gb a dĺžke optického vlákna od 20 do 100 km. Od dĺžky 20 do 50 km mali výsledky minimálnej bitovej chybovosti klesajúcu tendenciu. Pri dĺžke 60 km sa tieto hodnoty zlepšili z dôvodu navýšenia zisku zosilňovača z 5 na 10 dB.

Tabuľka 4
 Porovnanie výsledkov simulácií

ROF systém s rýchlosťou 1Gb/s pri vzdialenosti 20 km			
	Max. Q faktor	Min. BER	Výška oka
Prvý kanál	10.167	1.385×10^{-24}	1.406×10^{-5}
Druhý kanál	10.849	9.656×10^{-28}	7.884×10^{-6}
ROF systém s rýchlosťou 1Gb/s pri vzdialenosti 30 km			
	Max. Q faktor	Min. BER	Výška oka
Prvý kanál	8.243	7.757×10^{-17}	2.260×10^{-5}
Druhý kanál	8.974	1.370×10^{-19}	6.3×10^{-6}
ROF systém s rýchlosťou 1Gb/s pri vzdialenosti 40 km			
	Max. Q faktor	Min. BER	Výška oka
Prvý kanál	8.321	4.116×10^{-17}	2.506×10^{-4}
Druhý kanál	8.461	1.193×10^{-17}	2.728×10^{-5}
ROF systém s rýchlosťou 1Gb/s pri vzdialenosti 50 km			
	Max. Q faktor	Min. BER	Výška oka
Prvý kanál	8.637	2.561×10^{-18}	3.558×10^{-5}
Druhý kanál	9.678	1.775×10^{-22}	9.318×10^{-6}
ROF systém s rýchlosťou 1Gb/s pri vzdialenosti 60 km			
	Max. Q faktor	Min. BER	Výška oka
Prvý kanál	9.679	1.848×10^{-22}	6.569×10^{-5}
Druhý kanál	8.702	1.582×10^{-18}	6.471×10^{-6}
ROF systém s rýchlosťou 1Gb/s pri vzdialenosti 100 km			
	Max. Q faktor	Min. BER	Výška oka
Prvý kanál	9.141	2.869×10^{-20}	6.813×10^{-5}
Druhý kanál	7.991	6.124×10^{-16}	5.567×10^{-6}

IV. ZÁVER

V dnešnej dobe si spoločnosť zakladá hlavne na vyššej prenosovej rýchlosti, a preto je 1Gb/s RoF systém vhodnejší na použitie ako systém 100 Mb/s. Vo svojej výskumnej práci sme sa pokúsili nasimulovať a optimalizovať RoF prenosový systém tak, aby správne pracoval aj pri bitovej rýchlosti 10 Gb/s. Systém s takou rýchlosťou sa nám však nepodarilo dostatočne optimalizovať a nedosahovali sme výsledky, ktoré by boli vhodné na použitie. Pre dosiahnutie požadovaných výsledkov pri systéme s bitovou rýchlosťou 10 Gb/s by sme potrebovali lepšie hardwarové vybavenie.

POĎAKOVANIE

Táto práca bola podporovaná nasledujúcimi výskumnými projektami: APVV-17-0208 – „Odolné mobilné siete na doručovanie obsahu“ a VEGA 1/0260/23 – „Umelá inteligencia, robustné MANET a viacpreskokové D2D siete integrované do siete 6.generácie“.

LITERATÚRA

- [1] P. Sahu, "Fundamentals Of Optical Networks And Components", in Florida: CRC Press, 2020, 365 pp.
- [2] T. Javed, et al., "Radio over Fiber Technology", in <https://www.ijser.org/researchpaper/Radio-over-Fiber-Technology.pdf>, 2016.
- [3] M. Kowalczyk, et al., "Microwave and Optical Technology Letters", in <https://doi.org/10.1002/mop.24277>, 2009.
- [4] B. Sighn, et al., "A Review on Advantages and Applications of Radio over Fiber System", in <https://inpressco.com/wp-content/uploads/2016/06/Paper581042-1044.pdf>, 2016.
- [5] "What is RF over fiber technology and what are the benefits?", in <https://www.vialite.com/resources/guides/what-is-rf-over-fiber-technology/>, 2022.
- [6] S. Das, et. al., "Modeling and Performance Analysis of RoF System for Home Area Network with Different Line Coding Schemes Using OptiSystem", in <https://docplayer.net/85030516-Modeling-and-performance-analysis-of-rof-system-for-home-area-network-with-different-line-coding-schemes-using-optisystem.html>, 2014.
- [7] J. Johny and S. Shashidharan, "Design and Simulation of a Radio Over Fiber System", Hindustan University, Chennai, India: Department of Electronics and Communication Engineering, Hindustan Institute of Technology and Science, in <https://cdn.optiwave.com/wp-content/uploads/2015/07/1569641487.pdf>, 2012.
- [8] M. K. Ibrahim, et al., "Four Wave Mixing Mitigation in Already Existing Dense Wavelength Division Multiplexing Networks", Bagdad: Al-Nahrain University, in https://kalaharijournals.com/resources/FebV7_I2_8.pdf, 2022, ISSN: 0974-5823.
- [9] M. Pooja, et al., "Advantages and limitations of radio over fiber system", *International Journal of Computer Science and Mobile Computing*, 4(5), 2015, pp. 506-511, ISSN 2320-088X.

Detecting Melanoma in Images Using Neural Networks: A Hybrid Application for Early Diagnosis

¹*Martin NGUYEN*

¹Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

¹martin.nguyen@student.tuke.sk

Abstract – Melanoma is a deadly form of skin cancer that can spread quickly if not detected early. Unfortunately, it is not very uncommon that the patient himself overlooks such diseases and often comes for examination in the late stages of the disease. In recent years, there has been a growing interest in using computer vision and machine learning techniques to assist doctors in detecting diseases from an image. This paper proposes a hybrid application that uses a convolutional neural network (CNN) to classify skin lesions as either malignant or benign. Our application helps the user in deciding whether to visit a doctor or not. In addition, the application contains articles with information about this type of disease and map of ambulances in the user's area. In addition to ordinary people, the proposed application can be used as a diagnostic tool to aid dermatologist trainees in identifying melanoma at an early stage, potentially saving lives and improving patient outcomes. We presented our application at the Živé IT Projekty 2023 competition and awarded first place among 47 competing projects.

Keywords – Convolutional Neural Networks, Image analysis, Melanoma disease, ResNet-50

I. INTRODUCTION

Melanoma is a form of skin cancer that can be deadly if not detected and treated early. Despite the availability of various screening methods, many cases of melanoma go undetected until they have reached an advanced stage. Melanoma is a growing public health concern, with increasing incidence and deaths. To lower future incidence and mortality, more effective public education programs need to be developed [1]. As a result, there has been a growing interest in developing computer-aided diagnostic systems that can help the user in deciding whether to visit a doctor or not by creating an application for detecting melanoma from an image even at an early stage. In recent years, machine learning and computer vision techniques have shown promising results in this regard, such as their ability to detect tumors [2] or recognizing COVID-19 through image analysis [3].

In this paper, we propose an application that utilizes a convolutional neural network (CNN) to classify skin lesions as either malignant or benign. Our application not only helps users decide whether to visit a doctor but also provides information on melanoma and the location of ambulances in the user's area. This hybrid system has the potential to improve patient outcomes and potentially save lives. Furthermore, our proposed application can also be used as a diagnostic tool to aid dermatologist trainees or students of medicine in identifying melanoma. In this paper, we describe the design and implementation of our proposed system, and present the results of our experiments.

II. DEVELOPMENT PLANNING AND OBTAINING USER REQUIREMENTS

For the development of our proposed application for identifying melanoma in images, we consulted with two dermatovenerologists to gain insights and feedback about the project. During the meeting, we discussed the application's purpose, potential features, and the target audience. Based on the feedback and discussions, we decided to target three groups: ordinary people, medical students, and dermatovenerology trainees.

We collected feature ideas for the application from the specialists, which included basic and extended functionalities of the application. The flow diagram of the basic feature is displayed in the figure 1. User will upload or take a photo of the skin lesion. The application then analyzes the image and returns a malign/benign result with accuracy percentage. If the result is malignant, the app offers to show a map of specialists in the user's area, making it easier for the user to seek immediate medical attention. The application provides informative articles about the disease on its blog section, regardless of whether the result is positive or negative.

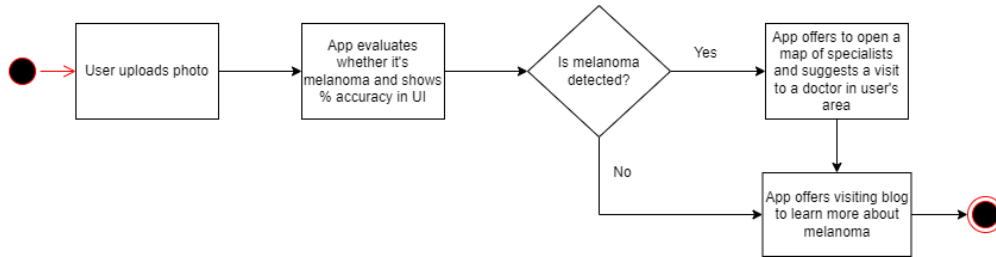


Fig. 1 Flow diagram of the basic feature

In order to reach a wider audience and provide accessible melanoma detection and information to more people, we decided to create a separate website and a multi-platform application that targets iOS and Android mobile devices.

III. EXISTING RESEARCHES

There are many research papers that addressed the issue of detecting melanoma in images. Very common is the use of deep learning methods, such as using convolutional neural networks for the classification. One of the example research is [4], where authors used deep learning methods to identify melanoma from dermoscopy images. The system produces a higher accuracy compared to the average of eight expert dermatologists. Their solution use deep learning methods and is trained on the ISIC public benchmark dataset. They evaluated the accuracy of their model to be more than 90%, whereas during the user testing, they achieved accuracy of 76%.

Another research [5] proposed solution using deep convolutional neural network Encoder-Decoder. They achieved very high accuracy of 95% on the benchmark ISIC dataset, which is more effective than most of existing state-of-art methods. However, the use of deep learning methods is not mandatory to achieve high accuracy. Authors in the research [6] used different techniques to extract features, such as Principal component analysis (PCA), and Support Vector Machine (SVM) for classification purposes. Authors state that the achieved classification accuracy is 92%, which is comparable to deep learning methods.

Authors in [7] propose an accessible mobile solution for melanoma detection using mobile image analysis. They share the same goal as our project, to create an application that could be employed by the general public for preliminary self-screening or assist the general physicians during the diagnosis process. They propose a detection system that is optimized to run entirely on the resource-constrained smartphone, so there is no need to have model running externally on the server. For that reason, they used SVM for classification purposes and achieved accuracy around 89%.

IV. MACHINE LEARNING MODEL TRAINING

The first step before performing the training process of our machine learning model was to obtain a dataset of skin lesion images. We utilized the ISIC Archive dataset hosted by the International Skin Imaging Collaboration. The dataset provides a vast amount of high-quality images for our training purposes. Once we had the dataset, we performed several preprocessing steps to ensure that the images were suitable for training. The first step was to remove hair from the images, which was accomplished using an hair-removal algorithm. After that, we applied an auto-cropping algorithm to the images to focus on the skin lesion, eliminating any irrelevant background. Example of preprocessed image is shown in figure 2.

The next step was to prepare the data for training by encoding data that were not a number, for example we mapped malignant feature as 1 and benign feature as 0. Then we split the dataset into subsets for training and testing. To increase the size of the dataset and reduce overfitting, we employed data augmentation techniques such as rotating, scaling, flipping, and adding noise to the images. With that, we obtained at least 8 more images from one single image. Example is shown in figure 3.

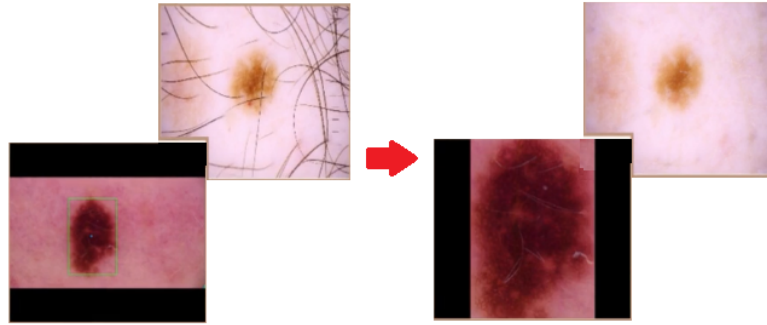


Fig. 2 Example of used data augmentation techniques



Fig. 3 Example of used data augmentation techniques

We then trained the classification model using a convolutional neural network (CNN) with a ResNet-50 architecture and the Adam optimizer. ResNet-50 is a convolutional neural network that is 50 layers deep and its architecture was successfully used in many image diagnosis applications, for example [8] which proposed an accurate and inclusive computational breast cancer diagnosis framework or [9] where transfer learning based classification of malarial infected cells was implemented. The CNN consisted of multiple layers, including Flatten, relu Dense 254, relu Dense 128, and softmax Dense 2. During training, we optimized the weights to minimize the loss function and achieve a high level of accuracy. After training, we evaluated the model's performance using the test dataset and measured the accuracy offline. As shown in the figure 4, the model achieved an accuracy of approximately 85%, indicating that it was successful in classifying skin lesions as either malignant or benign.

The false negatives rate of 0.16 implies that approximately 16% of actual malignant lesions were incorrectly classified as benign. This means that some malignant cases were mistakenly labeled as non-threatening, which can be concerning as it may lead to a delay in appropriate medical intervention.

On the other hand, the false positives rate of 0.11 suggests that around 11% of benign lesions were incorrectly classified as malignant. While false positives may not have immediate health consequences for the patients, they can cause unnecessary anxiety and potentially lead to unnecessary medical procedures. Further analysis and fine-tuning may help address these misclassification issues.

The final step was to deploy the trained model to the backend of the application for online evaluation. We tested the model on a small set of real-world images and found it to perform well, with satisfactory results. The application was then made available to users, allowing them to upload or take photos of skin lesions and receive an immediate classification result.

V. MOBILE AND WEB APPLICATION DEVELOPMENT

The backend of the application is the interface between the frontend/mobile application and the trained machine learning classification model. In order to accomplish this, an API in Flask (Python) was developed. This API is responsible for uploading images and returning a malignant/benign result with percentage accuracy. In addition, the API also provides article data for the blog section of the application as well as information about specialists in the page visitor's area. To ensure the efficient management of data, the application uses a PostgreSQL database. Additionally, Docker was used for containerization to facilitate the deployment and scaling of the application.

The frontend of the application was developed using the Angular framework. This user interface allows users to easily upload pictures and see the results. If malign melanoma is detected, the application offers user to redirect into page with the map of specialists in the user's area. The user can also access the blog section, where informative articles about the disease are displayed.

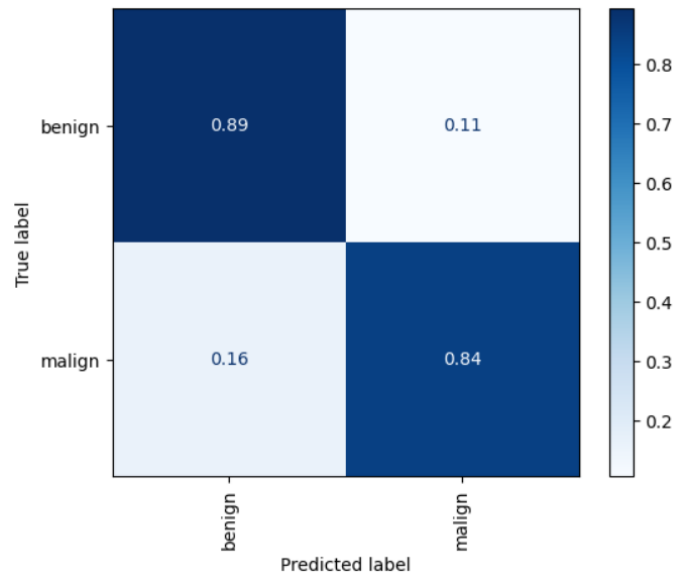


Fig. 4 Confusion matrix for the trained model

The mobile application was built using the open-source Flutter framework. Using Flutter allowed us to create a mobile application that is visually appealing, responsive, and compatible with both Android and iOS platforms, all while reducing development time and effort. It allowed us to build natively compiled mobile application for both iOS and Android operating systems using a single codebase.

The API developed for the backend of the application is designed to make it easy for users to upload images for analysis. Upon uploading an image, the API creates a request for the trained machine learning model to analyze the image and determine whether the skin lesion is benign or malignant. The result is returned with a percentage accuracy, providing users with important information that can help them make informed decisions about seeking medical assistance.

In addition to providing image analysis, the backend of the application also provides access to a blog section that contains informative articles about melanoma. These articles are designed to provide users with important information about the disease, including risk factors, warning signs, and treatment options. The blog section is easily accessible from the frontend of the application, making it easy for users to access the information they need to make informed decisions about their health.

To further aid users in finding help, the application also provides information about specialists in the user's area. This information is displayed on a map, making it easy for users to find nearby specialists who can help with the diagnosis and treatment of melanoma. The map is interactive, allowing users to click on specialists' locations to view additional information, such as contact information and hours of operation.

VI. PRESENTING THE APPLICATION IN THE COMPETITION

Presenting our application at the Živé IT Projekty 2023 competition was an exciting opportunity to showcase the work and compete with other innovative projects. During our presentation, we first defined the problem that our project aimed to solve - the difficulty of detecting melanoma at an early stage, which can often lead to fatal outcomes. We highlighted the importance of early detection and the limitations of traditional detection methods, which rely on visual inspection by dermatologists.

We then presented how our project solves this problem - by using computer vision and machine learning techniques to classify skin lesions as either malignant or benign. We explained that our application uses a convolutional neural network (CNN) trained on a dataset of skin images to make accurate predictions about the nature of skin lesions. By offering users a quick and easy way to upload a photo of a suspicious lesion and receive an instant diagnosis, our application empowers people to take control of their health and seek treatment early on.

We emphasized the impact that our project can have on the world around us. Early detection of melanoma can significantly improve patient outcomes and reduce healthcare costs associated with late-stage treatment. By providing accessible and affordable diagnostic tools, our application can potentially save lives and promote public health.

We also described the target audience for our application, which includes ordinary people, medical students, and dermatology trainees. Our application can be used as a diagnostic tool to aid in the early detection of melanoma, as well as a learning tool to help trainees improve their diagnostic skills.

We highlighted how our project is up to date with the latest trends in computer vision and machine learning. We used state-of-the-art techniques such as the ResNet-50 CNN architecture and data augmentation to train our model and achieve high accuracy on our test dataset. We also explained how our application incorporates features such as a blog section and a map of specialists to provide users with additional information and resources.

Finally, we presented our business plan, which involves offering our application as a free download on both mobile and web platforms. We plan to generate revenue through partnerships with dermatology clinics and hospitals, who can use our application to refer patients and streamline the diagnostic process.

Our application was recognized with the first prize in the competition. We received only positive feedback, emphasizing the need to continue working on the project.

VII. CONCLUSION

We successfully trained a machine learning model for the identification of melanoma in skin lesion images using deep learning techniques. The model was trained using a convolutional neural network with a ResNet-50 architecture and achieved an accuracy of approximately 85%. The application was made available on multiple platforms, including both mobile and website app, to reach a wider audience and provide accessible melanoma detection and information to more people. Overall, our proposed application has the potential to assist dermatologists in detecting melanoma at an early stage, potentially saving lives and improving patient outcomes.

The backend of the application provides an efficient interface between the frontend/mobile application and the trained machine learning model. The API in Flask (Python) is responsible for uploading images, returning malign/benign results, and providing access to informative articles and specialist information. The frontend of the application is designed using the Angular framework, offering users an intuitive and easy-to-use interface that allows for quick image analysis, access to informative articles, and specialist information.

Overall, presenting our application at the Živé IT Projekty 2023 competition was a valuable experience that allowed us to showcase our work, receive feedback from industry experts, and gain recognition for our efforts.

ACKNOWLEDGMENT

We are grateful to everyone who contributed to the success of this project and the research in the field. We would especially like to thank our team members Stanislav Ivanov, Oleksandr Fetkovych, Martina Kuchtova, Dominik Džama and František Kurimský, who developed the application under our supervision and provided valuable input, insights, and assistance at every stage of the project.

REFERENCES

- [1] D. S. Rigel, "Epidemiology of melanoma," *Seminars in Cutaneous Medicine and Surgery*, vol. 29, no. 4, p. 204–209, 2010.
- [2] D. Vasanthi, U. Palani, M. Vishnupriya, and K. Saundariya, "Detection of tumor by fusing medical images using artificial neural network," *2021 International Conference on System, Computation, Automation and Networking (ICSCAN)*, 2021.
- [3] J. Qin, J. Hu, J. Li, and P. Yan, "Convolutional neural network for computer-aided identification: Detection of lung ct images using cnn for covid-19 identification," *2022 International Conference on Big Data, Information and Computer Network (BDICN)*, 2022.
- [4] N. C. Codella, Q.-B. Nguyen, S. Pankanti, D. A. Gutman, B. Helba, A. C. Halpern, and J. R. Smith, "Deep learning ensembles for melanoma recognition in dermoscopy images," *IBM Journal of Research and Development*, vol. 61, no. 4/5, 2017.
- [5] A. A. Adegun and S. Viriri, "Deep learning-based system for automatic melanoma detection," *IEEE Access*, vol. 8, p. 7160–7172, 2020.
- [6] H. Alquran, I. A. Qasmieh, A. M. Alqudah, S. Alhammouri, E. Alawneh, A. Abughazaleh, and F. Hasayen, "The melanoma skin cancer detection and classification using support vector machine," *2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, 2017.
- [7] T.-T. Do, T. Hoang, V. Pomponiu, Y. Zhou, Z. Chen, N.-M. Cheung, D. Koh, A. Tan, and S.-H. Tan, "Accessible melanoma detection using smartphones and mobile image analysis," *IEEE Transactions on Multimedia*, vol. 20, no. 10, p. 2849–2864, 2018.
- [8] Q. A. Al-Haija and A. Adebajo, "Breast cancer diagnosis in histopathological images using resnet-50 convolutional neural network," *2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*, 2020.
- [9] A. S. Reddy and D. S. Juliet, "Transfer learning with resnet-50 for malaria cell-image classification," *2019 International Conference on Communication and Signal Processing (ICCS)*, 2019.

Návrh navigačného systému na báze neuro-fuzzy prístupu s využitím cloud computing-u

Ján VAŠČÁK

Katedra Kybernetiky a Umelej Inteligencie, Fakulta Elektrotechniky a Informatiky,
Technická Univerzita v Košiciach, Slovenská Republika

jan.vascak@tuke.sk

Abstrakt – Jednou zo základných úloh mobilných robotov je ich navigácia v prostredí s prekážkami. Existuje celá rada navigačných prístupov pre riadenie obchádzania prekážok avšak tieto sú väčšinou spojené s vysokými výpočtovými nárokmi, čo vedie k potrebe umiestniť na robotovi výkonné výpočtové zariadenie a čo má tým pádom za následok predražovanie riešenia a ďalšie technické komplikácie. Jednou z možností je distribuovať výpočet aj na zariadenia mimo robota, k čomu je možné využiť technológie cloud computing-u. Táto práca sa zaoberá overením možnosti využitia cloud computing-u pri navigácii mobilných robotov v simulačnom prostredí Webots s využitím navigačného systému na báze neuro-fuzzy prístupu. Experimenty potvrdili schopnosť sa úspešne vyhýbať prekážkam ako aj skutočnosť, že využitie cloud computing-u predstavuje výrazné zrýchlenie komunikácie oproti iným typom sieťových aplikácií riešiacich podobné úlohy.

Príučové slová – cloud, navigácia, neuro-fuzzy systém, robot

I. ÚVOD

Aplikácie inteligentnej robotiky prenikajú do mnohých oblastí, ako je napr. priemysel, vesmír, poľnohospodárstvo, obrana a doprava, či sociálna oblasť. Mobilný robot vykonáva mnoho úloh, ako je záchranná operácia, hliadkovanie, podvodný prieskum, pomoc pri katastrofách a planetárny prieskum atď. Základnou úlohou mobilných robotov je pritom navigácia v prostredí s prekážkami so zámerom dosiahnuť stanovený cieľ. Autonómna mobilná robotická navigácia je jednou z náročných úloh pre všetky techniky soft computing-u. Fuzzy logika a neurónová sieť sa široko používajú na navigáciu a riadenie mobilných robotov, pretože tieto metódy sú schopné zvládnuť neistotu systému. Fuzzy logika je vo všeobecnosti kombináciou fuzzy pravidiel a funkcií príslušnosti. Podobne neurónová sieť môže byť aplikovaná na rôzne typy systémov a ktorá dokáže riešiť problémy pomocou súboru empirických údajov (experimentálnych alebo predpovedaných). Neurónová sieť s fuzzy logikou zlepšuje rýchlosť rozhodovania mobilného robota pre hľadanie cieľov a vyhýbanie sa prekážkam [1]. Z tohto dôvodu je v tejto práci popísaný kaskádny neuro-fuzzy prístup [2], ktorý sa dokáže samostatne vyhnúť prekážke a bezpečne dosiahnuť cieľ v danom prostredí.

Hľadanie cieľa a vyhýbanie sa prekážkam sú dve dôležité úlohy pre každého mobilného robota v prostredí. Jednou z prvých úloh, kde bola využitá radiálna neurónová sieť (RBFNS) na riešenie plánovania cesty mobilného robota, bola práca [3], kde sa fuzzy logika v podobe fuzzy regulátora typu Takagi–Sugeno–Kang (TSK) používa na popis neurčitosti prostredia a neurónová sieť na vyladenie parametrov funkcií príslušnosti. Vývoj tohto typu prostriedku pokračoval v rozširovaní kritérií, na základ ktorých má byť cesta navrhnutá, ako sú napr. zakrivenie cesty, či ďalšie požiadavky na zabezpečenie bezpečnej jazdy autonómnych vozidiel, ako v prípade práce [4]. V [5] bol navrhnutý systém riadenia otáčok jednosmerného motora založený na mikroregulátoroch Arduino Uno pomocou viacvrstvovej neurónovej siete a PID regulátora. V [6] bola použitá biologicky inšpirovaná neurónová sieť na vytvorenie cesty bez kolízie v nestacionárnom prostredí. V [7] autori navrhli vylepšený fuzzy regulátor založený na optimalizácii pomocou mravčej kolónie (Reinforcement Ant Optimized Fuzzy Controller – RAOFC) a použili ho pre kolesového mobilného robota, ktorý sledoval stenu pomocou reinforcement learning-u. Vstupy navrhovaného regulátora sú sonarové senzory na zistenie voľného priestoru a výstupom je uhol riadenia robota. V [8] bola kombinovaná fuzzy logika s inými technikami soft computing-u, ako sú genetický algoritmus (GA), neurónové siete (NS) a optimalizácia roja častíc (Particle Swarm

Optimization – PSO) pre účely optimalizácie parametrov funkcie príslušnosti fuzzy regulátora na zlepšenie navigačného výkonu mobilného robota. Mobilná robotická navigácia založená na fuzzy posilňovaní bola predstavená spoločnosťami Beom a Cho [9] pre zložité prostredia. V [10] autori vytvorili správanie založené na neuro-fuzzy architektúre pre navigáciu mobilných robotov v neštruktúrovanom prostredí. Vývoj pokračoval ďalej a v súčasnosti sa tento prístup využíva aj pre prostredia s vysokým stupňom neurčitosti [11]. Vzhľadom na náročnosť výpočtov sa v prácach zaoberajúcich sa problematikou navigácie a riadenia dopravy stále viac využívajú prvky cloud a edge computing-u [12], ktoré sa v našej práci budeme snažiť využiť a popíšeme ich v nasledujúcej časti.

II. VYUŽITIE CLOUD COMPUTING-U PRI NAVIGÁCIÍ VOZIDIEL

V skratke je možné navrhnutú úlohu pre navigáciu autonómneho robota, ktorý sa bude samostatne vyhýbať prekážkam, popísať tak, že pomocou zosnímaných údajov o prekážkach a polohe robota budeme na cloud-e trénovať neurónovú sieť na predikciu uhla natočenia robota, ktorý je následne poslaný na vstup navigačného neuro-fuzzy systému.

A. Základné charakteristiky cloud computing-u

Cloud Computing poskytuje služby IT založené na nástrojoch. Služby sú k dispozícii ako pay-as-you-go model. Cloud poskytuje organizáciám výhodu pri nastavovaní základného hardvéru pre požiadavky používaného softvéru, t. j. namiesto nákupu hardvéru alebo softvéru sa môžu používať softvérové cloudové služby. Dostupnosť cloudových služieb kedykoľvek a kdekoľvek z neho robí uskutočniteľným riešením pre mnoho ďalších aplikácií. Cloudové služby sú obmedzené niektorými parametrami ako kvalita služieb (QoS), efektívne využitie cloudových zdrojov, užívateľského rozpočtu, užívateľských termínov, spotreby energie atď. [13].

Cloud computing je potrebné vnímať ako systém pozostávajúci z troch hlavných služieb: softvér ako služba (SaaS), infraštruktúra ako služba (IaaS) a platforma ako služba (PaaS). SaaS zahŕňa licencovanie softvérovej aplikácie zákazníkom. Licencie sa zvyčajne poskytujú prostredníctvom modelu pay-as-you-go alebo na požiadanie. Tento typ systému možno nájsť v 365 balíku Microsoft Office. IaaS zahŕňa metódu dodávania všetkého od operačných systémov po servery a úložisko prostredníctvom pripojenia založeného na IP ako súčasť služby na požiadanie. Klienti sa môžu vyhnúť potrebe nákupu softvéru alebo serverov a namiesto toho tieto zdroje obstarávať v externej službe na požiadanie. Medzi populárne príklady systému IaaS patria IBM Cloud a Microsoft Azure. PaaS sa považuje za najkomplexnejšiu z troch vrstiev cloudovej výpočtovej techniky. PaaS zdieľa určité podobnosti so SaaS, hlavným rozdielom však je, že namiesto poskytovania softvéru online je to vlastne platforma na vytváranie softvéru dodávaného prostredníctvom internetu [14]. Tento model zahŕňa platformy ako Salesforce.com a Heroku.

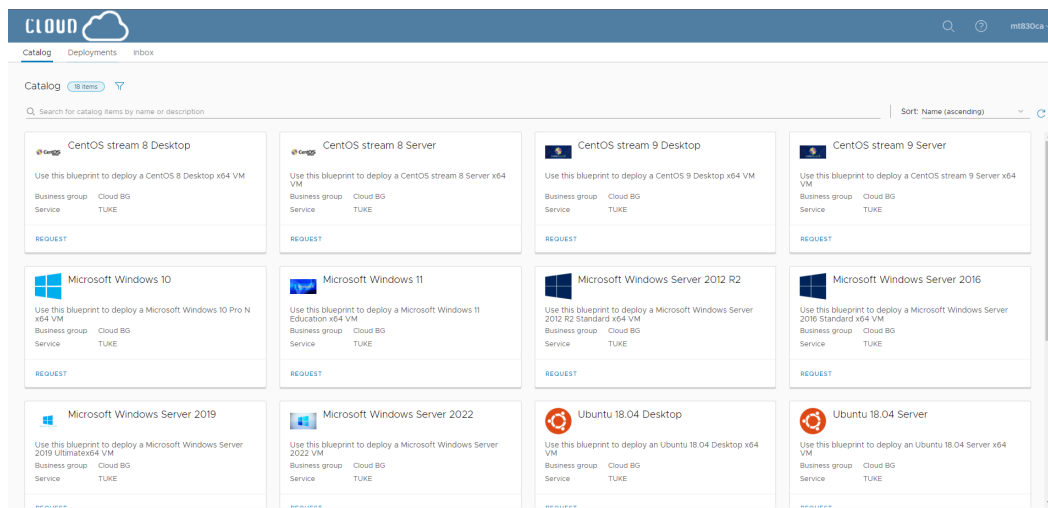
Cloudový prístup ponúka spoločnostiam zo všetkých sektorov množstvo výhod vrátane možnosti používať softvér z ľubovoľného zariadenia prostredníctvom natívnej aplikácie alebo prehliadača. Výsledkom je, že používatelia môžu prenášať svoje súbory a nastavenia na iné zariadenia úplne bezproblémovým spôsobom. Pred zavedením cloudových technológií boli spoločnosti nútené nakupovať, budovať a udržiavať nákladné technológie a infraštruktúru na správu informácií. Takto môžu ušetriť na nákladných serverových centrách a IT oddeleniach. Umožňuje tiež používateľom inovovať softvér rýchlejšie než tradičným spôsobom cez inštalácie záplat.

So všetkou rýchlosťou, efektívnosťou a inováciami, ktoré prichádzajú s cloud computingom, existujú samozrejme aj riziká. Bezpečnosť bola vždy veľkým problémom cloud-u, najmä pokiaľ ide o citlivé osobné, či finančné záznamy alebo technologické údaje. Servery spravované spoločnosťami v oblasti cloud computing-u sa môžu stať aj obeťami rôznych napr. prírodných katastrof, či ozbrojených útokov a ich výpadok môže paralyzovať rozsiahle geografické oblasti. Preto sa otázke bezpečnosti cloudových technológií venuje v súčasnosti vysoká pozornosť [15].

B. TUKE Cloud

V rámci Technickej univerzity v Košiciach (TUKE) sa taktiež nachádza vlastné cloudové riešenie – TUKE Cloud (<https://cloud.tuke.sk/>), ktorý patrí do oblasti tzv. privátnych cloud-ov a poskytuje svoje služby všetkým študentom a zamestnancom TUKE, ktorí sa môžu doň po schválení žiadosti o prístup prihlásiť pomocou svojho osobného login-u. Následne je užívateľovi umožnený prístup do rozhrania TUKE Cloud, kde je možné si okamžite vyžiadať virtuálny stroj (VM), viď Obr. 1.

Pre potreby tejto práce vzhľadom na potreby tréningu neurónovej siete bol zvolený VM s operačným systémom Windows 11 a hardvérom vybaveným osemjadrovým CPU a 8 GB RAM. Ďalej, bolo potrebné sa pripojiť pomocou remote desktop protokolu a nainštalovať potrebné

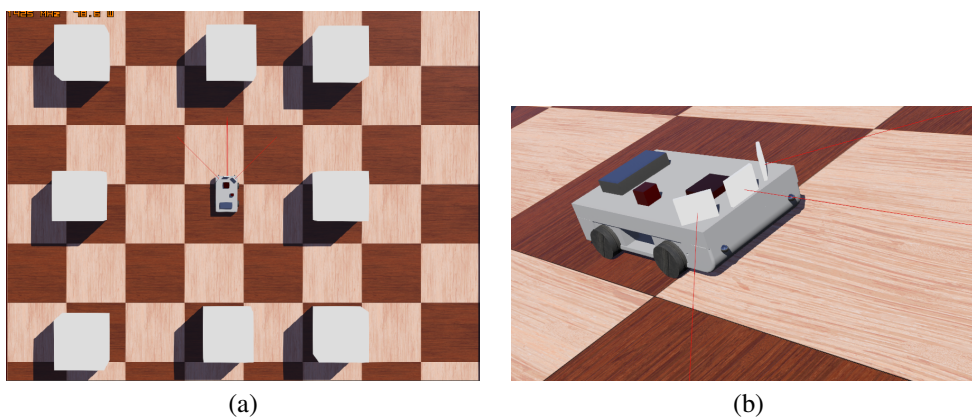


Obr. 1 Ponuka operačných systémov pre VM.

softvérové vybavenie k realizácii tréningu neurónovej siete. To zahŕňalo softvér VSCoDe, Python verzie 3.9.12 a knižnice ako je TensorFlow, NumPy a Keras. Po nainštalovaní softvérového vybavenia bolo spustené tréningu neurónovej siete a natrénovaný model bol uložený na úložisku VM pre lokálne stiahnutie a využitie v robotickom simulátore Webots [16] na predikciu uhla otáčania robota.

III. NÁVRH RIEŠENIA

Návrh navigácie mobilného robota s využitím cloud computing-u bol realizovaný na simulačných experimentoch v prostredí simulátora Webots, preto ako prvé bolo potrebné v ňom navrhnuť simulačné prostredie, v ktorom sa bude robot pohybovať, vid' Obr. 2 (a). Následne bol vytvorený model robota s tromi ultrazvukovými senzormi a štyrmi motormi na zabezpečenie otáčania sa robota, vid' Obr. 2 (b).



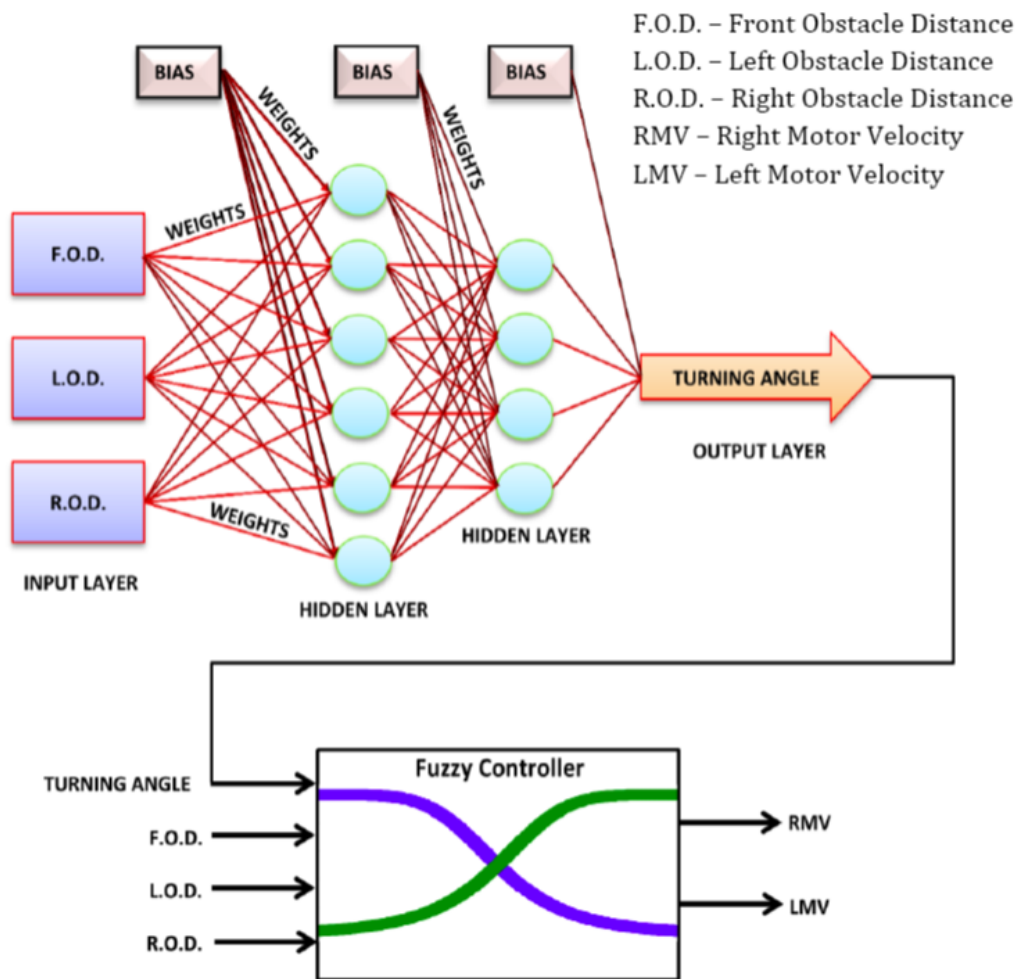
Obr. 2 Ilustračná ukážka prostredia v simulátore Webots (a) a navrhnutý model robota v simulátore Webots (b).

Pre potreby tréningu neurónovej siete bol vytvorený skript v jazyku Python na ručné ovládanie robota, pomocou ktorého boli zozbierané údaje potrebné pre tréningu ako vzdialenosť od objektu vľavo – LOD, vzdialenosť od objektu vpravo – ROD, vzdialenosť od objektu pred robotom – FOD a uhol natočenia robota.

Na odosielanie údajov bolo vytvorené s použitím knižnice Flask (v jazyku Python) vlastné API. Po úspešnom otestovaní API skriptu bol tento umiestnený do služby GitHub, ktorá má integráciu s cloudovou platformou Heroku, vďaka ktorej mohol byť realizovaný hosting API priamo u poskytovateľa služby Heroku.

Po úspešnom prijatí dát, boli tieto spracované ďalším vlastným skriptom vytvoreným v jazyku Python, v ktorom bola napokon navrhnutá NS pozostávajúca z troch vrstiev na platforme TUKE Cloud, vid' Obr. 3. Navigačný model reprezentovaný NS bol tréningu v 300 cykloch a cieľom bolo minimalizovať strednú kvadratickú odchýlku. Napokon bol navrhnutý fuzzy regulátor, do

ktorého vstupujú LOD, ROD, FOD a uhol natočenia robota, ktorý vystupuje z neurónovej siete. Výstupom je rýchlosť pravého a ľavého motora robota, t. j. RMV a LMV.



Obr. 3 Schéma použitého neuro-fuzzy systému pre potreby návrhu navigačného modelu [2].

IV. VÝSLEDKY

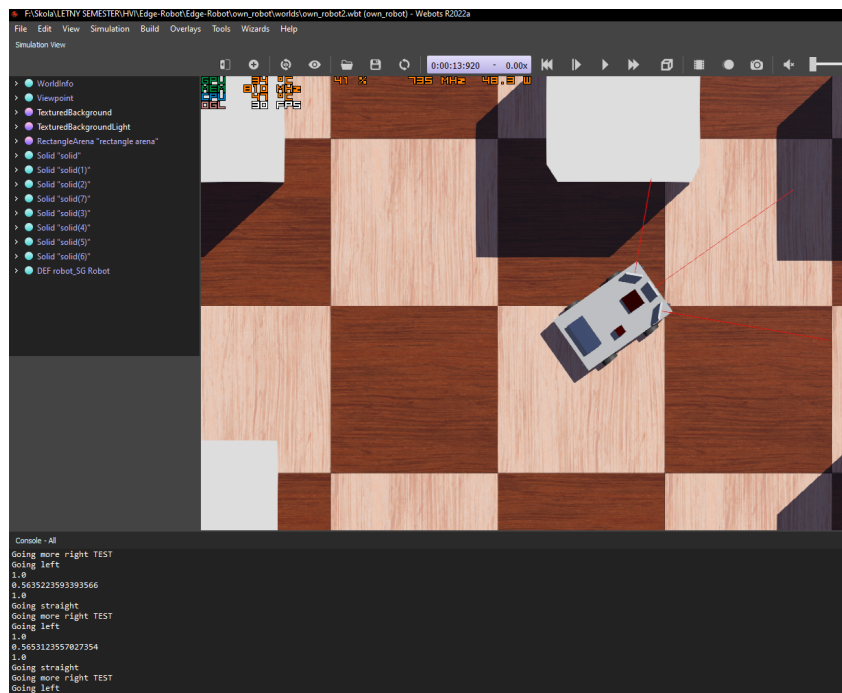
Vstupom do neurónovej siete boli štyri hodnoty zo senzorov, ktoré boli pred vstupom predspracované tak, aby sa neurónová sieť vedela správne naučiť. Počas tréningu boli vyskúšané rôzne kombinácie aktivačných funkcií, vrstiev a iných parametrov, aby boli dosiahnuté čo najlepšie výsledky. Počas procesu učenia bolo možné pozorovať kontinuálny pokles strednej kvadratickej odchýlky.

Navrhnutý systém bol úspešne otestovaný na robotovi v simulátore Webots, ktorý sa úspešne vyhýbal všetkým prekážkam nachádzajúcim sa v jeho ceste, vid' Obr. 4.

Taktiež bolo otestované pripojenie na VM v prostredí TUKE Cloud, aby bolo možné porovnať rýchlosť pripojenia v lokálnom prostredí oproti rýchlosti pripojenia na cloud-e. Priemerná rýchlosť v lokálnom prostredí bola 500 Mb/s, pričom v prostredí TUKE Cloud bola v priemere 1,2 Gbps. Najvyššia nameraná hodnota v prostredí TUKE Cloud bola 18 Gbps, čiže prístup s využitím cloud-u poskytuje v priemere minimálne dvojnásobnú rýchlosť pripojenia, než iné spôsoby sieťovej komunikácie v úlohách tohto typu.

V. ZÁVER

V tejto práci bol úspešne vytvorený simulačný model autonómneho robota na báze neuro-fuzzy systému, ktorý robotovi umožňuje sa úspešne vyhýbať prekážkam. Taktiež sa podarilo úspešne implementovať API na platforme Heroku pre odosielanie údajov do databázy a tréning neurónovej siete v prostredí TUKE Cloud. Týmto spôsobom bola aj otestovaná možnosť spolupráce aplikácií implementovaných v dvoch rozdielnych cloudových prostrediach, jednom verejnom (Heroku) a jednom privátnom (TUKE Cloud), ktoré sú súčasťou jednej spoločnej úlohy.



Obr. 4 Ukážka vyhýbania sa robota prekážkam v prostredí Webots.

V budúcnosti vzhľadom na časovú kritickosť týchto typov úloh bude potrebné sa podrobnejšie zaoberať problematikou distribúcie úloh medzi jednotlivými cloudovými službami ako aj prostriedkami počítania na hrane (edge-computing) a za týmto účelom vytvorenia tzv. *off-loading* mechanizmu pre dynamickú distribúciu výpočtov [17]. Ďalším možným rozšírením do budúcnosti by bolo vytvorenie monitorovacieho softvéru pre robota.

POĎAKOVANIE

Táto práca je finančne podporená Agentúrou na podporu výskumu a vývoja prostredníctvom projektu APVV-20-0247, ENISaC — Edge-enabled inteligentné snímanie a výpočty.

LITERATÚRA

- [1] X. Ma, X. Li, and H. Qiao, "Fuzzy neural network-based real-time self-reaction of mobile robot in unknown environments," *Mechatronics*, vol. 11, no. 8, pp. 1039–1052, 2001.
- [2] A. Pandey and K. Burse, "Cascade neuro-fuzzy architecture based mobile-robot navigation and obstacle avoidance in static and dynamic environments," *International Journal of Advanced Robotics and Automation*, vol. 1, pp. 1–9, Oct. 2016.
- [3] J. Godjevac and N. Steele, "Neuro-fuzzy control of a mobile robot," *Neurocomputing*, vol. 28, no. 1-3, pp. 127–143, 1999.
- [4] L. Yu, X. Wang, Z. Hou, Z. Du, Y. Zeng, and Z. Mu, "Path planning optimization for driverless vehicle in parallel parking integrating radial basis function neural network," *Applied Sciences*, vol. 11, no. 17, 2021.
- [5] N. Rai and B. Rai, "Neural network based closed loop speed control of dc motor using arduino uno," *International Journal of Engineering Trends and Technology*, vol. 4, no. 2, pp. 137–140, 2013.
- [6] S. X. Yang and M. Meng, "Neural network approaches to dynamic collision-free trajectory generation," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 31, no. 3, pp. 302–318, 2001.
- [7] C.-F. Juang and C.-H. Hsu, "Reinforcement ant optimized fuzzy controller for mobile-robot wall-following control," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 10, pp. 3931–3940, 2009.
- [8] M. Algabri, H. Mathkour, H. Ramdane, and M. Alsulaiman, "Comparative study of soft computing techniques for mobile robot navigation in an unknown environment," *Computers in human behavior*, vol. 50, pp. 42–56, 2015.
- [9] H. R. Beom and H. S. Cho, "A sensor-based navigation for a mobile robot using fuzzy logic and reinforcement learning," *IEEE transactions on Systems, Man, and Cybernetics*, vol. 25, no. 3, pp. 464–477, 1995.
- [10] W. Li, C. Ma, and F. M. Wahl, "A neuro-fuzzy system architecture for behavior-based control of a mobile robot in unknown environments," *Fuzzy Sets and systems*, vol. 87, no. 2, pp. 133–140, 1997.
- [11] M. Goharimanesh, A. Mehrkish, and F. Janabi-Sharifi, "A fuzzy reinforcement learning approach for continuum robot control," *J. Intell. Robotic Syst.*, vol. 100, no. 3, pp. 809–826, 2020.
- [12] S. Liu, L. Liu, J. Tang, B. Yu, Y. Wang, and W. Shi, "Edge computing for autonomous driving: Opportunities and challenges," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1697–1716, 2019.
- [13] A. Montazerolghaem, M. H. Yaghmaee, and A. Leon-Garcia, "Green cloud multimedia networking: Nfv/sdn based energy-efficient resource allocation," *IEEE Transactions on Green Communications and Networking*, vol. 4, no. 3, pp. 873–889, 2020.
- [14] L. Qian, Z. Luo, Y. Du, and L. Guo, "Cloud computing: An overview," in *Cloud Computing*, M. G. Jaatun, G. Zhao, and C. Rong, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 626–631.

- [15] D.-A. Pandey, R. Boddu, M. Tiwari, and T. Tiwari, "An analysis of data security and privacy in cloud computing," in *Recent Trends in Applied Sciences and Computing Engineering*. Association of Global Academician and Researchers, 2021, pp. 1481–1535.
- [16] Webots, "http://www.cyberbotics.com," open-source Mobile Robot Simulation Software. [Online]. Available: <http://www.cyberbotics.com>
- [17] D. Herich, J. Vaščák, I. Zolotová, and A. Brecko, "Automatic path planning offloading mechanism in edge-enabled environments," *Mathematics*, vol. 9, no. 23, 2021.

Návrh mobilnej aplikácie pre vizualizáciu poveternostných podmienok v monitorovanom priestore

¹Luboš OVSENÍK, ²Jakub ORAVEC, ³Samuel ANDREJČÍK, ⁴Ján GIŇOVSKÝ

^{1,2,3,4} Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹lubos.ovsenik@tuke.sk, ²jakub.oravec@tuke.sk, ³samuel.andrejck@tuke.sk,
⁴jan.ginovsky@student.tuke.sk

Abstrakt — Príspevok sa zaoberá návrhom mobilnej aplikácie určenej pre vizualizáciu a analýzu dát získaných prostredníctvom senzorov monitorujúcich poveternostné podmienky v ľubovoľnom skúmanom priestore. Cieľom tejto práce bolo vytvoriť intuitívne a používateľsky prívetivé rozhranie mobilnej aplikácie v platforme Android Studio, ktorá poskytuje údaje o počasí vizuálne prístupným a informatívnym spôsobom. Navrhovaná aplikácia využíva moderné technológie, techniky vizualizácie údajov a zásady dizajnu zameraného na človeka s cieľom poskytnúť používateľovi lepší zážitok.

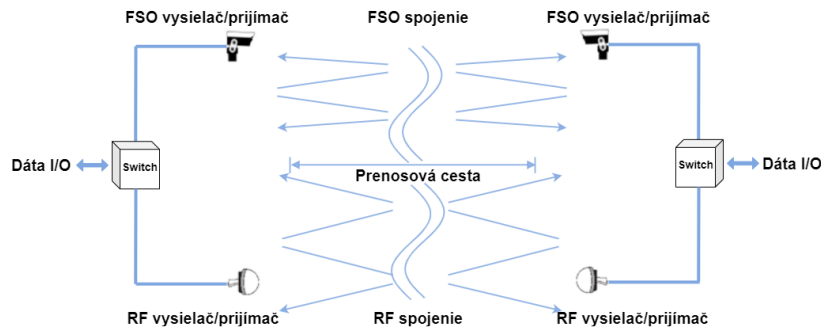
Kľúčové slová — Android, FSO, hybridný FSO/RF systém, mobilná aplikácia, RF

I. ÚVOD

V súčasnej dobe sa zvyšuje dopyt po monitorovaní poveternostných podmienok v rôznych priestoroch, ako napríklad priemyselných areáloch alebo mestských oblastiach. Vzhľadom na to, že rôzne systémy pre prenos dát, ako napríklad FSO (Free Space Optic), RF (Radio Frequency) a hybridné FSO/RF systémy, sú závislé na kvalite poveternostných podmienok, je nevyhnutné mať spoľahlivý spôsob vizualizácie týchto podmienok. Cieľom tejto práce bol návrh a implementácia mobilnej aplikácie, ktorá umožní vizualizáciu poveternostných podmienok v monitorovanom priestore. Aplikácia bola vytvorená s použitím knižnice MPAndroidChart a platformy Android Studio, ktoré nám umožnili vytvoriť prehľadnú a funkčnú aplikáciu. Vďaka tejto knižnici sme boli schopní vizualizovať poveternostné podmienky a sledovať vplyv prostredia na kvalitu prenášaného signálu RSSI (Received Signal Strength Indicator). Táto mobilná aplikácia je vhodná aj pre menej skúsených užívateľov, pretože jej použitie je intuitívne a jednoduché. Aplikácia bude schopná zobrazovať históriu nameraných dát z meteorologickej stanice o poveternostných podmienkach, ako sú napríklad teplota, vlhkosť a rýchlosť vetra. Okrem toho, aplikácia bude umožňovať aj zobrazovanie údajov o kvalite signálu RSSI.

II. HYBRIDNÝ FSO/RF PRENOSOVÝ SYSTÉM

Predstavovaná aplikácia pre analýzu dát získaných prostredníctvom senzorov monitorujúcich poveternostné podmienky počasia slúži k správne nastaveniu vhodnej prenosovej cesty v hybridnom FSO/RF prenosovom systéme. FSO je systém pre optický prenos dát voľným prostredím (vzduchom), ktorý funguje na podobnom princípe ako prenos dát optickými vláknami a teda tak, že informácie sú prenášané pomocou lúčov svetla. Hlavným rozdielom medzi týmito systémami je prostredie, v ktorom je tento prenos realizovaný. Zatiaľ čo optické vláknové spoje na prenos používajú najčastejšie optické vlákna na báze SiO₂, ktorých nevýhodou je ich cena a náročná inštalácia, FSO používa atmosférický nelicencovaný kanál, vďaka čomu je tento systém lacnejší a ľahko inštalovateľný pri porovnateľných prenosových rýchlostiach a šírke prenosového pásma. Nevýhodou je však nutnosť priamej viditeľnosti medzi FSO hlavicami a pomerne krátkej vzdialenosti (cca 20 km), na ktorej operujú. Vplyv počasia brániaci priamej viditeľnosti tak môže mať negatívne účinky. Doplnením systému prenosu dát voľným prostredím o prenos dát rádiovým RF signálom vzniká hybridný FSO/RF systém (Obr. 1), vyzdvihujúci dobré vlastnosti jednotlivých systémov, čím sa stáva odolnejším voči vplyvom počasia [1, 3].



Obr. 1 Hybridný FSO/RF systém.

Hybridná FSO/RF komunikácia je implementovaná tak, aby kombinovala výhody oboch systémov pre zabezpečenie spoľahlivého rýchleho širokopásmového komunikačného spojenia. Aby sme neznehodnotili výhody FSO systému, je potrebné nájsť kompromis v nastaveniach hybridnej siete. RF linka by mala byť dostupná hneď v momente, ako sa FSO stáva nedostupnou. Šírka pásma RF linky by mala dosahovať približné hodnoty FSO linky [2, 4]. Veľké rozdiely v šírke pásma by spôsobovali zníženie prenosovej rýchlosti pri prepnutí z FSO na RF linku. Neexistuje však „najlepšie“ nastavenie hybridného prenosového systému, všetko to závisí od konkrétnej situácie a požiadaviek, ako aj konkrétneho miesta implementácie. Napríklad kombinácia FSO/RF s vysokými prenosovými rýchlosťami, ktoré si vyžadujú veľké nosné frekvencie nie sú odporúčané v kontinentálnom podnebí, pretože sú obe ovplyvňované hmlou, dažďom, či snehom. Naopak, implementáciou tejto kombinácie v tropickom suchom podnebnom pásme môžeme získať stabilnú komunikačnú sieť s vysokými prenosovými rýchlosťami. Komerčné RF linky sa pohybujú vo frekvenčných pásmach od 60 GHz do 80 GHz. Frekvencie okolo 60 GHz sú nelicencované, no zároveň nemôžu byť použité na väčšie vzdialenosti kvôli vysokému útlmu ovzduším, konkrétne kyslíkom, ktorý spôsobuje absorpciu až 15 dB/km. Frekvencie od 70 GHz až 80 GHz RF linky používané modernejšími zariadeniami dokážu prenášať dáta na vzdialenosti do 3 km rýchlosťou 1 Gb/s nezávisle od počasia, avšak tieto frekvencie bývajú dostupné až po získaní licencie. Použitím nižších frekvencií, ktoré nevyžadujú licencie sa znižujú aj prenosové rýchlosti, no výhodou je malý vplyv atmosférických javov na prenos [1, 5].

III. NÁVRH MOBILNEJ APLIKÁCIE

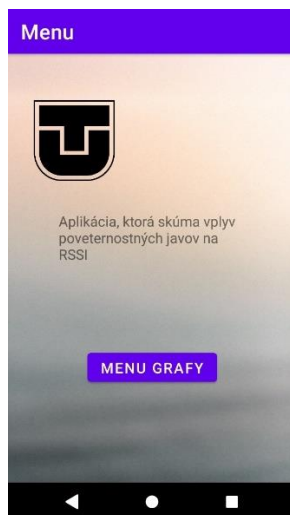
Mobilná aplikácia bola navrhnutá s použitím knižnice MPAndroidChart, ktorá umožňuje vizualizovať poveternostné podmienky v monitorovanom priestore. Android je open-source operačný systém, ktorý beží na linuxovom jadre a jeho aplikácie sú vyvíjané pomocou jazyka Java.

Mobilná aplikácia sa skladá z troch častí. V prvej časti je úvodná obrazovka, ktorá slúži na uvítanie užívateľa a na jej pozadí bežia procesy na stiahnutie dát z internetu. V druhej časti sa nachádza menu s možnosťou výberu grafov. V tretej časti sa nachádza zobrazenie grafov. Toto bolo realizované pomocou voľne dostupnej knižnice MPAndroidChart, ktorá slúži na vykreslenie získaných dát v podobe grafu.

A. Úvodná obrazovka

Ak spustíme aplikáciu tak môžeme vidieť úvodnú obrazovku (Obr. 2), na ktorej je logo technickej univerzity, krátky text v ktorom je stručne napísané, načo je mobilná aplikácia určená a tlačidlo s nápisom MENU GRAFY. Po kliknutí na toto tlačidlo sa nám otvorí druhá časť aplikácie, kde sa nachádza menu so všetkými grafmi, ktoré chceme zobraziť.

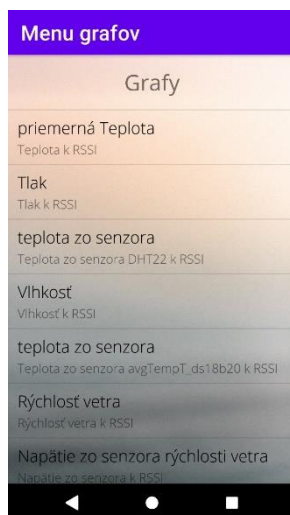
Na pozadí úvodnej obrazovky pri spustení aplikácie beží proces v ktorom sa skontroluje, či je vytvorený súbor s názvom data1.json. Ak tento súbor neexistuje, tak sa vytvorí. Ak už existuje, tak sa kontrola preskočí. To je kontrola, či je mobilné zariadenie pripojené k internetu. Ak nie sme pripojení k sieti tento krok sa preskočí a použijú sa už uložené dáta. No ak sme pripojení k internetu nasleduje stiahnutie dát z URL adresy.



Obr. 2 Úvodná obrazovka aplikácie.

B. Menu s grafmi

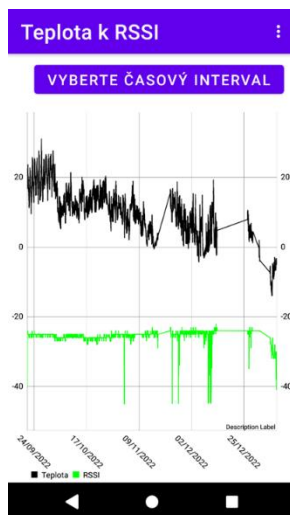
Po kliknutí na tlačidlo MENU GRAFY sa nám otvorí trieda GraphBar. V tejto triede nie je nič zaujímavé. Nachádza sa tu len menu grafov, kde sú zobrazené všetky triedy grafov (Obr. 3). Po kliknutí napríklad na „priemerná Teplota“ sa nám otvorí a spustí trieda, ktorá nám vykreslí graf v ktorom bude zobrazené ako vplýva teplota na RSSI použitého FSO kanála (Obr. 4).



Obr. 3 Menu grafov.

C. Zobrazenie grafov

V tejto časti sa grafy vykresľujú a zobrazia sa získané dáta, ktoré sme získali z URL adresy. Na vykreslenie grafov bola použitá voľne dostupná knižnica MPAndroidChart. MPAndroidChart je knižnica grafických komponentov pre Android, ktorá umožňuje tvorbu a zobrazovanie rôznych typov grafov v aplikácii. Knižnica podporuje množstvo typov grafov vrátane stĺpcových grafov, čiarových grafov, koláčových grafov, radarových grafov a mnohých ďalších. MPAndroidChart umožňuje prispôbenie grafu pomocou rôznych konfiguračných možností, ako napríklad nastavenie farieb, štýlu a popisov grafu. Knižnica podporuje interakciu s používateľom, vrátane funkcií ako zoomovanie. Knižnica MPAndroidChart má tiež podporu pre animácie, ktoré umožňujú plynulé prechody medzi rôznymi hodnotami grafu. Knižnica je k dispozícii ako súčasť maven alebo Gradle buildu a jej použitie nevyžaduje žiadne zvláštne oprávnenia. Knižnica umožňuje tvorbu rôznych typov grafov, ktoré môžu byť prispôbené podľa požiadaviek aplikácie [6]. Preto bola táto knižnica zvolená ako vhodný výber pre vytvorenie mobilnej aplikácie na vizualizáciu poveternostných podmienok. Umožňuje nám vykresliť viac ako len jednu hodnotu na grafe.

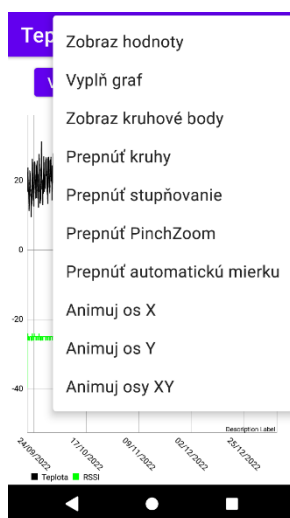


Obr. 4 Vykreslenie grafu „Vplyv teploty na RSSI“.

Aplikácia taktiež ponúka možnosť upraviť graf. Metóda `onCreateOptionsMenu` vytvára menu, ktoré sa zobrazí v hornej časti obrazovky a obsahuje rôzne možnosti pre úpravu grafu (Obr. 5). Menu obsahuje nasledujúce položky:

- Zobraz hodnoty- umožňuje zobrazit' alebo skryt' hodnoty (textové popisy) na grafe.
- Vyplň graf- umožňuje zobrazit' alebo skryt' farebné výplne pod krivkami na grafe.
- Zobraz kruhové body- umožňuje zobrazit' alebo skryt' kruhové body, ktoré označujú jednotlivé dáta na grafe.
- Prepnúť kruhy- umožňuje prepnúť medzi cubic a lineárnou interpoláciou medzi bodmi na grafe.
- Prepnúť stupňovanie- umožňuje prepnúť medzi schodovitou a lineárnou interpoláciou medzi bodmi na grafe.
- Prepnúť PinchZoom- umožňuje zapnúť alebo vypnúť pinch-to-zoom funkcionality na grafe.
- Prepnúť automatickú mierku- umožňuje zapnúť alebo vypnúť automatické prispôsobenie osí grafovým dátam.
- Animuj os X- spúšťa animáciu posúvania osi X grafu.
- Animuj os Y- spúšťa animáciu posúvania osi Y grafu.
- Animuj osy XY- spúšťa animáciu posúvania oboch osí grafu.

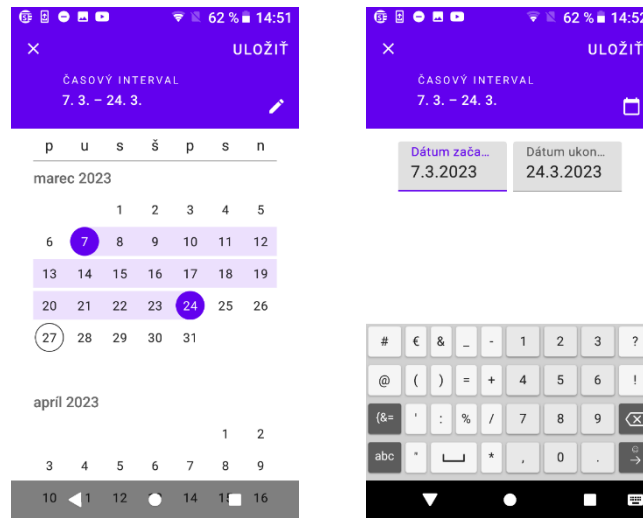
V podstate ide o rôzne nastavenia a animácie, ktoré môžu používatelia použiť na úpravu vzhľadu a správania grafu v aplikácii.



Obr. 5 Menu možností ako vylepšiť graf.

Napríklad úprava (výber) časového intervalu znamená, že si zvolíme časový úsek, v ktorom chceme vizualizovať vybrané údaje. Na výber máme z dvoch možností ako to realizovať. Buď sa v kalendári zvolí časový úsek tak, že sa klikne na nami požadovaný dátum, to je dátum začiatku, a tak sa klikne na ďalší dátum, to je dátum ukončenia, to zobrazí časovú periódu (Obr. 6a). Ďalšia

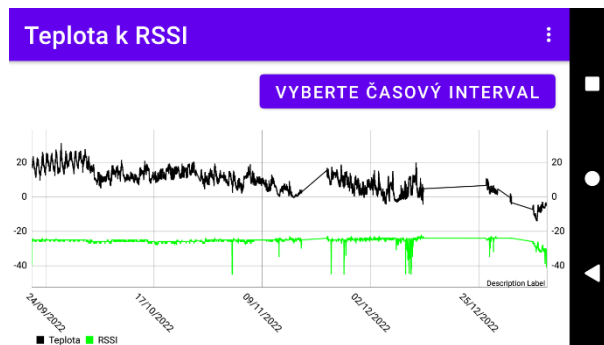
možnosť je, že do dvoch polí vpíšeme dátum začiatku a dátum ukončenia (Obr. 6b). V oboch prípadoch klikneme na tlačidlo uložiť. Po odkliknutí sa v grafe zobrazí nami zvolený časový úsek. Časový úsek obsahuje dáta pre zvolenú časovú periódu. Túto funkciu sme realizovali pomocou knižnice Material Components for Android, konkrétne pomocou komponentu MaterialDatePicker.



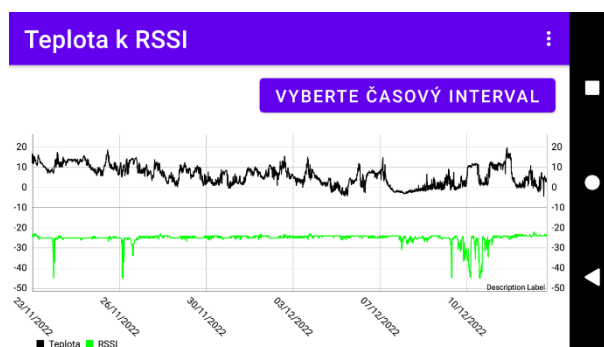
Obr. 6 Výber časového úseku: a) pomocou kalendára, b) pomocou polí.

Na základe získaných dát z meteorologickej stanice sa dajú vytvorili grafy, ktoré nám umožňujú vizuálne sledovať, ako sa menia hodnoty "RSSI" v závislosti od meteorologických podmienok. V grafoch sme zaznamenali niektoré zaujímavé vzory a trendy, ktoré poukazujú na to, že niektoré meteorologické podmienky môžu mať významný vplyv na hodnoty "RSSI". Skúmané časové obdobie bolo od 21.9.2022 až po 9.1.2023. Podrobnejšie bolo preskúmané časové obdobie 23.11.2022-15.12.2022 a 9.12.2022-13.12.2022, kde bolo viditeľné, že RSSI malo na grafe najväčšie výkyvy.

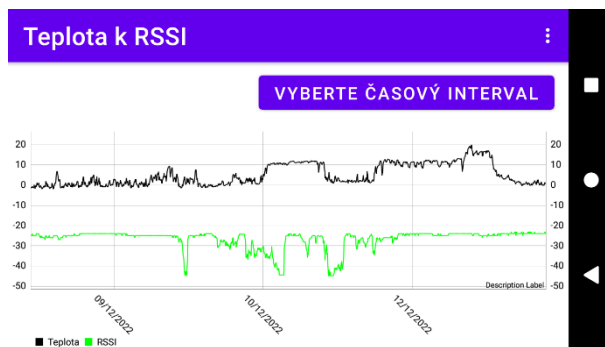
Na obrázkoch (Obr. 7, Obr. 8 a Obr. 9) je vizualizovaná závislosť zmeny RSSI od teploty v sledovaných obdobiach.



Obr. 7 Vplyv teploty na RSSI od 21.9.2022 do 9.1.2023.

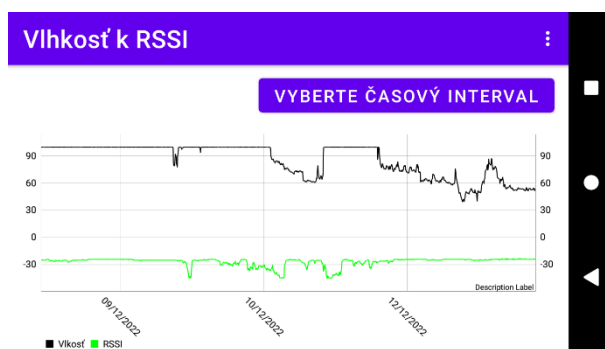


Obr. 8 Vplyv teploty na RSSI od 23.11.2022 do 15.12.2022.



Obr. 9 Vplyv teploty na RSSI od 9.12.2022 do 13.12.2022.

Na Obr. 10 je vplyv vlhkosti na RSSI v období 9.12.2022-13.12.2022.



Obr. 10 Vplyv vlhkosti na RSSI od 9.12.2022 do 13.12.2022.

IV. ZÁVER

Výsledky prezentovanej práce ukazujú, že mobilná aplikácia pre vizualizáciu poveternostných podmienok v monitorovanom priestore môže byť veľmi užitočná pre monitorovanie a zlepšovanie kvality prenášaného dátového signálu v rôznych podmienkach. Použitím FSO a RF technológií v hybridnom systéme je dokázané, že kombinácia týchto technológií môže viesť k výraznému zlepšeniu prenosových rýchlostí a spoľahlivosti prenosu dát. Mobilná aplikácia bola navrhnutá s použitím knižnice MPAndroidChart, ktorá umožňuje vizualizovať poveternostné podmienky v monitorovanom priestore a sledovať vplyv prostredia na RSSI. Použitie platformy Android Studio a jazyka Java nám umožnilo vytvoriť funkčnú a prehľadnú aplikáciu, ktorá je ľahko použiteľná aj pre menej skúsených užívateľov.

POĎAKOVANIE

Táto práca bola podporovaná nasledujúcimi výskumnými projektami: APVV-17-0208 – „Odolné mobilné siete na doručovanie obsahu“ a VEGA 1/0260/23 – „Umelá inteligencia, robustné MANET a viacpreskokové D2D siete integrované do siete 6.generácie“.

LITERATÚRA

- [1] M. N. Khan, et al., “Maximizing throughput of hybrid FSO-RF communication system: An algorithm”, in *IEEE Access*, Vol. 6, 2018, pp. 30039-30048.
- [2] M. A. Amirabadi, et al., “A novel hybrid FSO/RF communication system with receive diversity”, in *Optic*, Vol. 184, 2019, pp. 293-298.
- [3] A. Munandar, et al., “Design of real-time weather monitoring system based on mobile application using automatic weather station”, in *2nd International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology (ICACOMIT 2017)*, IEEE, 2017, pp. 44-47.
- [4] A. Sampathkumar, et al., “A novel scheme for an IoT-based weather monitoring system using a wireless sensor network”, in *Integration of WSN and IoT for smart cities*, 2020, pp. 181-191.
- [5] S. Dewitte, et al., “Artificial intelligence revolutionises weather forecast, climate monitoring and decadal prediction”, in *Remote Sensing*, Vol. 13, no 16 (2021), pp. 3209.
- [6] P. Jahoda, “MPAndroidChart”, in <https://weeklycoding.com/mpandroidchart/>.

Návrh web rozhrania pre vizualizáciu poveternostných podmienok v monitorovanom priestore

¹Luboš OVSEŇÍK, ²Jakub ORAVEC, ³Samuel ANDREJČÍK, ⁴Richard PRIBULA

^{1,2,3,4} Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹lubos.ovsenik@tuke.sk, ²jakub.oravec@tuke.sk, ³samuel.andrejczik@tuke.sk, ⁴richard.pribula@student.tuke.sk

Abstrakt — Príspevok sa zaoberá problematikou návrhu webového rozhrania na vizualizáciu poveternostných podmienok v monitorovanej oblasti. Cieľom tejto práce bolo vytvoriť intuitívne a používateľsky prívetivé webové rozhranie, ktoré poskytuje údaje o počasi vizuálne prítlačným a informatívnym spôsobom. Navrhované webové rozhranie využíva moderné webové technológie, techniky vizualizácie údajov a zásady dizajnu zameraného na človeka s cieľom poskytnúť používateľovi lepší zážitok.

Kľúčové slová — FSO, hybridný FSO/RF systém, RF, webová aplikácia, webové rozhranie

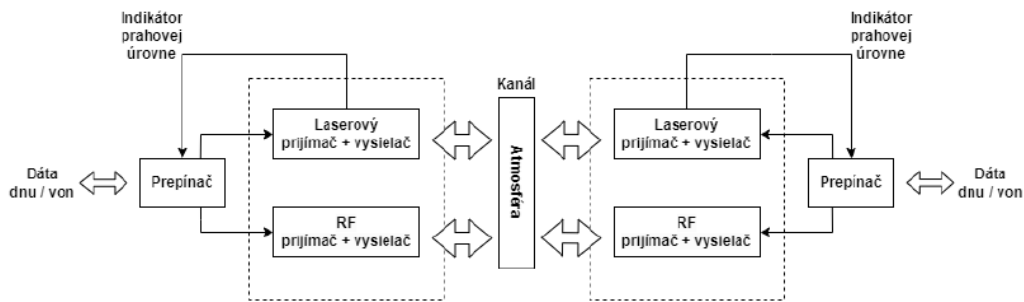
I. ÚVOD

Monitorovanie a vizualizácia počasia hrá veľký význam pri prenose dát vzduchom, obzvlášť pomocou hybridných FSO/RF komunikačných systémoch. S príchodom pokročilých technológií monitorovania počasia a dostupnosťou obrovského množstva údajov o počasi sa potreba efektívnych webových rozhraní na vizualizáciu poveternostných podmienok v monitorovanej oblasti stala prvoradou. V tejto práci sme sa sústredili na vytvorenie webového prostredia na vizualizáciu a kontrolu nameraných dát z monitorového priestoru. Naš výskum zahŕňal štúdium hybridných FSO/RF komunikačných systémov, aký vplyv majú zmeny počasia na prenosový kanál v takomto komunikačnom systéme. Webové rozhranie je navrhnuté tak, aby zobrazovalo rôzne parametre počasia, ako sú teplota, vlhkosť, viditeľnosť, rýchlosť vetra, silu prijatého signálu a atmosférický tlak pomocou interaktívnych grafov. Rozhranie umožňuje používateľom vybrať konkrétne časové obdobia a parametre počasia na vizualizáciu a poskytuje možnosti prispôsobenia a porovnania údajov. V práci je použitý React.js ako front-end framework, Node.js ako back-end framework a MySQL ako databázový systém na ukladanie údajov o počasi. Webové rozhranie je nasadené na linuxovom serveri, na ktorom je webová aplikácia umiestnená [1,2,3].

II. HYBRIDNÉ FSO/RF SYSTÉMY

Hybridné systémy FSO/RF majú oproti samostatným systémom FSO alebo RF niekoľko výhod. Komunikácia FSO je veľmi citlivá na atmosférické podmienky, ako je hmla, dážď a sneh, ktoré môžu spôsobiť výrazný útlm signálu a zhoršiť kvalitu komunikácie. Na druhej strane, RF komunikácia je menej ovplyvnená atmosférickými podmienkami a môže poskytovať spoľahlivú komunikáciu na veľké vzdialenosti [4]. Kombináciou týchto dvoch technológií môžu hybridné systémy FSO/RF využiť silné stránky oboch technológií na zabezpečenie spoľahlivej a vysoko výkonnej bezdrôtovej komunikácie. V takýchto systémoch poskytuje RF spojenie záložný komunikačný kanál pre spojenie FSO, čo pomáha zmierniť vplyv atmosférických podmienok na spojenie FSO [5]. Niekoľko výskumných štúdií skúmalo použitie hybridných komunikačných systémov FSO/RF v rôznych aplikáciách, ako sú bezdrôtové senzorové siete, záloha pre bunkové siete a komunikácia medzi bodmi. Tieto štúdie ukázali, že hybridné systémy FSO/RF môžu zlepšiť spoľahlivosť a výkonnosť komunikácie v porovnaní so samostatnými systémami FSO alebo RF. Celkovo je RF komunikácia rozhodujúcou bezdrôtovou komunikačnou technológiou

používanou v rôznych aplikáciách. V spojení s inými technológiami, ako je FSO, môže poskytnúť ešte väčšie výhody a umožniť spoľahlivú a vysoko výkonnú bezdrôtovú komunikáciu v náročných prostrediach [5].

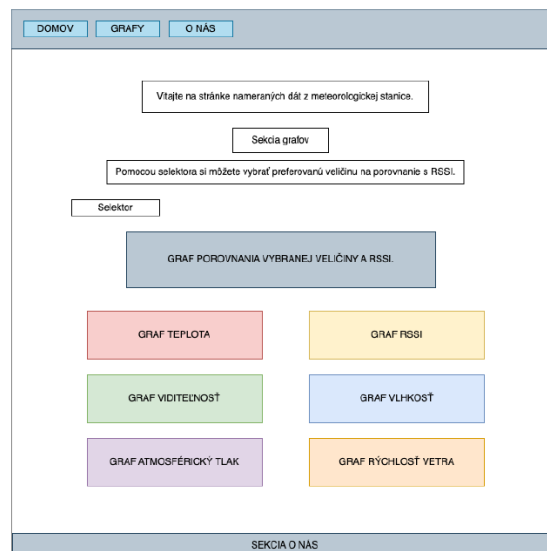


Obr. 1 Hybridný FSO/RF komunikačný systém.

Hybridné systémy FSO/RF však nie sú bez problémov. Jednou z hlavných výziev je navrhnuť systém tak, aby spojenia FSO a RF fungovali v tandeme a poskytovali spoľahlivú a vysoko výkonnú komunikáciu. To si vyžaduje dôkladné zváženie faktorov, ako je prenosový výkon a citlivosť spojov FSO a RF, použitá modulačná schéma a mechanizmus prepínania medzi oboma spojmi. Ďalšou výzvou je synchronizácia oboch spojov. Keďže spojenia FSO a RF pracujú na rôznych frekvenciách a majú rôzne charakteristiky šírenia, je nevyhnutné zabezpečiť, aby signály prichádzajúce do prijímača boli správne synchronizované. V opačnom prípade môže dôjsť k skresleniu alebo strate prijímaného signálu, čo vedie k slabšej komunikačnej výkonnosti [4]. Napriek týmto problémom sa hybridné systémy FSO/RF ukázali ako veľmi sľubné v rôznych aplikáciách, najmä v prostrediach, kde tradičné bezdrôtové komunikačné technológie nemusia byť účinné. Napríklad v mestských oblastiach s vysokými budovami môžu hybridné systémy FSO/RF poskytovať vysoko rýchlostnú a spoľahlivú bezdrôtovú komunikáciu tam, kde iné technológie môžu byť nepraktické alebo príliš drahé. Celkovo možno povedať, že kombinácia komunikačných technológií FSO a RF v hybridnom systéme ponúka prístup k bezdrôtovej komunikácii, ktorá môže zlepšiť výkon a spoľahlivosť, najmä v náročných prostrediach [5]. Na Obr. 1 je znázornená bloková schéma hybridného FSO/RF komunikačného systému.

III. WEBOVÁ APLIKÁCIA

Na začiatku tvorby webovej aplikácie sme si určili, že chceme mať na vrchu aplikácie navigačné tlačidlá na navigáciu v našej webovej aplikácii, pod navigačnými tlačidlami text s úvodnými informáciami čo užívateľ nájde na našej webovej aplikácii, o aké dáta sa jedná a ako s nimi nakladať. Ďalej sa nachádza sekcia grafov, ktorá reprezentuje hlavnú časť celej aplikácie. V tejto sekcii sú zobrazované jednotlivé grafy našich sledovaných veličín a selektor na prepínanie medzi grafmi, ktoré zobrazujú porovnanie hodnoty RSSI (Received Signal Strength Indicator) a vybranej veličiny akú si užívateľ zvolí. Na Obr. 2 je možné vidieť diagram usporiadania jednotlivých prvkov vo webovej aplikácii.



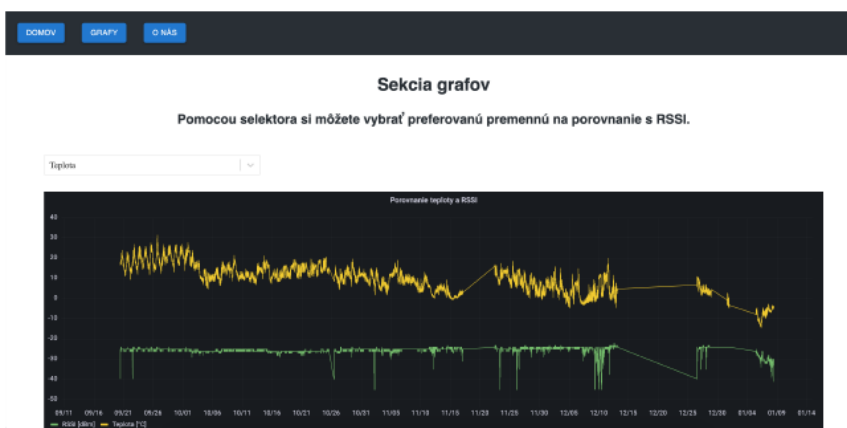
Obr. 2 Diagram usporiadania prvkov vo webovej aplikácii.

Na Obr. 3 je vidieť vrchnú časť webovej aplikácie kde je popísané čo môže užívateľ v aplikácii nájsť.



Obr. 3 Ukážka vrchnej časti webovej aplikácie.

Na Obr. 4 je zobrazený začiatok sekcie grafov, konkrétne pohľad na graf porovnanie teploty a hodnoty RSSI, ktorý sa zobrazuje základne pri otvorení aplikácie. Pomocou selektora si užívateľ môže zvoliť akú premennú chce s RSSI porovnávať a po zvolení sa graf vygeneruje na tom istom mieste ako predchádzajúci graf.



Obr. 4 Ukážka začiatku sekcie grafov spolu so selektorom na porovnanie veličín.

Pod grafom porovnania veličín s hodnotou RSSI sa nachádzajú samostatné grafy kde reprezentujeme jednotlivé veličiny. Užívateľ si môže podľa potreby bližšie nastaviť presne aké časové okno potrebuje. Na obrázkoch (Obr. 5, Obr. 6 a Obr. 7) sú prezentované grafy pre všetky samostatné veličiny reprezentované v čase.



Obr. 5 Ukážka grafov pre reprezentáciu teploty a RSSI.



Obr. 6 Ukážka grafov pre reprezentáciu vlhkosti a atmosférického tlaku.



Obr. 7 Ukážka grafov pre reprezentáciu rýchlosti vetra a viditeľnosti.

Na konci nášho vývoja webovej aplikácie sme aplikáciu presunuli na náš Linux server aby bola prístupná online, nie len v našom lokálnom počítači. Vývoj celej aplikácie sme robili v aplikácii Visual Studio Code, ktorá podporuje množstvo prídavných modulov a pomocou jedného z nich, konkrétne Remote – SSH od spoločnosti Microsoft sme nadviazali SSH spojenie s našim serverom. Pomocou tohto modulu sme si v aplikácii Visual Studio Code otvorili súborový systém nášho servera a jednoducho presunuli celú aplikáciu na server.

IV. ZÁVER

Pri hodnotení vyvinutej webovej aplikácie na vizualizáciu nameraných údajov z monitorového priestoru, sme využili viaceré kritéria. Z pohľadu výkonnosti je webová aplikácia veľmi rýchla, keďže bola vyvinutá jedným z najnovších JavaScript frameworkom React.js. Užívateľ má dojem, že sa pohybuje po reálnej aplikácii, nie po webovej stránke, pretože po prístupe na webovú aplikáciu, sa celá pred-načíta na strane užívateľa, a preto užívateľ nemusí čakať za serverom, keď sa chce presunúť na inú časť aplikácie. Užívateľské prostredie je veľmi jednoduché a prehľadné, čo dáva používateľovi pocit istoty pri práci vo webovej aplikácii. Jednou z nevýhod je vlastnosť aplikácie Grafana, ktorá sa stará o vykresľovanie grafov, že po tom ako si užívateľ dopytuje iné časové okno pre daný skúmaný graf, musí tento graf aplikácia Grafana na strane servera z našich dát vygenerovať a vrátiť užívateľovi. Čo sa týka bezpečnosti, je aplikácia zabezpečená viacerými heslami pre vstup do MySQL databázy, do aplikácie Grafana a aj pre vstup na Linux server. Všade sa využívajú iné heslá. Jednou z požiadaviek bolo aby sa na webovú aplikáciu dalo dostať iba cez univerzitnú sieť, čo sme aj dodržali. Táto práca prináša veľký prínos pre účely výskumu dopadu poveternostných podmienok a celkovo zmien počasia na FSO/RF komunikačné systémy. Pomocou tejto aplikácie je možné tieto nadväznosti medzi zmenami počasia a zmenou hodnôt RSSI jasne identifikovať a zaznamenať.

POĎAKOVANIE

Táto práca bola podporovaná nasledujúcimi výskumnými projektami: APVV-17-0208 – „Odolné mobilné siete na doručovanie obsahu“ a VEGA 1/0260/23 – „Umelá inteligencia, robustné MANET a viacpreskokové D2D siete integrované do siete 6.generácie“.

LITERATÚRA

- [1] A. Munandar, et al., “Design of real-time weather monitoring system based on mobile application using automatic weather station”, in *2nd International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology (ICACOMIT 2017)*, IEEE, 2017, pp. 44-47.
- [2] A. Sampathkumar, et al., “A novel scheme for an IoT-based weather monitoring system using a wireless sensor network”, in *Integration of WSN and IoT for smart cities*, 2020, pp. 181-191.
- [3] S. Dewitte, et al., “Artificial intelligence revolutionises weather forecast, climate monitoring and decadal prediction”, in *Remote Sensing*, Vol. 13, no 16 (2021), pp. 3209.
- [4] M. N. Khan, et al., “Maximizing throughput of hybrid FSO-RF communication system: An algorithm”, in *IEEE Access*, Vol. 6, 2018, pp. 30039-30048.
- [5] M. A. Amirabadi, et al., “A novel hybrid FSO/RF communication system with receive diversity”, in *Optic*, Vol. 184, 2019, pp. 293-298.

Identifikácia kandidátov na neologizmy na báze slovenského internetového doménového priestoru

Petra JANKUROVÁ, Ján GENČI

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

petra.jankurova@student.tuke.sk, jan.genci@tuke.sk

Abstrakt — Cieľom práce bolo na základe textov zo slovenského doménového priestoru nájsť kandidáty na slovenské neologizmy. Na základe vykonanej analýzy existujúcich metód boli navrhnuté a implementované prístupy pre automatizovanú identifikáciu kandidátov na slovenské neologizmy. Článok opisuje postup získania a spracovania vstupných dát, ako aj zvolené metódy identifikácie neologizmov a ich aplikáciu na tieto dáta. Použitými krokmi identifikácie neologizmov boli porovnanie reťazcov, analýza frekvencií výskytu, aplikácia Levenshteinovej vzdialenosti, identifikácia neologizmov na základe metalingvistických značiek a manuálna analýza. Výsledky vykonaných experimentov dokázali výskyt 705 potenciálnych neologizmov v internetovom priestore.

Kľúčové slová — doménový priestor, internet, neologizmy, spracovanie prirodzeného jazyka

I. ÚVOD

Identifikácia nových slov, neologizmov, je dôležitým procesom pre sledovanie zmien jazyka, ktoré umožňuje aktualizáciu slovníkov či vylepšovanie a zabezpečenie aktuálnosti nástrojov využívajúcich spracovanie prirodzeného jazyka. Internetový priestor prináša do obehu množstvo nových textov, pre ktoré existuje potenciál výskytu nových slov. Zároveň umožňuje rýchle šírenie týchto textov naprieč celým svetom. Vzhľadom na ich veľké množstvo a rýchlosť šírenia nie je efektívne v textoch manuálne vyhľadávať nové slová. Preto techniky ich čiastočne automatizovanej identifikácie sú čím ďalej, tým viac rozšírené.

Cieľom našej práce bolo v dátach zozbieraných zo slovenského internetového doménového priestoru identifikovať potenciálne neologizmy pomocou metód čiastočne automatizovanej identifikácie neologizmov.

II. EXISTUJÚCE METÓDY IDENTIFIKÁCIE NEOLOGIZMOV

A. Manuálna analýza

Čítanie textov a súbežné vyhľadávanie nových slov uskutočňované človekom je najstarším spôsobom identifikácie neologizmov. Vzhľadom na náklady na čas i ľudské zdroje je tento spôsob nepostačujúci množstvu dát, ktoré sú v súčasnosti dostupné. Zároveň je manuálna analýza vysoko subjektívna.

B. Čiastočne automatizovaná identifikácia neologizmov

Automatizáciou krokov identifikácie neologizmov je možné proces výrazne urýchliť. Keďže automatizované kroky sú uskutočňované výpočtovou technikou na základe vopred definovaných pravidiel, predchádza sa do určitej miery aj subjektívnosti pri výbere neologizmov.

Problémom automatizovaných metód identifikácie neologizmov je nejednoznačnosť definície a vlastností neologizmov. Medzi slovami identifikovanými ako neologizmy sa teda s vysokou pravdepodobnosťou vyskytujú aj reťazce znakov falošne označené za neologizmy. Výstupy automatizovaných metód preto možno považovať za tzv. kandidáty na neologizmy, pričom tie si v poslednom kroku vyžadujú aj manuálnu analýzu. Takýto prístup sa vo všeobecnosti označuje za čiastočne automatizovanú, resp. semi-automatizovanú identifikáciu neologizmov.

Existujú viaceré metódy čiastočne automatizovanej identifikácie neologizmov opierajúce sa o rôzne známe definície a charakteristiky týchto slov. Ich použitie je možné samostatne, no

častokrát sú aplikované v kombinácii.

1) *Predspracovanie dát*

V rámci predspracovania dát je možné vykonať odstránenie reťazcov, ktoré nemajú potenciál byť neologizmami, napríklad vlastných mien, preklepov či nezmyselných reťazcov.

2) *Porovnanie reťazcov s vylučovacím zoznamom*

Najrozšírenejšia forma semi-automatizovanej identifikácie neologizmov je založená na vylúčení: neologizmom, resp. kandidátom na neologizmus, je označené každé slovo, ktoré sa nenachádza vo vopred určenom referenčnom zozname slov.

Ako referenčný zoznam slov, ktorého výber je podľa Kerremans a Prokić [1] kritický, je možné použiť slovníkové alebo korpusové dáta, prípadne ich kombináciu, ktorú Janssen [2] nazval „rozšírené lexikografické diachronické kritérium“.

Porovnanie reťazcov je možné rozšíriť o výpočet Levenshteinovej vzdialenosti. Levenshteinova vzdialenosť [3] je najmenší počet operácií vloženia, zmazania a výmeny znakov potrebných pre transformáciu jedného reťazca na druhý. Môže byť použitá ako parameter pri porovnaní reťazcov, ktorý stanovuje podobnosť kandidátov so slovami vylučovacieho zoznamu.

3) *Analýza frekvencií výskytu*

Analýza frekvencií výskytu skúma počet výskytov jednotlivých slov, môže však analyzovať i geografické či časové trendy ich používania. Ako poznamenali Freixa a Torner [4] a Metcalf [5], nie každé nové slovo je možné považovať za neologizmus, pokiaľ jeho rozšírenie je obmedzené na istú geografickú oblasť či úzku skupinu ľudí. Príliš nízke počty výskytov teda znamenajú vysokú pravdepodobnosť, že slovo nie je neologizmom a v mnohých prípadoch ani platným slovom.

4) *Strojové učenie*

K aplikácii strojového učenia na identifikáciu neologizmov boli predstavené viaceré prístupy. Zatiaľ čo Stenetorp [6] a tím Chiru a Rebedea [7] strojové učenie použili ako nosnú metódu svojich identifikátorov, projekt Logoscope [8] ho využíva ako doplnok vylučovacej metódy, resp. porovnania reťazcov.

Stenetorp [6], rovnako ako tím zostavujúci Logoscope [8], pracoval so zoznamom slov a ako tréningové dáta použil vopred definovaný zoznam neologizmov. Chiru a Rebedea [7] klasifikátor trénovali na evolučných krivkách vybraných slov.

Veľkým problémom aplikácie strojového učenia je však potreba dostatočne veľkého vopred získaného zoznamu neologizmov pre potreby tréningovania modelu.

5) *Metalingvistické značky*

Hľadanie metalingvistických značiek alebo tzv. „diskriminantov“, ktoré môžu signalizovať výskyt neologizmov, v textoch, je technika, ktorú použil Paryzek [9]. Je založená na koncepte, ktorý formuloval Chlebda [10], a to, že frazémy¹ sa vyskytujú v úvodzovkách alebo za istými frázami. Tieto diskriminanty možno rozdeliť do dvoch kategórií:

- lexikálne - frázy, o ktorých predpokladáme, že predchádzajú neologizmom,
- interpunkčné - úvodzovky, apostrofy a iné interpunkčné znaky, v ktorých sa môžu nachádzať neologizmy.

III. VSTUPNÉ DÁTA

Zdrojom vstupných dát pre identifikáciu neologizmov bol v rámci tejto práce slovenský internetový doménový priestor. V praxi to znamená obsah všetkých webových stránok registrovaných pod doménou .sk, ktoré boli dostupné v čase písania práce, resp. v čase zberu dát.

A. *Stiahnutie vstupných dát*

Aby bola zabezpečená rôznorodosť a dostatočné pokrytie doménového priestoru pri zbere dát, ako východisková databáza domén bol použitý ich zoznam z webovej stránky SK-NIC² aktuálny k dátumu 4.10.2022. Obsahuje denne generovaný zoznam všetkých registrovaných slovenských domén.

Obsahy webových stránok boli stiahnuté Bash skriptom využívajúcim príkaz *wget*, pričom boli uložené ako textové súbory. Zároveň bol vypracovaný protokol obsahujúci návratové hodnoty tohto príkazu, aby bolo spätne možno analyzovať dostupnosť, resp. nedostupnosť webových stránok a jej dôvody.

B. *Príprava východiskového zoznamu slov*

Žiadaným výstupom spracovania vstupných dát bol zoznam slov, v ktorom sa každé slovo vyskytuje práve raz, spolu s počtom výskytov daného slova. Stiahnuté obsahy webových stránok boli spracované do tejto podoby nasledujúcimi krokmi:

¹ „Frazéma je každý jazykový znak, bez ohľadu na jeho sémantický status a formálnu štruktúru, ktorý predstavuje názov pre obsahový potenciál (termín), ktorý hovoriaci vyslovuje ako jeho relatívne konštantný symbol.“ (Chlebda 1991: 27)

² <https://sk-nic.sk/>

- odstránenie HTML tag-ov a separácia slov,
- konverzia na malé písmená,
- odstránenie duplicity a spočítanie počtu výskytov.

IV. IDENTIFIKÁCIA NEOLOGIZMOV

Kroky identifikácie neologizmov boli realizované skriptami a programami písanými v skriptovacom jazyku Bash a v programovacom jazyku C++.

A. *Predspracovanie dát*

V rámci predspracovania dát boli vykonané dve operácie:

- odstránenie slov obsahujúcich symboly, ktoré nie sú súčasťou slovenskej abecedy alebo nepatrili do užšieho výberu povolených symbolov,
- odstránenie slov, ktoré možno považovať za preklepy alebo neplatné slová, teda slová obsahujúce ten istý znak vyskytujúci sa tri- a viackrát bezprostredne za sebou.

B. *Porovnanie reťazcov*

Pre porovnanie reťazcov boli využité referenčné dáta zo štyroch zdrojov:

- morfológická databáza SNK v aktuálnej verzii 2015-02-05,
- korpusové dáta vo verzii prim-6.1 z roku 2013,
- dáta z webových stránok z roku 2014 zozbierané v rámci záverečnej práce [11],
- databáza štandardizovaného geografického názvoslovia.

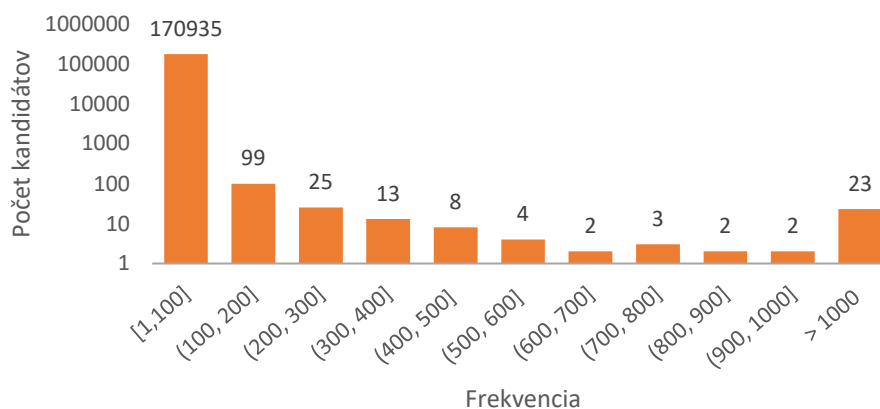
Dáta z každého zdroja boli upravené do podoby zoznamu slov, pričom získané zoznamy boli neskôr zlúčené do jedného vylučovacieho zoznamu.

Porovnanie reťazcov bolo realizované v dvoch fázach. Najprv bol zoznam vstupných slov porovnaný samostatne s každým referenčným zoznamom slov. Tento krok bol použitý najmä pre posúdenie relevancie jednotlivých referenčných zoznamov. Následne boli vstupné slová porovnané so zlúčeným vylučovacím zoznamom.

C. *Analýza frekvencií výskytu*

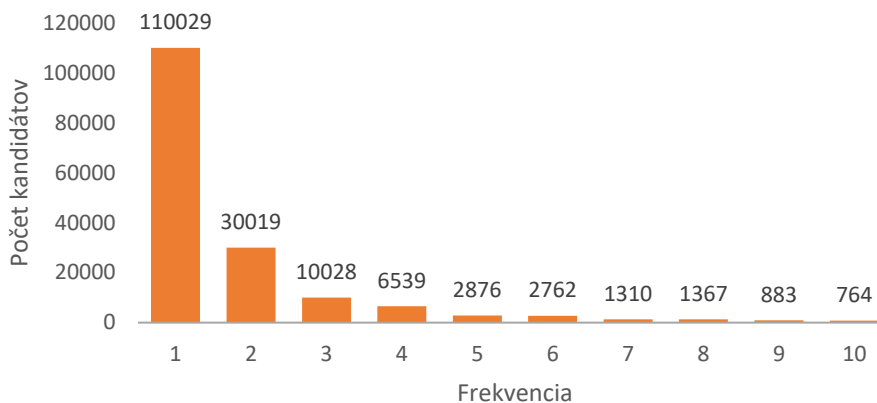
Každému slovu bola pri prvotných úpravách vyzbieraných dát vypočítaná tzv. frekvencia, resp. počet výskytov. Cieľom frekvenčnej analýzy bolo na základe posúdenia relevancie slov disponujúcich určitou frekvenciou nájsť tzv. hranicu predstavujúcu najnižší počet výskytov, ktorý ešte znamenal, že kandidát mal potenciál byť neologizmom. Pri hľadaní hranice bol posudzovaný nielen potenciál slova byť neologizmom, ale i platným slovom ako takým. Za hranicu bolo na záver stanovené číslo 4, ktoré bolo prijateľným kompromisom medzi rizikami prítomnosti príliš veľkého počtu falošných kandidátov a stratou skutočných neologizmov. Následne boli slová s frekvenciou nižšou ako 4 odstránené zo zoznamu kandidátov.

Štatistické rozdelenie frekvencií je zobrazené v histograme na Obr. 1. Histogram bol pre prehľadnejšie zobrazenie veľmi širokého rozsahu hodnôt počtu kandidátov budovaný v logaritmickej stupnici na zvislej osi využitím logaritmu so základom 10. Ako je možné vidieť, počty výskytov značnej väčšiny slov sa pohybovali v nízkych hodnotách.



Obr. 1 Štatistické rozdelenie frekvencie výskytov

Histogram na Obr. 2 bližšie zobrazuje rozdelenie frekvencií v rozsahu 1 až 10, pričom je zjavné, že viac ako polovica celkového počtu slov (171 116) mala zaznamenaný len 1 výskyt.



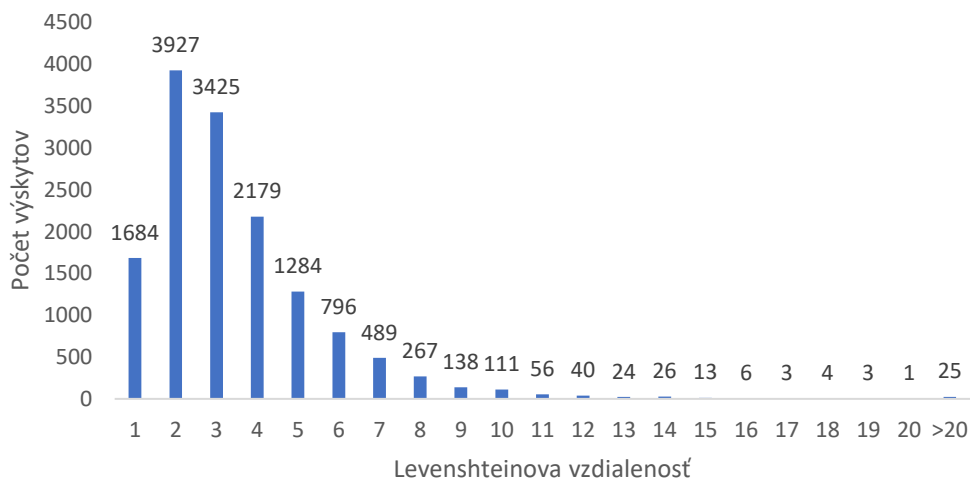
Obr. 2 Štatistické rozdelenie frekvencie výskytov pre frekvencie z intervalu 1 až 10

D. Aplikácia Levenshteinovej vzdialenosti

Výpočet Levenshteinovej vzdialenosti bol uskutočnený na kandidátoch vybraných v rámci frekvenčnej analýzy. Levenshteinova vzdialenosť bola pre tieto slová počítaná vo vzťahu k slovám z morfolologickej databázy, keďže mala primárne slúžiť na odhalenie nesprávne napísaných už existujúcich slov. Každému slovu zo zoznamu kandidátov bola nájdená najnižšia hodnota Levenshteinovej vzdialenosti, ktorú dosiahlo pri porovnávaní so slovami morfolologickej databázy. V našom prípade bola najnižšia možná hodnota vzdialenosti rovná 1, keďže reťazce identické so slovami databázy boli odstránené v kroku porovnania reťazcov, a teda hodnota 0 už nebola dosiahnuteľná.

Opísaný postup bol vykonávaný programom napísaným v jazyku C++, ktorý pre výpočet Levenshteinovej vzdialenosti využíva Wagner-Fischer algoritmus [12], ktorý počíta túto hodnotu pomocou tabuľky dynamického programovania.

Hodnoty získané výpočtom Levenshteinovej vzdialenosti sa nachádzali v rozpätí od 1 do 108, pričom hodnoty vyššie ako 15 sa vyskytovali pri veľmi malom počte slov. Štatistické rozdelenie vypočítaných hodnôt je zobrazené v grafe na Obr. 3.



Obr. 3 Štatistické rozdelenie hodnôt Levenshteinovej vzdialenosti

E. Identifikácia neologizmov na základe metalingvistických značiek

Identifikácia na základe prítomnosti metalingvistických značiek, alebo aj diskriminantov, bola použitá samostatne, nakoľko už v kroku úpravy vstupných dát bolo nutné vykonať zmeny. Zo súborov obsahujúcich stiahnuté webové stránky boli extrahované len tie slová, ktoré sa nachádzali medzi dvoma typmi interpunkčných diskriminantov:

- slovenské úvodzovky,
- anglické úvodzovky.

Úvodzovky boli následne odstránené a získané slová boli podrobené postupu spracovania analogickému k spracovaniu vstupných dát:

- konverzia na malé písmená,
- odstránenie duplikátov a výpočet frekvencií,
- predspracovanie dát.

Po spracovaní boli zo slov vybrané kandidáty na neologizmy technikou porovnania reťazcov s použitím zostaveného vylučovacieho zoznamu.

F. Porovnanie vstupných dát v rozpätí štyroch mesiacov

Experimentálnym krokom bolo opätovné stiahnutie vstupných dát z internetových domén a ich porovnanie s pôvodnými vstupnými dátami, ktoré v tomto prípade figurovali ako referenčné dáta.

Súbor domén bol druhýkrát stiahnutý k dátumu 25.2.2023, teda približne v rozpätí 4 mesiacov od stiahnutia pôvodných vstupných dát. Dáta boli podrobené rovnakému postupu ako pôvodné vstupné dáta:

- stiahnutie,
- odstránenie HTML tag-ov a extrahovanie slov,
- konverzia na malé písmená a odstránenie duplikátov,
- predspracovanie dát.

Následne boli slová porovnané s pôvodným zoznamom slov z októbra aj s vylučovacím zoznamom. Februárový zoznam kandidátov bol nakoniec vybraný použitím frekvenčnej analýzy s hranicou rovnakou ako pri októbrových dátach, aby bola zachovaná konzistencia postupu.

G. Manuálna analýza

Po vykonaní všetkých automatizovaných krokov identifikácie neologizmov bolo nutné zhodnotiť získané výsledky a z kandidátov vybrať slová s potenciálom byť neologizmami ich manuálnou analýzou.

Okrem samotného výberu potenciálnych neologizmov bol kladený dôraz na identifikáciu rôznych atribútov, ako ich pôvodu, spôsobu vzniku či slovného druhu a určenie ich základného tvaru. Dôležitá bola aj analýza týchto atribútov pri falošných kandidátoch, ktorá môže v budúcnosti vylepšiť a spresniť automatizovanú identifikáciu neologizmov. Falošné kandidáty je pritom možné priamo zaradiť do vylučovacieho zoznamu, zároveň ich spoločné charakteristiky môžu slúžiť pre definovanie pravidiel predspracovania dát.

Manuálna analýza bola vykonaná autorkou práce. Pre ďalšie spracovanie a využitie budú identifikované neologizmy posúdené lingvistami.

V. VYHODNOTENIE VÝSLEDKOV

Zoznam domén získaný zo stránky SK-NIC v čase stiahnutia obsahoval 446 402 záznamov, pričom počet úspešne stiahnutých domén bol 361 284. Z nich bolo extrahovaných 30 314 570 slov, z ktorých po odstránení duplikátov vznikol východiskový zoznam 2 071 519 slov.

Na tieto slová boli postupne aplikované kroky predspracovania dát, porovnania reťazcov s vylučovacím zoznamom a analýzy frekvencií výskytu, po ktorých bolo ponechaných 14 501 kandidátov. Podrobné číselné vyhodnotenie zobrazuje Tabuľka 1. Levenshteinova vzdialenosť bola pre riziko vyradenia skutočných neologizmov využitá len ako indikátor pravdepodobnosti slova byť falošným kandidátom pri manuálnej analýze.

Tabuľka 1

Číselné vyhodnotenie predspracovania dát, porovnania reťazcov a frekvenčnej analýzy uskutočnených na vstupných dátach

	Pôvodný počet slov	Počet odstránených slov	Počet kandidátov
Predspracovanie dát	2 071 519	1 452 059	619 460
Porovnanie reťazcov	619 460	448 344	171 116
Frekvenčná analýza	171 116	156 615	14 501

Výber slov nachádzajúcich sa vnútri metalingvistických značiek vyústil do nájdenia 152 092 slov, z toho:

- 1 850 sa nachádzalo v slovenských úvodzovkách,
- 150 242 sa nachádzalo v anglických úvodzovkách.

Odstránenie duplikátov, predspracovanie dát a porovnanie reťazcov s vylučovacím zoznamom viedlo k výberu 1 030 kandidátov na neologizmy. Presnejšie údaje zobrazuje Tabuľka 2.

Tabuľka 2

Číselné vyhodnotenie identifikácie neologizmov na základe metalingvistických značiek

	Pôvodný počet slov	Počet odstránených slov	Počet kandidátov
Slovenské úvodzovky	1 850	1 819	31
Anglické úvodzovky	150 242	149 243	999

Vstupné dáta opätovne stiahnuté po štyroch mesiacoch obsahovali dokopy 2 839 185 slov, ktorých počet sa po aplikovaní krokov predspracovania dát zredukoval na 962 760 slov. Z nich

441 060 sa nachádzalo aj vo februárových dátach. Získané nové slová boli následne porovnané s vylučovacím zoznamom slov, čo viedlo k získaniu 284 590 kandidátov. Po aplikovaní frekvenčnej hranice 4 sa počet kandidátov zredukoval na 8 815, ktoré boli neskôr analyzované manuálne. Podrobné údaje zobrazuje Tabuľka 3. Je zjavné, že obsah webových stránok sa dynamicky a často mení a je nutné neologizmy identifikovať pravidelne.

Tabuľka 3
Číselné vyhodnotenie identifikácie neologizmov na základe porovnania vstupných dát v rozpätí štyroch mesiacov

	Pôvodný počet slov	Počet odstránených slov	Počet kandidátov
Stiahnutie a úprava dát	-	-	2 839 185
Predspracovanie dát	2 839 185	1 876 395	962 790
Porovnanie reťazcov	962 790	678 200	284 590
Frekvenčná analýza	284 590	275 775	8 815

Manuálna analýza bolo vykonaná nezávisle na kandidátoch získaných zo samostatne stojacich slov z októbra 2022 i februára 2023 a slov nachádzajúcich sa v anglických a slovenských úvodzovkách. Počty extrahovaných neologizmov zobrazuje Tabuľka 4. Po odstránení duplikátov, ktorých počet bol 4, bol počet nájdených neologizmov 766. Nakoľko však tento zoznam obsahoval aj viacero tvarov toho istého slova, boli všetky neologizmy konvertované do ich základných tvarov. Výsledný počet identifikovaných neologizmov bol **705**.

Tabuľka 4
Číselné vyhodnotenie manuálnej analýzy kandidátov

	Počet kandidátov	Počet falošných kandidátov	Počet neologizmov
Slová – október	14 501	13 931	570
Slová – február	8 815	8 654	161
Slovenské úvodzovky	999	972	27
Anglické úvodzovky	31	19	12

Množina falošných kandidátov pritom obsahovala cudzie slová (*cyberawareness*), vlastné mená a názvy (*GooglePay*), zret'azené viacslovné pomenovania (*ajjachranimplanetu*), slangové výrazy (*knihožrúť*), iniciálové skratky (*GDPR*), slová súvisiace s tvorbou webov (*allowcookies*) a preklepy (*kombinvoané*).

Identifikované neologizmy boli utvorené odvodzovaním (*ebrána*), skladaním (*cykloreģión*), preberaním z cudzích jazykov (*cloudový*) či tvorením nových slovných druhov (*naskrutkovanie*). Pochádzali z oblastí internetového predaja (*balíkobox*), kryptomien (*kryptoplatforma*), vedy a medicíny (*pocovidový*), sociálnych sietí a internetu (*influencer*), technológií (*ransomvér*) i športu (*workoutový*).

VI. ZÁVER

Tento príspevok bol zameraný na identifikáciu neologizmov v slovenskom internetovom doménovom priestore. Na dáta stiahnuté z dostupných webových stránok registrovaných pod slovenskou doménou bolo aplikované predspracovanie dát a následne viaceré metódy čiastočne automatizovanej identifikácie neologizmov, a to porovnanie reťazcov s vylučovacím zoznamom, frekvenčná analýza, aplikácia Levenshteinovej vzdialenosti a identifikácia neologizmov na základe metalingvistických značiek. Manuálna analýza výsledkov potvrdila výskyt neologizmov v získaných zoznamoch slov, zároveň dokázala i prítomnosť falošných kandidátov. Na záver autori vyčlenili 705 potenciálnych neologizmov. Hlavnou motiváciou budúceho vývoja je zlepšenie presnosti, teda zredukovanie počtu falošných kandidátov.

POĎAKOVANIE

Táto práca bola podporená Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-18-0046.

ZDROJE

- [1] KERREMANS, Daphné; PROKIĆ, Jelena. Mining the web for new words: Semi-automatic neologism identification with the NeoCrawler. *Anglia*. 2018, roč. 136, č. 2, s. 239–268.
- [2] JANSSEN, Maarten. Orthographic Neologisms. Selection Criteria and Semi-Automatic Detection. 2011.
- [3] LEVENSHEIN, Vladimir I et al. Binary codes capable of correcting deletions, insertions, and reversals. In: *Soviet physics doklady*. 1966, zv. 10, s. 707–710. Č. 8.
- [4] FREIXA, Judit; TORNER, Sergi. Beyond frequency: On the dictionarization of new words in Spanish. *Dictionaries: Journal of the Dictionary Society of North America*. 2020, roč. 41, č. 1, s. 131–153.
- [5] METCALF, Allan A. Predicting new words: The secrets of their success. Houghton Mifflin Harcourt, 2004.

- [6] STENETORP, Pontus. Automated extraction of swedish neologisms using a temporally annotated corpus. Citeseer, 2010.
- [7] CHIRU, Costin-Gabriel; REBEDEA, Traian Eugen. Archaisms and neologisms identification in texts. In: 2014 RoEduNet Conference 13th Edition: Networking in Education and Research Joint Event RENAM 8th Conference. 2014, s. 1–6.
- [8] FALK, Ingrid; BERNHARD, Delphine; GÉRARD, Christophe. The Logoscope: A semi-automatic tool for detecting and documenting French new words from the linguistic project to the web interface. 2018. Diz. pr. Université Strasbourg.
- [9] PARYZEK, Piotr. Comparison of selected methods for the retrieval of neologisms. *Investigationes linguisticae*. 2008, roč. 16, s. 163–181.
- [10] CHLEBDA, Wojciech. *Elementy frazematyki: wprowadzenie do frazeologii nadawcy*. Oficyna Wydawn. Leksem, 2003.
- [11] BAČO, Martin. *Statické a dynamické charakteristiky slovenského webu*. Technická univerzita v Košiciach, 2015.
- [12] WAGNER, Robert A; FISCHER, Michael J. The string-to-string correction problem. *Journal of the ACM (JACM)*. 1974, roč. 21, č. 1, s. 168–173.

Meranie aspektov použiteľnosti systémov pre podporu rozhodovania lekárov

¹Oliver LOHAJ, ²Ján PARALIČ, ³Ján JEVČÁK

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

¹oliver.lohaj@tuke.sk, ²jan.paralic@tuke.sk, ³jan.jevcak@student.tuke.sk

Abstrakt — Systémy na podporu klinického rozhodovania (CDSS) sa stali dôležitým nástrojom v zdravotníctve na zlepšenie starostlivosti o pacientov v rôznych fázach tohto procesu. Na zabezpečenie účinnosti, efektívnosti a spokojnosti s týmito systémami je však nevyhnutné zohľadniť viaceré aspekty ich použiteľnosti. Táto práca predstavuje referenčný model návrhu CDSS, ktorý podporuje použiteľnosť, založený na troch kľúčových zložkách: (1) model zjednotenej teórie akceptácie a používania technológií (UTAUT), (2) referenčný model pre CDSS a (3) model rozhodovania (RPD) na simuláciu rozhodovania používateľov. Aspekty použiteľnosti sa merali pomocou štandardných dotazníkov ASQ [12] a PSSUQ [13] a na zber kvalitatívnych údajov sa použila metóda hlasného premýšľania (Thinking aloud method). Výsledky boli porovnané aj s predchádzajúcimi štúdiami. Tento článok poskytuje poznatky pre navrhovanie CDSS s dôrazom na použiteľnosť, čím sa zvyšuje účinnosť, efektívnosť a spokojnosť zdravotníckych pracovníkov a pacientov s týmito systémami.

Kľúčové slová — Použiteľnosť, Aspekty použiteľnosti, Systém na podporu rozhodovania, Referenčný model

I. ÚVOD

Vo výskume sú často študované rôzne aspekty systémov pre podporu rozhodovania, ktoré sú úspešne nasadzované v rôznych aplikačných oblastiach. Pre ich plné využitie je však potrebné zohľadniť viaceré aspekty ovplyvňujúce ich použiteľnosť. Tento článok popisuje koncepčný výsledok výskumu týkajúceho sa merania aspektov použiteľnosti systémov pre podporu rozhodovania lekárov. Pri rozhodovaní lekárov môžu byť nápomocné systémy na podporu rozhodovania, ktoré zvyšujú bezpečnosť pacientov a tým pádom aj ich spokojnosť. V prieskume šiestich komunitných nemocníc v oblasti mesta Boston sa podľa [1] dalo 9 z 10 nežiaducich účinkov súvisiacich s liekmi predísť a všetkým týmto nežiaducim účinkom sa dalo predísť pomocou systému CDSS, ktorý by vedel upozorňovať lekárov na konflikty medzi liekmi a pacientom, či nesprávne dávkovanie.

Okrem predchádzania nesprávnych medikácií je častým cieľom nasadenie CDSS optimalizácia nákladov. Napríklad štúdia [2] hovorí o znížení finančných nákladov o 11% pri hospitalizáciách a 27% na oddelení urgentného príjmu. CDSS nielen napomáhajú pri rozhodovaní alebo zvyšovaní bezpečnosti a efektivity ale znižujú aj náklady.

Pre správne fungovanie CDSS je veľmi dôležité maximalizovať jeho použiteľnosť. Použiteľnosť je termín používaný v oblasti návrhu a hodnotenia interaktívnych systémov, ktorý sa zaoberá schopnosťou používateľa úspešne a efektívne používať daný systém s minimálnym úsilím [14]. Norma ISO 9241-11:2018 definuje nasledujúce hlavné aspekty použiteľnosti [15]:

- Účinnosť: Miera, do akej používateľa dokážu dosiahnuť svoje ciele s príslušným systémom presne a úplne, pričom využívajú správne množstvo zdrojov a bez straty času.
- Efektívnosť: Miera, do akej používateľa dosahujú požadované výsledky s daným systémom pri použití určeného množstva zdrojov
- Spokojnosť: Miera, do akej sú používatelia spokojní s používaním systému.

Ako vhodne testovať a merať rôzne aspekty použiteľnosti systémov na podporu rozhodovania vo všeobecnosti a špecificky CDSS, nie je zďaleka uspokojivo preskúmané. Na podporu testovania použiteľnosti je niekoľko dôkazov v podobe štúdií. Napríklad [3] a [4] hovoria, že pri meraniach rôznymi metódami je aj metódou s najnižšou efektívnosťou vysoká pravdepodobnosť odhalenia chýb. Ďalej tieto štúdie poukázali na schopnosť jednotlivé metódy kombinovať, čím sa zvýši efektívnosť. Dôležitým doplnkom je aj fakt, že takéto formy testovania použiteľnosti fungujú

aj keď nie sú vykonávané expertom v danej tematike. Výsledkom práce Carolla [5] je zistenie rozdielov medzi očakávaniami v použiteľnosti CDSS medzi lekármi a sestričkami. Vďaka spätnej väzbe, ktorá ponúkala pohľady dvoch rôznych strán, bolo možné urobiť lepší a použiteľnejší systém.

V našom výskume sme sa zamerali na aspekty, ktoré sú v dnešnej dobe dôležité, ako napríklad zabezpečiť účinnosť, efektívnosť a spokojnosť so systémami pre podporu rozhodovania (CDSS). Pri návrhu týchto systémov je však nevyhnutné zohľadniť aj ďalšie aspekty ich použiteľnosti. Referenčný model návrhu CDSS prináša rámcový prístup na zlepšenie použiteľnosti týchto systémov. Referenčný model zdôrazňuje hlavné aspekty použiteľnosti, ako je účinnosť (dosahovanie požadovaných výsledkov s minimálnym úsilím), efektívnosť (dosahovanie výsledkov s dostupnými zdrojmi) a spokojnosť používateľov. Okrem toho sa zohľadňujú aj ďalšie dôležité faktory, ako je skúsenosť s daným systémom, či dobrovoľnosť používania a pokrytie používateľských potrieb v rôznych situáciách. Táto práca o referenčnom modeli návrhu CDSS prispieva k lepšiemu porozumeniu a zabezpečeniu použiteľnosti systémov pre podporu rozhodovania, čo môže viesť k ich vylepšeniu a zvýšeniu spokojnosti používateľov.

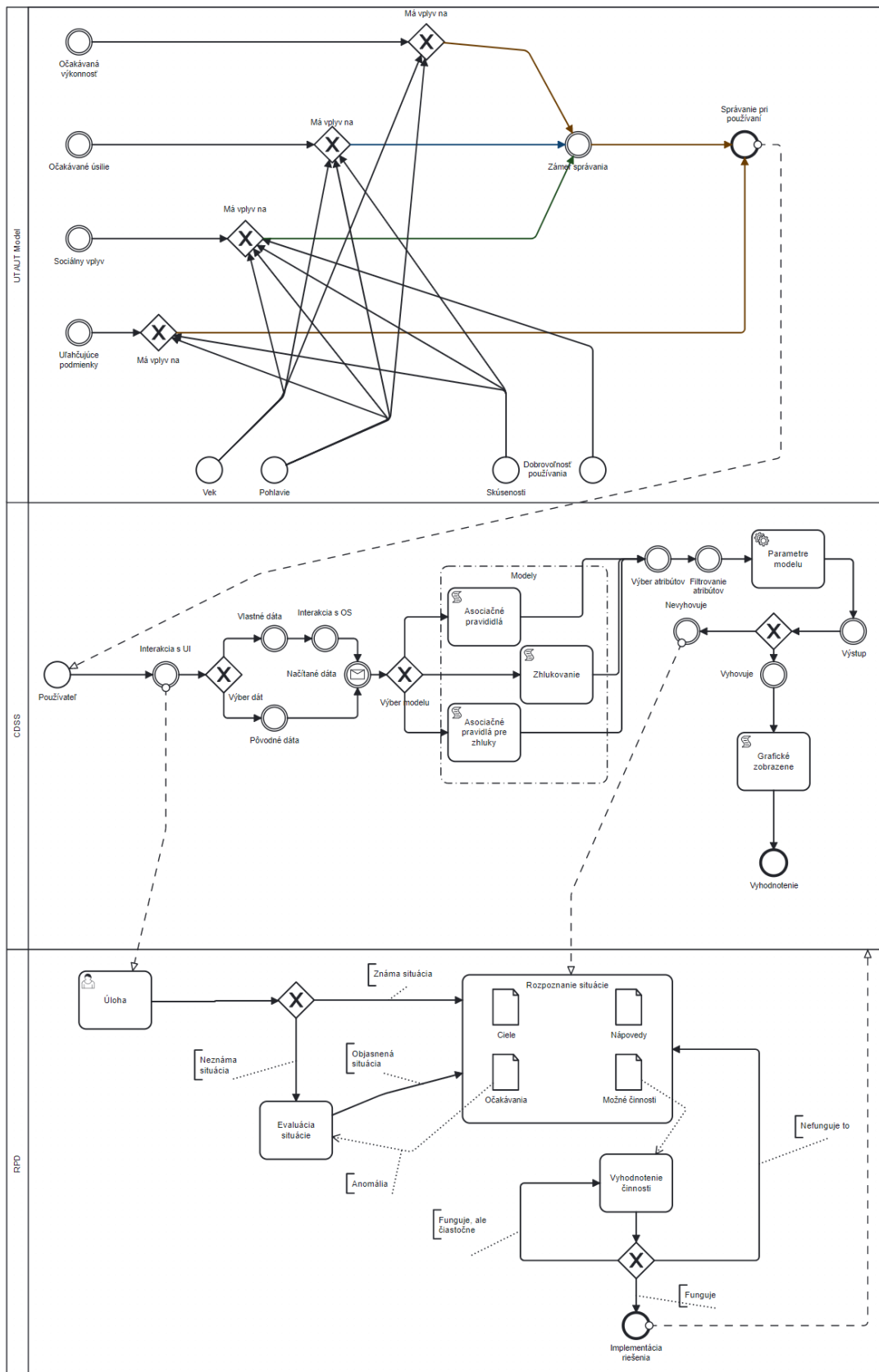
V kapitole II. je predstavený a vysvetlený navrhnutý referenčný model. Kapitola III. približuje metodiku experimentálnej časti výskumu zameranú na testovanie konkrétneho CDSS [16] na vybranej skupine používateľov s využitím štandardizovaných dotazníkov doplnených o metódu hlasného premýšľania [17], ako aj výsledky experimentu. Záverečná kapitola IV. sumarizuje najdôležitejšie závery prezentovaného výskumu.

II. NÁVRH RIEŠENIA

Problematika použiteľnosti CDSS nie je dostatočne preskúmaná [19]. Chýba koncepčné prepojenie kľúčových aspektov použiteľnosti špecificky dôležitých pre informačné systémy typu CDSS. Z toho dôvodu sme navrhli referenčný model (pozri Obr. 1), ktorý pomenúva jednotlivé komponenty a súčasti daného systému, zohľadňuje kľúčové aspekty použiteľnosti a rozhodovacieho procesu. Následne v experimentálnej časti potom ukážeme spôsob využitia tohto modelu pre zostavenie používateľskej štúdie na otestovanie konkrétneho CDSS s cieľom identifikovať najslabšie články jeho použiteľnosti.

Navrhnutý model môže napomôcť programátorom, ktorí sa s daným typom systému ešte nestretli, k lepšiemu porozumeniu. Keďže aj použiteľnosť vieme definovať viacerými spôsobmi, navrhnutý referenčný model sa opiera o viacero pohľadov na vec. Tento postup je založený na práci S. Kaparthy [6], ktorí pracovali s cloudovým CDSS. Naše tri pohľady pozostávajú z :

- **Použiteľnosť** - Veľký vplyv na použiteľnosť má aj samotný používateľ a jeho vlastnosti v podobe veku, skúseností, vnútornej a vonkajšej motivácie. Práve týmto faktorom sa venuje rámec UTAUT (Unified Theory of Aceptance and Use of Technology). Tento model je schopný napomôcť pri podchytení prvkov, ktoré by mohli softvéru brániť v jeho schopnosti nadobúdať používateľov a následne taktiež sťažovať tvorcom zdokonaľovanie systému [7].
- **Tvorba rozhodnutí** - Aby sme porozumeli ako sa jednotlivé rozhodnutia tvoria, vznikol model s názvom RPD. Ten popisuje, ako sa venujú tvorbe rozhodnutí experti v danom obore, pretože tí neanalyzujú viacero možností ale vytvárajú postup po jednom, pretože dokážu identifikovať dobrú možnosť ako prvú, ktorú zvažujú. Experti si vyberajú alternatívu, ktorá je „dostatočne dobrá“, ale nie nevyhnutne najlepšia možná, čo znamená že sa sústredia na efektívnosť a na rýchlosť [8].
- **Systém** - Existuje referenčný model, ktorý vnáša trochu svetla do problematiky tvorby a fungovania klinických systémov na podporu rozhodovania. Zikos a DeLellis [9] vytvorili model, ktorý sa síce venuje použitiu daného CDSS systému v nemocničnom prostredí no základy sú rovnaké bez ohľadu na prostredie, v ktorom bude používaný.



Obr.1 Navrhnutý referenčný model pre testovanie použiteľnosti CDSS

III. METODIKA A VYHODNOTENIE EXPERIMENTU

Experimentálna časť nášho výskumu bola realizovaná formou používateľskej štúdie. Tá pozostávala z testovania systému, zberu spätnej väzby troma navzájom sa dopĺňujúcimi metódami: dvoma typmi dotazníkov, jeden po každej úlohe a jeden až po skončení všetkých úloh, ktoré nám umožnili kvantifikovať niektoré aspekty použiteľnosti. Použili sme aj metódu Thinking aloud na zachytenie kvalitatívnej spätnej väzby. Po vyhodnotení získanej spätnej väzby

a na základe aspektov použiteľnosti boli následne navrhnuté spôsoby, ako vylepšiť CDSS a zlepšiť tak jeho použiteľnosť.

Celkovo sme vybrali 5 testerov, keďže na základe zdroja [18] sme zistili, že testovanie s piatimi používateľmi predstavuje rovnováhu medzi potrebou dostatočných údajov a praktickými obmedzeniami času a zdrojov. Výber respondentov bol vykonaný pomocou UTAUT modelu, kde každý z testerov mal pokrývať aspoň jeden z bodov modelu. Respondenti boli vybraní na základe ich profesie a teda testovacia vzorka obsahovala študentov medicíny, zdravotnú sestru, a softvérového developera.

Testovaným softvérom bola aplikácia SPPRK [16]. Je to aplikácia, ktorá vznikla v spolupráci s VÚSCH v Košiciach a je primárne určená pre kardiológov na zistenie rizika KV ochorenia. Predmetom testovania bola časť aplikácie venovaná asociačným pravidlám a časť aplikácie zameraná na zhlukovanie.

Pri testovaní boli vyhodnocované tri hlavné aspekty použiteľnosti, a to účinnosť, efektívnosť a spokojnosť. Prvou analyzovanou a vyhodnocovanou bola *efektívnosť*. Tú sme merali pomocou dotazníka PSSUQ [13] s využitím otázok: 2,3,5,7,8,9,11,12. Tento dotazník používa sedembodovú Likertovú škálu v rozsahu od „úplne nesúhlasím“ po „úplne súhlasím.“ Pomocou týchto otázok dostávame hodnotu 3,21. Táto hodnota nám hovorí o kladnom, pozitívnom výsledku. Taktiež je dôležité podotknúť jednotlivé chyby u jednotlivých testerov a ich celkový čas, ktorý bol v priemere 30 sekúnd na jednu chybu, čo bolo spôsobené vysvetľovaním, kde nastala chyba a ako sa vrátiť späť aby sa mohlo pokračovať v testovaní. Takýmto spôsobom vieme vypočítať o koľko sa efektívnosť jednotlivých používateľov zhoršila.

$$\text{Percentuálna strata} = 100 - \left[\frac{T - (n \times \bar{x})}{T} \right] \times 100\% \quad (1)$$

- T = celkový čas testovania
- n = počet chýb
- \bar{x} = Priemerná strata času na úlohu

Takýmto spôsobom sme vyčíslili nasledovné straty v percentách:

Respondent₁₋₅ : (2.2, 5.3, 4.3, 4.9, 6.5)

Ďalej bola analyzovaná a vyhodnocovaná *účinnosť*. Opisom tohto aspektu použiteľnosti je aby mal systém vlastnosti, ktoré svojim používateľom dovoľia splniť svoje ciele. V prípade testovania sú tieto ciele dané scenáre a v nich úlohy. Ak chceme vyjadriť účinnosť, je potom následne potrebné vedieť koľko úloh používateľa úspešne dokončili a následne aj naopak.

$$\text{Účinnosť} = \frac{\text{Počet úspešne splnených úloh}}{\text{Celkový počet úloh}} \times 100\% \quad (2)$$

Pomocou tohto vzorca vieme vyhodnotiť mieru úspešnosti jednotlivých respondentov v dokončení všetkých úloh. Ideálna hodnota je 100%, no tá je v praxi nemožná a hlavne ak systém testujú respondenti, ktorí s ním nemajú predchádzajúce skúsenosti. Pre porovnanie, štúdia od J. Saura [10] kde bolo vykonaných 115 testovaní s 3472 používateľmi poukazuje, že priemerná efektívnosť je na úrovni 78%. Z tohto vieme usúdiť, že nami testovaný systém na podporu rozhodovania u lekárov má porovnateľnú a priemernú efektívnosť, vid'. výsledky nižšie.

Respondent₁₋₅ – (92.30, 69.23, 76.92, 76.92, 61.53) - Priemer - 75,38%

Posledným aspektom je *spokojnosť*. Meranie spokojnosti spočívalo na vyhodnocovaní dotazníkov. Ako sme už spomínali v predchádzajúcej sekcii, oba dotazníky – ASQ a PSSUQ používajú sedembodovú Likertovú škálu, preto sme ich pre tento prípad použitia zlúčili v otázkach venujúcim sa spokojnosti a spriemerovali ich.

- ASQ - 3,06, PSSUQ - 3,38 - **Spolu - 3,3**

Výsledkom je hodnota 3,3 ktorá je znova ako zvyšok systému priemerná s jemným náznakom kladnej reakcie zo strany používateľov. Aby sa tento aspekt ako aj ostatné vylepšili, je potrebné rozobrať chyby, ktoré tieto výsledky ovplyvnili a vyriešiť ich alebo zamedziť ich dodatočnému opakovaniu. Pomocou dotazníka PSSUQ vieme aj zistiť s ktorými časťami systému neboli naši respondenti spokojní. Dotazník pozostáva z troch častí: *užitočnosť systému*, *kvalita informácií* a *kvalita používateľského prostredia*. Nami dosiahnuté výsledky sme porovnali s výsledkami publikovanými v práci [11] Saura a Lewisa, ktorí robili 21 testovaní s 210 respondentmi (vid'. Tab. 1).

Tab.1 Hodnotenie častí systému

	Sauro A Lewis	CDSS
SYSUSE (užitočnosť systému)	2,80	3,5
INFOQUAL (kvalita informácií)	3,02	3,17
INTERQUAL (kvalita používateľského prostredia)	2,49	3,55
Celkovo	2,82	3,39

Okrem dotazníkov sme využili na získanie cennej spätnej väzby aj metódu Thinking aloud, ktorá ukázala medzery nami testovaného systému najmä v kvalite používateľského prostredia. Pomocou metódy hlasného premýšľania bolo možné získať vecné rady ktoré vznikli prepisom verbálneho prejavu počas interakcie s aplikáciou a hovoria nám, ako vylepšiť systém a zvýšiť všetky aspekty použiteľnosti a najmä používateľského prostredia (zosumarizované podľa jednotlivých aspektov použiteľnosti, ako aj podľa jednotlivých respondentov ich možno vidieť v Tab. 2):

1. Zakomponovanie vysvetliviek jednotlivých atribútov v dátach do aplikácie pomocou metódy „hover over“
2. Jednoduchý animovaný tutoriál na úvod
3. Podpora slovenského jazyka.
4. Podrobné vysvetlenie v rozhraní ako fungujú LHS a RHS pri asociačných pravidlách.
5. Polia pre vybrané atribúty by mali byť štandardne prázdne.
6. Personalizácia grafických zobrazení.
7. Väčší počet chybových hlášok a upozornení, farebné odlíšenie.
8. Zdôrazniť ideálny postup pre danú metódu napr. šípkami alebo farebným označením.
9. Pri metóde zhlukovania, ak pole v ktorom sú vybrané atribúty bude mať počet prvkov menší ako 2 tak aplikácia zamedzí pokračovaniu a zrealizuje reštart.
10. Zvýrazniť správy o potvrdení voľby.
11. Chýbajúce normy hodnôt daných parametrov napr. chorý a zdravý
12. Numerické značenie postupu pri metódach nie je dostatočne výrazné
13. Chýba mód pre farbosleposť.
14. Na úvod používateľské prostredie pôsobí ako veľká spleť informácií.
15. Ak by mal byť systém používaný pre prax potrebuje byť jednoduchší.

Rozdielom pri analýze týchto odporúčaní je, že sa jedná o kvalitatívne dáta a nie kvantitatívne. Týmto spôsobom vieme povedať, že výstupom sú priamo odporúčania na spôsob riešenia vzniknutých problémov, ktoré by mali byť schopné maximálne pozdvihnúť kvalitu nami používaných aspektov použiteľnosti. Ak napríklad zhrnieme 7. a 8. bod, tak používateľom chýbalo farebné odlíšenie, vďaka ktorému by vedeli odlíšiť závažnosť zobrazovaných informácií. Pri 9. bode neexistoval spôsob ktorým by sa používatelia mohli dozvedieť, že nemožnosť pridávať atribúty bola spôsobená porušením podmienky - v textovom poli musia byť aspoň 2 atribúty. Samotný protokol alebo metodika Thinking aloud pomohla pri používateľskej štúdií dostať z respondentov informácie, ktoré priamo napomáhajú pri zvyšovaní kvality systému a aj použiteľnosti.

Tab.2 Sumár spätnej väzby získanej metódou hlasného premýšľania (Thinking aloud)

	Tester 1	Tester 2	Tester 3	Tester 4	Tester 5	Efektívnosť	Efektivita	Spokojnosť
1.	•		•	•	•	•		
2.	•					•	•	
3.		•				•	•	•
4.		•				•		
5.		•	•	•	•	•		•
6.				•	•			•
7.		•	•	•	•	•	•	
8.		•	•	•	•	•		•
9.		•	•	•	•	•	•	•
10.			•	•		•		
11.			•	•	•	•		•
12.				•		•		•
13.				•		•	•	•
14.					•	•		•

15.					•		•	•
-----	--	--	--	--	---	--	---	---

IV. ZÁVER

Na základe výskumu a analýzy vykonanej v tomto článku možno konštatovať, že použiteľnosť je rozhodujúcim faktorom pri navrhovaní a vývoji systémov na podporu klinického rozhodovania (CDSS). Navrhovaný referenčný model, ktorý zahŕňa prvky z modelu UTAUT, modelu CDSS a modelu RPD môže slúžiť ako užitočný rámec na navrhovanie CDSS, ktoré sú pre zdravotníckych pracovníkov intuitívnejšie. Taktiež môže slúžiť ako ukazovateľ toho, ako testovať použiteľnosť a ako došlo k chybám. Okrem toho môže pomôcť ich vysvetliť a zároveň aj pochopiť. Výsledky testovania použiteľnosti, ktoré zahŕňalo použitie dotazníkov a metódy hlasného premýšľania (Thinking aloud) poskytli cenné poznatky o účinnosti, efektívnosti a spokojnosti s testovaným CDSS. Zo zistení vyplýva, že zdravotnícki pracovníci uprednostňujú CDSS, ktoré sa ľahko používajú, poskytujú presné a relevantné informácie a zlepšujú ich rozhodovacie schopnosti. Okrem toho porovnanie výsledkov s výsledkami predchádzajúcich štúdií poukazuje na dôležitosť zohľadnenia špecifického kontextu a požiadaviek používateľov pri navrhovaní CDSS. Táto práca prispieva k pokračujúcemu úsiliu o zlepšenie použiteľnosti CDSS a zdôrazňuje potrebu pokračovať vo výskume a vývoji v tejto oblasti s cieľom podporiť zdravotníckych pracovníkov pri poskytovaní kvalitnej starostlivosti.

POĎAKOVANIE

Táto práca vznikla vďaka podpore Vedeckej grantovej agentúry MŠVVaŠ SR a SAV projekt č. 1/0685/21 a Agentúry na podporu výskumu a vývoja na základe Zmluvy č. APVV-16-0213 a Zmluvy č. APVV-17-0550.

POUŽITÁ LITERATÚRA

- [1] Beeler, P. E., Bates, D. W., & Hug, B. L. (2014). Clinical decision support systems. In *Swiss Medical Weekly* (Vol. 144). EMH Swiss Medical Publishers Ltd. <https://doi.org/10.4414/smw.2014.14073>
- [2] Khan S, Maclean CD, Littenberg B. The effect of the Vermont Diabetes Information System on inpatient and emergency room use: results from a randomized trial. *Health Outcomes Res Med*. 2010
- [3] J. Nielsen, Usability engineering at a discount, in: *Proceedings of the 3rd International Conference on Human Computer Interaction*, 1989, pp. 2–8.
- [4] C.M. Karat, R. Campbell, T. Fiegel, Comparison of empirical testing and walkthrough methods in user interface evaluation, in: *Proceedings of CHI'92 Conference on Human Factors in Computing Systems*, 1992, pp. 397–404
- [5] Carroll, C., Marsden, P., Soden, P., Naylor, E., New, J., & Dornan, T. (2002). Involving users in the design and usability evaluation of a clinical decision support system. In *Computer Methods and Programs in Biomedicine* (Vol. 69). www.elsevier.com/locate/cmpb
- [6] Kaparathi, S., Mann, A., & Power, D. J. (2021). An Overview of Cloud-Based Decision Support Applications and a Reference Model. *Studies in Informatics and Control*, 30(1), 5–18. <https://doi.org/10.24846/v30i1y202101>
- [7] Marikyan, D. & Papagiannidis, S. (2021) Unified Theory of Acceptance and Use of Technology: A review. In S. Papagiannidis (Ed), *TheoryHub Book*. Available at <http://open.ncl.ac.uk/> / ISBN: 9781739604400
- [8] E. Dalinger, D. Ley, Fraunhofer Institute for Communication, „A Reference Model for Designing Decision Support Systems in Novel Work Domains“ *International Conference on Systems* (2011)
- [9] Zikos, D., & Delellis, N. (2018). CDSS-RM: A clinical decision support system reference model. *BMC Medical Research Methodology*, 18(1). <https://doi.org/10.1186/s12874-018-0587-6>
- [10] Sauro, Jeff. "A practical guide to measuring usability." *Measuring Usability LLC*, Denver 12 (2010)..
- [11] Lewis, James R. "Measuring perceived usability: The CSUQ, SUS, and UMUX." *International Journal of Human-Computer Interaction* 34.12 (2018): 1148-1156.
- [12] LEWIS, James R. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *ACM Sigchi Bulletin*, 1991, 23.1: 78-81.
- [13] LEWIS, James R. Psychometric evaluation of the PSSUQ using data from five years of usability studies. *International Journal of Human-Computer Interaction*, 2002, 14.3-4: 463-488.
- [14] Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 249-256.
- [15] ISO 9241-11:2018. Ergonomics of human-system interaction - Part 11: Usability: Definitions and concepts. International Organization for Standardization.
- [16] Pavlíček, A. Popisné modely špecifických skupín pacientov vzhľadom na ich kardiovaskulárne riziko – Diplomová práca (2022). KKUI FEI TU v Košiciach.
- [17] PAYNE, John W. Thinking aloud: Insights into information processing. *Psychological Science*, 1994, 5.5: 241-248.
- [18] J. Nielsen (2000), Why you only need to test with 5 users, Nielsen Norman Group, <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>
- [19] Gong, Y., Kang, H. (2016). Usability and Clinical Decision Support. In: Berner, E. (eds) *Clinical Decision Support Systems*. Health Informatics. Springer, Cham. https://doi.org/10.1007/978-3-319-31913-1_4

Porovnanie klasifikačných modelov z pohľadu interpretovateľnosti

Lenka DOBRANSKÁ, Tomáš ADAM, 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, tomas.adam@tuke.sk, frantisek.babic@tuke.sk

Abstrakt — Publikácia sa zaoberá vytvorením interaktívnej webovej aplikácie s navrhnutým hodnotiacim mechanizmom, ktorý dokáže určiť poradie klasifikačných modelov na základe preferencií používateľa. Hodnotiaci mechanizmus využíva multikriteriálne rozhodovanie, metódu TOPSIS a príslušné prístupy na hodnotenie výstupov dátovej analytiky. V rámci hodnotenia interpretovateľnosti rozhodovacích modelov sú použité viaceré metriky vrátane identity, separability a stability. Navrhované riešenie má potenciál zlepšiť dôveru v predikcie a zvýšiť pravdepodobnosť prijatia algoritmov strojového učenia v klinickej praxi. Z používateľského testovania vyplýva, že vytvorená webová aplikácia spolu s interpretačnými metódami Lime a Shap, poskytuje transparentné vysvetlenia rozhodovacích procesov a napomáha netechnicky zainteresovaným stranám pochopiť faktory ovplyvňujúce výsledné predikcie modelov.

Kľúčové slová — Dátová analytika, medicínska diagnostika, interpretovateľnosť

I. ÚVOD

Integráciou strojového učenia do kritických oblastí, akými sú zdravotníctvo, či výroba autonómnych vozidiel, sú rozhodnutia, ktoré kedysi prijímali ľudia, aktuálne prijímané automaticky pomocou algoritmov. S cieľom zabezpečiť spoľahlivosť takýchto rozhodnutí je potrebné pochopenie dôvodov, pre ktoré boli tieto rozhodnutia prijaté. Modely strojového učenia sú zvyčajne vo svojej podstate len čiernymi skrinkami a neposkytujú vysvetlenia, ako a prečo dospeli k daným rozhodnutiam. Vďaka tomu môžu byť tieto modely nespoľahlivé a nedôveryhodné. Na poskytnutie prehľadu o rozhodovacích procesoch týchto modelov boli navrhnuté rôzne interpretačné metódy, ako aj metriky na hodnotenie interpretovateľnosti.

V posledných rokoch je kladený väčší dôraz na uznávanie zdravotníckych riešení, ktoré sú podporované strojovým učením [1]. Zainteresované strany postupom času zisťujú, že analýza dát znižuje náklady na zdravotnú starostlivosť a prispôsobuje lekárske služby každému pacientovi jednotlivo [2]. Aby koncoví používatelia a zdravotnícki pracovníci porozumeli a následne prijali alebo odmietli predpovede, musia správne pochopiť výsledky predikčných modelov. Nájdenie spoločného jazyka na výmenu znalostí medzi lekárskeho expertom a dátovým analytikom môže byť v praxi náročnejšie ako úloha extrakcie znalostí z modelov generovaných počítačovými systémami, ktoré vykonáva analytik. Autori [3] zdôrazňujú, že lekárske expert môže od dátového analytika požadovať niekoľko požiadaviek vrátane: záruk interpretovateľnosti a vysvetliteľnosti, ktoré sú prispôbené špecifickým požiadavkám medicínskeho problému, či modelovania súladu s klinickými protokolmi a usmerneniami pre daný problém. Na druhej strane, dátový analytik môže od lekára požadovať jasné lekárske požiadavky týkajúce sa interpretovateľnosti a vysvetliteľnosti, realistické pochopenie obmedzení interpretovateľnosti a možností analytických modelov, či jasný opis skutočného procesu lekárskeho rozhodovania.

Nami navrhnuté riešenie môže byť nápomocné pri budovaní dôvery v predpovede a zvýšiť pravdepodobnosť prijatia algoritmov strojového učenia v klinickej praxi. Vďaka jednotlivým funkciám by oznamovanie dôležitých zistení z údajov netechnicky zainteresovaným stranám mohlo prebiehať jasným a stručným spôsobom.

A. Metriky zlepšujúce interpretovateľnosť modelov

Autori práce [4] sa zaoberali šiestimi agnostickými technikami interpretovateľnosti, teda takými, ktoré možno použiť pre všetky modely strojového učenia. Nakoľko nepotrebujú znalosti o vnútornom fungovaní modelu, potrebujú však prístup k údajom a výsledkom predikcií. Jednalo sa o nasledovné techniky: LIME, SHAP, Anchors, LORE, ILIME“ a MAPLE, ktoré použili na tabuľkových a textových medicínskych dátach, pričom na vyhodnotenie využili sedem rôznych

metrik: identity, stability, separability, similarity, execution time, bias detection a trust. Autori sa zamerali na model náhodného lesa, pri predikcii pozitívnej (viac ako päť hviezdíčiek) alebo negatívnej (päť a menej hviezdíčiek) recenzie v dátach, ktoré obsahovali informácie o liekoch a taktiež pri predikcii stupňa vedľajších účinkov (mierny, stredný, ťažký, závažný, mimoriadne závažný) v dátach týkajúcich sa vedľajších účinkov. V tomto prípade autori zistili, že neexistuje najlepšia technika interpretovateľnosti, ktorá by dosiahla najlepší výsledok pre všetky metriky.

Autori použili v práci [5] model logistickej regresie, ich cieľom bolo kategorizovať knihy a klasifikovať binárny sentiment IMDb recenzií. Využili dve metódy interpretovateľnosti: LIME a COVAR. Pri anotovaní recenzií zobrazovali najskôr samotné recenzie a recenzie, v ktorých boli zvýraznené niektoré slová so sentimentom na základe LIME interpretácií a následne na základe COVAR interpretácií. Autori vyhodnocovali modely interpretovateľnosti na základe metriky rýchlosti prenosu informácií (ITR) a koeficientu dôvery. Výsledkom bolo, že recenzie s vyznačenými slovami vďaka COVAR dosiahli lepšie výsledky aj v rýchlosti aj presnosti anotácií. Práve to zdôrazňuje potenciál využitia jednoduchých metód, ako je COVAR, pre modely strojového učenia. Okrem hodnotenia interpretovateľnosti ich experimenty poskytujú aj empirické dôkazy o tom, ako môže interpretovateľnosť zlepšiť interakciu medzi ľuďmi a systémami umelej inteligencie: keď im audítori poskytnú vysvetlenia, môžu viac ako zdvojnásobiť svoju produktivitu v úlohách anotácie.

Práca [6] sa zameriava na dve definície interpretovateľnosti, ktoré boli zavedené v literatúre: simulovateľnosť, čo je schopnosť používateľa spustiť model na danom vstupe a lokálnej interpretácie „čo ak“ čo predstavuje schopnosť používateľa správne určiť predpoveď modelu na základe zmien na vstupe, ak poznajú pôvodnú predpoveď modelu. Autori navrhujú metriku celkového počtu "prevádzkových operácií" (number of runtime operation counts) vykonaných modelom pri spustení, pričom tvrdia, že celkový počet ovplyvňuje schopnosť používateľa pochopiť interpretácie modelu. Autori natrénovali modely rozhodovacieho stromu, logistickej regresie a neurónovú sieť. Vďaka "čo ak" analýze zmenou vstupných hodnôt zistili, že rozhodovací strom a logistická regresia sú lokálne interpretovateľné, zatiaľ čo neurónová sieť nie. Analýzou tiež zistili, že so zvyšujúcim sa počtom opakovaných spustení účastníkom trvá dlhšie, kým lokálne interpretujú model a výsledky sú menej presné.

II. NÁVRH VLASTNÉHO RIEŠENIA

Víziou riešenia je vytvorenie interaktívnej webovej aplikácie, ktorá bude obsahovať vhodný hodnotiaci mechanizmus kombinujúci rôzne metriky miery interpretovateľnosti rozhodovacích modelov. Medzi spomínané metriky by sme zahrnuli identity, separability a stability [4]. Hodnotiaci mechanizmus by na základe multikritériálneho rozhodovania a TOPSIS metódy určil, ktorý vytvorený model je najlepší z pohľadu preferencií cieľového používateľa. Webová aplikácia by zároveň prostredníctvom interpretačných metód Lime a Shap poskytovala transparentné interpretácie, ako algoritmy dospeli k svojej predikcii. Vygenerované výsledky môžu pomôcť pri pochopení faktorov, ktoré riadia predikciu.

A. Hodnotiaci mechanizmus

Hodnotiaci mechanizmus [7] bude fungovať nasledovne:

1. Dátový analytik si zvolí preferencie medzi jednotlivými metrikami. Hodnoty využívajú stupnicu od 1 do 9, kde 1 predstavuje rovnakú dôležitosť a 9 predstavuje extrémnu dôležitosť.
2. Na základe zvolených preferencií sa vytvorí Saatyho matica porovnaním každého kritéria s každým iným kritériom a priradením číselnej hodnoty reprezentujúcej relatívnu dôležitosť jedného kritéria voči druhému [8].
3. Matica sa následne použije na výpočet súboru váh, pomocou normalizovaného geometrického priemeru, ktoré predstavujú relatívnu dôležitosť každého kritéria [8].

Ďalším postupom je použitie TOPSIS metódy, ktorej kroky sú nasledovné [9]:

4. Zostavíme rozhodovaciu maticu s alternatívami (modelmi) ako riadkami a kritériami (metrikami) ako stĺpcami. Z dát vyberáme náhodnú vzorku pacientov o veľkosti sto, pre ktorých sú vypočítané jednotlivé metriky. Parameter je voliteľný a jeho hodnotu je možné zmeniť. Priradíme skóre (získaný výpočet metriky) každej alternatíve pre každé kritérium.
5. Normalizujeme rozhodovaciu maticu.
6. Stanovíme váženú normalizovanú rozhodovaciu maticu tak, že vynásobíme normalizovanú rozhodovaciu maticu váhou každého kritéria.
7. Vypočítame ideálne najlepšie a ideálne najhoršie riešenie určením najlepších a najhorších hodnôt pre každé kritérium.

8. Vypočítame vzdialenosť každej alternatívy od ideálneho najlepšieho a ideálneho najhoršieho riešenia.
9. Z týchto hodnôt je vypočítaný relatívny ukazovateľ vzdialenosti modelu od ideálneho najlepšieho a ideálneho najhoršieho riešenia.
10. Výsledné poradie modelov zoradíme zostupne, pričom dátový analytik získa zoznam odporúčaných modelov.

B. Popis aplikácie

Vytvorená aplikácia poskytuje používateľovi viaceré analytické funkcie rozdelené do základných sekcií. Prvá sekcia predstavuje zobrazenie nahratých dát a základných štatistík, ako sú napríklad agregované dáta podľa tried, alebo priemerné hodnoty jednotlivých atribútov.

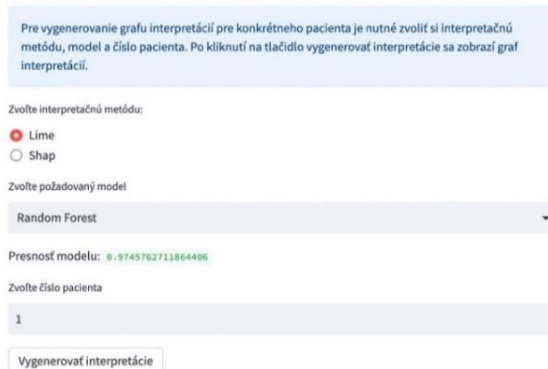


The image shows two white rectangular boxes with rounded corners. The top box contains the text 'Zobrazíť dáta' followed by a downward-pointing chevron icon. The bottom box contains the text 'Zobrazíť základné štatistiky' followed by a downward-pointing chevron icon.

Obr. 1 Funkcia zobrazenia dát a štatistík.

Druhá sekcia predstavuje generovanie grafov interpretácií pre konkrétnych pacientov. V tejto sekcií si používateľ zvolí požadovanú interpretačnú metódu a model. Po natrénovaní zvoleného modelu sa automaticky vypočíta a zobrazí úspešnosť, senzitivita a špecifickosť zvoleného modelu. Následne si používateľ vyberie pacienta, pre ktorého si želá vygenerovať interpretácie. Po vygenerovaní sa v aplikácii zobrazí graf interpretácií s príslušnými vysvetlivkami pre lepšie pochopenie výsledku.

Generovanie interpretácií pre konkrétneho pacienta

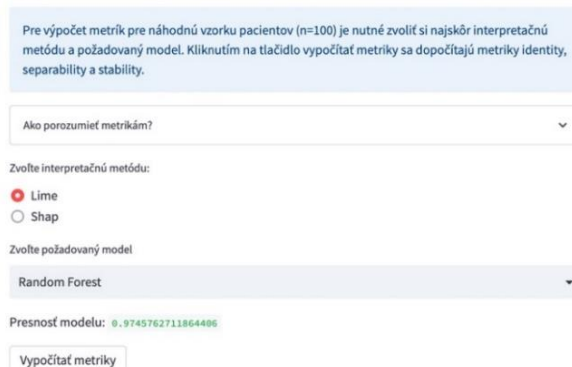


The image shows a web interface for generating patient interpretations. It features a blue header with the title 'Generovanie interpretácií pre konkrétneho pacienta'. Below the header is a light blue box containing instructions: 'Pre vygenerovanie grafu interpretácií pre konkrétneho pacienta je nutné zvoliť si interpretačnú metódu, model a číslo pacienta. Po kliknutí na tlačidlo vygenerovať interpretácie sa zobrazí graf interpretácií.' Below this are three sections: 'Zvoľte interpretačnú metódu:' with radio buttons for 'Lime' (selected) and 'Shap'; 'Zvoľte požadovaný model:' with a dropdown menu showing 'Random Forest'; and 'Zvoľte číslo pacienta:' with a text input field containing '1'. At the bottom, there is a 'Presnosť modelu: 0.9745762711864486' and a 'Vygenerovať interpretácie' button.

Obr. 2 Sekcia generovania interpretácií pre konkrétneho pacienta

Tretia sekcia predstavuje výpočet metrík pre vzorku pacientov. Funkcionalita je zameraná na výpočet metrík identity, separability a stability. Tieto metriky sú počítané pre náhodnú vzorku pacientov ($n=100$). Podobne ako v predchádzajúcej časti aplikácie si používateľ zvolí interpretačnú metódu a požadovaný model. Kliknutím na tlačidlo „Vypočítať metriky“ sa spustí výpočet a následne zobrazí výsledok.

Výpočet metrík pre vzorku pacientov



The image shows a web interface for calculating metrics for a patient sample. It features a blue header with the title 'Výpočet metrík pre vzorku pacientov'. Below the header is a light blue box containing instructions: 'Pre výpočet metrík pre náhodnú vzorku pacientov ($n=100$) je nutné zvoliť si najskôr interpretačnú metódu a požadovaný model. Kliknutím na tlačidlo vypočítať metriky sa dopočítajú metriky identity, separability a stability.' Below this are three sections: 'Ako porozumieť metrikám?' with a dropdown menu; 'Zvoľte interpretačnú metódu:' with radio buttons for 'Lime' (selected) and 'Shap'; and 'Zvoľte požadovaný model:' with a dropdown menu showing 'Random Forest'. At the bottom, there is a 'Presnosť modelu: 0.9745762711864486' and a 'Vypočítať metriky' button.

Obr. 3 Sekcia výpočtu metrík pre náhodnú vzorku pacientov

Vyhodnotenie modelov prebieha na základe zvolených preferencií od používateľa, pomocou príslušnej zvolenej interpretačnej metódy a implementovaného hodnotiaceho mechanizmu.

Hodnotiaci mechanizmus

V tejto časti je možné použiť hodnotiaci mechanizmus, ktorý na základe multikriteriálneho rozhodovania a TOPSIS metódy určí, ktorý model je najlepší.

Metriky zodpovedajú kritériam a modely alternatívam.

Zvoľte preferencie metrick:

Je nutné zvoliť preferencie metrick, teda, ktorá z dvojíc metrick je pre vás dôležitejšia.

Ako zvoliť preferencie? ▾

Preferencia identity/separability
1/3 ▾

Preferencia identity/stability
1/3 ▾

Preferencia identity/accuracy
1/3 ▾

Obr. 4 Funkcia vyhodnotenia modelov

C. Testovanie

Príprava na testovanie spočívala zo stanovenia konkrétnych testovacích úloh a vytvorenia dotazníka, v ktorom sa respondenti vyjadrili k samotnej práci s aplikáciou. V prvej časti dotazníka sa nachádzali testovacie úlohy, ktoré sa testerom podarilo vyriešiť. Dotazník ďalej obsahoval príslušné tvrdenia, ku ktorým mali zaujať svoj postoj výberom odpovede. Z ich odpovedí vyplynuli nasledujúce zistenia:

1. S realizáciou úloh testerami nemali žiadne väčšie problémy.
2. Jeden z testerov považuje aplikáciu za prehľadnú, zatiaľ čo druhý navrhuje úpravy používateľského rozhrania.
3. Jednotlivé vysvetlivky v aplikácii boli testerom nápomocné pri realizácii úloh, ako aj pri pochopení vygenerovaných výsledkov.
4. Jednotlivé vygenerované grafy pre pacientov sa testerom zdali byť prehľadné a vďaka vysvetlenému popisu aj ľahko pochopiteľné.
5. Jednému z testerov sa voľba preferencií metrick zdala byť jasná, zatiaľ čo druhý tester navrhol doplniť dodatočné vysvetlenia.
6. Testerom bolo jasné, akým spôsobom bolo generované výsledné poradie modelov.

Na základe výsledkov testovania sme zvažili možné úpravy aplikácie. Rozhodli sme sa zmeniť úvodné rozloženie aplikácie, kedy sme nástroj pre nahratie dát presunuli z bočného ľavého panelu do hlavnej časti aplikácie. Odstránili sme červenú hlášku, ktorá upozorňovala používateľa, že do aplikácie nenahral dáta - na používateľov mohla pôsobiť ako chyba aplikácie. Ďalšou zmenou bolo pridanie doplňujúcich informácií o spôsobe voľby preferencií jednotlivých metrick do časti s implementovaným hodnotiacim mechanizmom.

III. ZÁVER

Webovú aplikáciu sme vytvorili vďaka aplikačnému rámcu Streamlit v programovacom jazyku Python. Táto aplikácia pomocou multikriteriálneho rozhodovania a TOPSIS metódy určí poradie modelov na základe preferencií používateľov a umožní dátovému analytikovi komunikovať s doktorom už prefiltrovaný zoznam výsledkov. Týmto spôsobom je predkladané riešenie nápomocné pri budovaní dôvery v predpovede a zvýši pravdepodobnosť prijatia vygenerovaných rozhodovacích modelov v klinickej praxi.

ZOZNAM POUŽITEJ LITERATÚRY

- [1] Ahmad, A. M., Eckert, C., Teredesai, A., & McKelvey, G. (2018). Interpretable machine learning in healthcare. In IEEE intelligent informatics bulletin (pp. 1–7).

- [2] Wang, L., & Alexander, C. A. (2020). Big data analytics in medical engineering and healthcare: methods, advances and challenges. In *Journal of Medical Engineering & Technology* (Vol. 44, Issue 6, (pp. 267–283).
- [3] Vellido, A. (2020). The importance of interpretability and visualization in machine learning for applications in medicine and health care. *Neural Comput & Applic* 32, 18069–18083.
- [4] El Shawi, R., Sherif, Y., Al-Mallah, M., & Sakr, S. (2019). Interpretability in HealthCare A Comparative Study of Local Machine Learning Interpretability Techniques. In *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*.
- [5] Schmidt, P., & Biessmann, F. (2019). Quantifying Interpretability and Trust in Machine Learning Systems (Version 1).
- [6] Slack, D., Friedler, S. A., Scheidegger, C., & Roy, C. D. (2019). Assessing the Local Interpretability of Machine Learning Models (Version 2).
- [7] Anderková, V., & Babič, F. (2022). How to Reduce the Time Necessary for Evaluation of Tree-Based Models. In *Lecture Notes in Computer Science* (pp. 290–305).
- [8] Saaty, T. L. (1990). How to make a decision: the analytic hierarchy process. *European Journal of Operational Research*, 48(1), 9-26.
- [9] Hwang, C. L., & Yoon, K. (1981). *Multiple attribute decision making: Methods and applications*. Springer-Verlag.

FEI's digital assistant in augmented reality

Jozef DOCI, Tomáš ADAM, František BABIČ

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

jozef.doci@student.tuke.sk, tomas.adam@tuke.sk, frantisek.babic@tuke.sk

Abstract — The use of digital assistant technology in the field of universities represents a great potential, not only as a teaching aid but to help students with their assignments, improve the flow of individual lessons or as an assistant for professors. Combining these technologies can create a friendlier, more real assistant whose spectrum of abilities is broad enough to help students in every aspect of their college life. The created FEI's digital assistant offers a simple and compact way to answer frequent questions asked by students. This version like Minimum Viable Product contains questions about five subjects from Business informatics study program at Faculty of Electrical Engineering and Informatics, Technical University of Košice. The results of acceptance testing confirmed the high potential of this solution, but also pointed to the high expectations of students due to current trends in this area.

Keywords — Augmented reality, Digital assistant, Unity, Speech recognition

I. INTRODUCTION

Various technological tools are supposed to bring a higher level of entertainment, to make it more efficient and to ensure a higher level of productivity in less time and with the use of fewer resources. There have been numerous mobile applications that provide informational support or entertainment since the first augmented reality mobile application was developed in 1997. The use of augmented reality in sports, games, and entertainment shows that this approach is able to promote social interactivity and learning between users. Tourism uses augmented reality as a tool to enhance the tourist experience by providing archaeological information at cultural heritage sites [1] or to enrich exhibitions and museums [2]. In the gaming industry, the possibilities for augmented reality are great. A few years ago, a phenomenon called Pokémon Go was born and it demonstrated how important it is to use augmented reality in the gaming industry. After that, games like Minecraft Earth or Jurassic World Alive were made. In the field of medicine, a project was created [3] in which the authors focused on navigation inside a building for blind people.

Currently, several digital assistants are commonly used. Among them is the globally widespread digital assistant from Apple Siri, as well as Google Now from Google, which is found on all Android devices these days. Assistants such as Cortana on mobile devices from Microsoft, virtual assistant Braina on PC devices with the Windows operating system, Alexa from Amazon are also used. All these assistants have one essential thing in common, without which we cannot imagine such assistants nowadays. They are connected by a voice recognition system and text to speech technology. Technologies that allow the user to speak a request in natural language [4] enable the assistant to fulfill it and reply in natural language [5].

The number of universities developing and deploying virtual assistants is increasing. There are several international and British institutions using this technology. Using various platforms, users can access their schedules, monitor attendance, book appointments, and more. The digital assistant called ASK L.U. by Lancaster University centralizes information about the university in one place. This digital assistant, like assistants from well-known companies, is made up of a small portable device similar to Amazon Alexa or Google Nest. Digital assistant Becky by Leeds Beckett University serves as a chatbot with the inbuilt function of directed dialogue. It was created as one of the first digital assistants of its kind in the world. This assistant can answer students' frequently asked questions about semester dates, classes, everyday student life or navigate through campus to a certain building.

II. DIGITAL ASSISTANT

A digital assistant enabling to mediate requests and questions from the user regarding basic questions about the subjects of the Technical University of Košice and subsequently answering

them would be a great method to improve the course of individual lectures and exercises. After the incremental development and testing, we would have created an application which consists of processing natural language into text form, searching for the request in the database and afterward responding to the user in natural language. Thanks to a large database of subjects, the application would have been used in every year and semester of study, not only in the first years. The digital assistant would be guiding new users to the necessary information thanks to the technology of controlled dialogue. It would also include the function of introducing itself and explaining the basics of using the application to the user. Application would have the option of searching for questions from database directly from application. Thanks to these functions, we could simplify the work of teachers and provide general answers to students.

A. Question database

The pilot database of the created digital assistant consists of 28 questions commonly used by students. This version contains multiple questions in Slovak language about five subjects from bachelor and master study levels. The questions are organized in a form of an inference network. This network is separated into two main groups: teachers and subjects, see Figure 1.

As an example, we chose 5 questions used at user testing:

- Kto je prednášajúci na predmete... / Who is the lecturer of subject...
- Aký e-mail má prednášajúci predmetu... / What is an e-mail of the lecturer of subject...
- Aká je úspešnosť predmetu... / What is the success rate of the subject...
- Koľko заданий musím odovzdať na predmete... / How many assignments do I have to submit on subject...
- Koľko zápočtov bude na predmete... / How many tests will there be on subject...

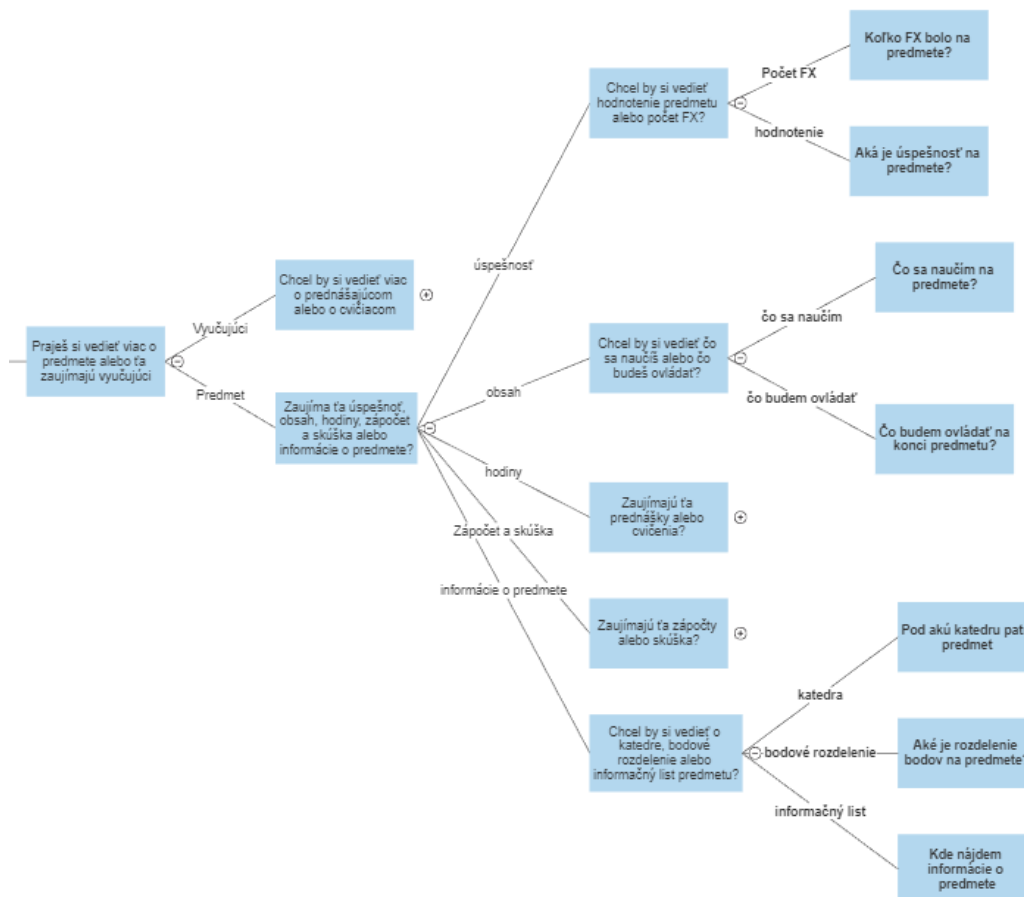


Fig. 1 Part of designed inference network

Questions were collected in cooperation with students through questionnaires, prototype testing and based on personal experiences. User can navigate through this database directly from the application through F.A.Q. panel opened by a button in the top left corner of the user interface.

B. Engine

From a technical point of view, it is important that the device meets two requirements. The first one is that the device supports ARCore technology. The full list of supported devices is

available online¹. It ensures that the device is technologically capable enough to generate augmented reality and has a powerful enough CPU to ensure good performance and real-time calculations. The second requirement is the minimum required Android API level 7.0 “Nougat” (API level 24). Activating this requirement takes place in Unity in the Player tab (Figure 2. red marked part) in the Other Settings section (Figure 2. blue marked part) and it is necessary to adjust the Minimum API Level setting. In this section, it is also possible to edit the name of the assistant or default icon (Figure 2. yellow marked part).

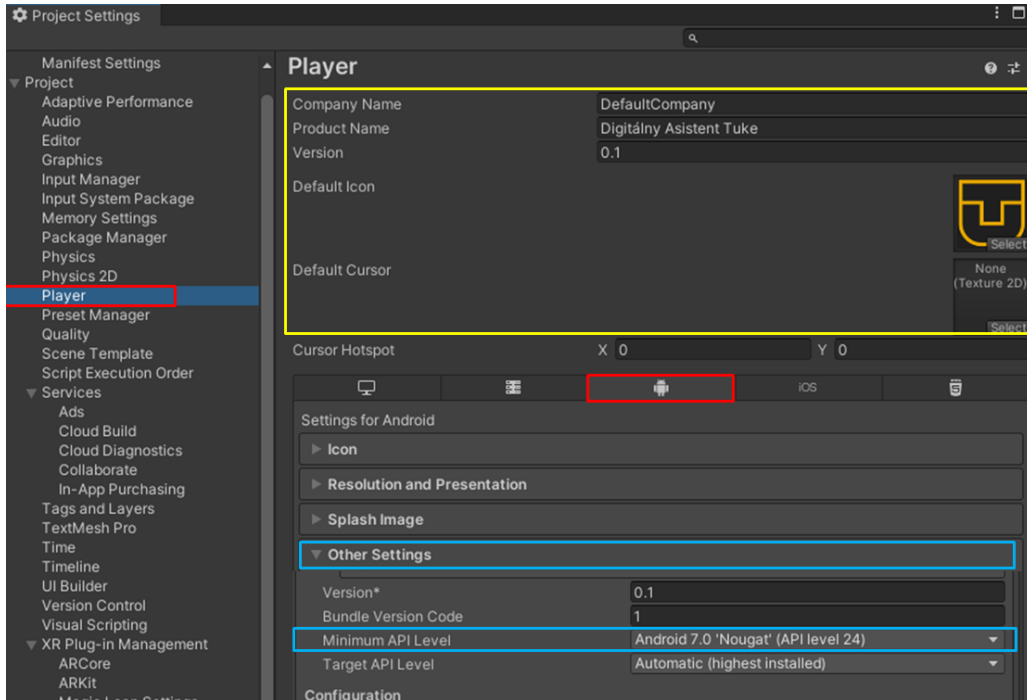


Fig. 2 Setting the minimum version of the application.

During the development of the assistant, we decided to create a marker-based augmented reality. Marker-based augmented reality displays virtual content on a predefined marker. That's why we created enough detailed marker, enabling the model to be generated.

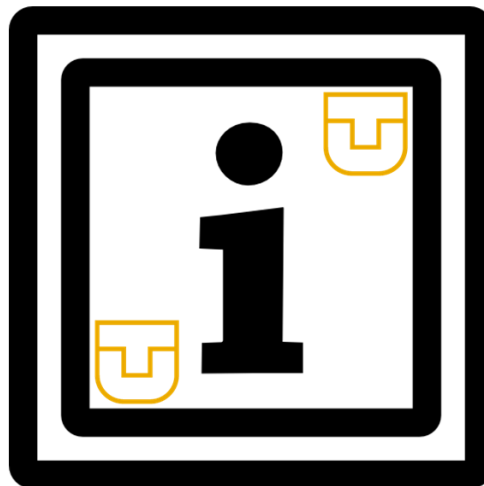


Fig. 3 Created marker.

C. User Interface

The entire user screen is made up of a camera that allows augmented reality to generate a virtual model of the assistant. The interface consists of two buttons located in the upper right and left corner, a white panel showing the assistant's response to the user, and a button located above it allowing the user to talk to the assistant. The current version of the interface has been designed to satisfy the requirements of students, with a focus on providing Slovak localization for the intended target audience.

¹ <https://developers.google.com/ar/devices>

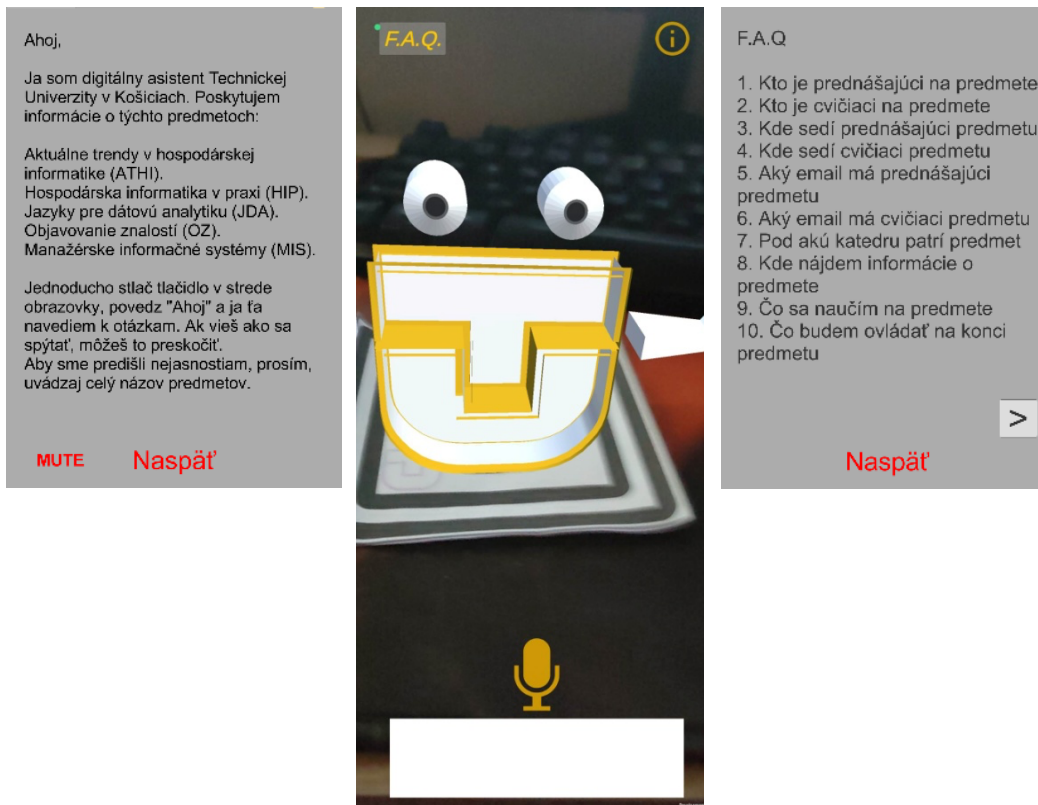


Fig. 4 User interface of digital assistant, left – information panel, right – FAQ panel, middle – user interface

When the user presses the F.A.Q. button the device screen will be cleared of other buttons and a panel with a list of questions will appear. Due to the larger number of questions in the database, the displayed list is divided into parts, therefore a button to switch pages with questions and a button to close the panel are also displayed. There are 10 questions from the database on each sheet of the list with 3 sheets total. After closing this panel, the other buttons will be displayed again.

When the user presses the information button in the upper right corner of the screen, the screen is cleared of other buttons and a panel with text introducing the assistant appears. The digital assistant will start reading this text using the text to speech technology. In the lower part of the panel, there are buttons for stopping pre-reading and closing the panel. After closing it, the other buttons will be displayed again.

A directed dialogue function has been created for new users of the application. This function guides the user to each question from the database. Directed dialogue works on the principle that the assistant gives the user a choice of several options to continue the conversation. After choosing one of them, the assistant will proceed to the next step of dialogue. This is repeated until the dialogue reaches the desired question. We will explain this function on one question from database listed above.

User question: Kto je prednášajúci na predmete objavovanie znalostí. / Who is the lecturer of the subject Knowledge discovery.

Directed dialogue process:

1. The user presses microphone button and says "Ahoj / Hello".
2. Assistant answers "Ahoj, chceš vedieť viac o predmete alebo ťa zaujímajú vyučujúci? / Hello, do you want to know more about the subject or are you interested in the teachers?"
3. The user says " Zaujímajú ma vyučujúci/ I'm interested in teachers" after pressing the microphone button.
4. The assistant will say "Chcel by si vedieť viac o prednášajúcom alebo o cvičiacom? / Would you like to know more about the lecturer or the practical teacher?"
5. The assistant replies to the user „Chceš vedieť kto to je, kde sedí alebo jeho email? / Do you want to know who he is, where he sits or his email?“.
6. The user presses the microphone button and replies "Kto to je? / Who he is?".
7. The assistant will answer "Spýtaj sa ma otázku „Kto je prednášajúci na predmete“ a názov zvoleného predmetu / Ask me the question "Who is the lecturer of the subject" and the name of the chosen subject".

After that user can ask this question and receive an answer from assistant.

III. CONCLUSION

The created application offers a simple and compact way to answer frequent questions asked by students. One of the features implemented is the ability to guide users through directed dialogue. An important factor in the development of the application was simplicity and transparency for new users as well as for frequent users of the assistant. In our future work we will focus on expanding the existing database, exchanging the technical implementation of this database and user testing within other target groups of users (different study programs, different years of study).

REFERENCES

- [1] V. Vlahakis, M. Ioannidis, J. Karigiannis, M. Tsotros, M. Gounaris, D. Stricker, T. Gleue, P. Daehne, and L. Almeida, "Archeoguide: An augmented reality guide for archaeological sites," *Computer Graphics and Applications, IEEE*, vol. 22, no. 5, pp. 52–60, 2002.
- [2] S. Feiner, B. MacIntyre, T. Hollerer, and A. Webster, "A tour- ing machine: Prototyping 3d mobile augmented reality systems for exploring the urban environment," *Personal Technologies*, vol. 1, no. 4, pp. 208–217, 1997.
- [3] Liu, Y., Stiles, N. R., & Meister, M. (2018). Augmented reality powers a cognitive assistant for the blind. *ELife*, 7, e37841.
- [4] Arora, Shipra J., and Rishi Pal Singh. "Automatic speech recognition: a review." *International Journal of Computer Applications* 60.9 (2012).
- [5] Taylor, Paul. *Text-to-speech synthesis*. Cambridge university press, 2009.

Vyvinutú metódu možno považovať za univerzálnu, pretože sa dá použiť na akomkoľvek type obrazu, či už na farebnom, v odtieňoch sivej alebo binárnom. Navrhovaná metóda zaručuje pomer kompresie 3:1 s vysokou pravdepodobnosťou a zároveň sa deklaruje vykonávanie v reálnom čase bez straty dát. Vyšší kompresný pomer je možné dosiahnuť, ak má obraz vyššiu redundanciu údajov. Princíp tejto metódy sa dá ľahko opísať nasledovnými krokmi, ktoré je nutné vykonať v procese akvizície a kódovania údajov [5]:

1. Krok Načítaj prvý riadok pixelov a nájdi jedinečné pixely.
2. Krok Vzostupne zorad' načítané pixely a pridel' im poradové číslo.
3. Krok Vytvor dvojice z pixelov a nahrad' ich príslušným poradovým číslom.
4. Krok Aplikuj kódovanie na každú zo vzniknutých dvojíc pixelov.
5. Krok Načítaj ďalší riadok pixelov a opakuj kroky 2 až 5.

Výsledok: Komprimovaný obraz

Je zjavné, že dôležitou súčasťou vyvinutého kódovania je presné usporiadanie a zoradenie pixelov. Celé kódovanie je neskôr závislé od tohto usporiadania a preto je proces akvizície a kódovania údajov veľmi dôležitý a opíšeme si ho detailnejšie.

Akvizícia údajov. Tento proces je lineárny, pretože je založený na načítavaní a spracovaní údajov riadok po riadku zľava doprava. Pracujeme stále iba s jedným riadkom, ktorý sa ukladá do pamäte. Forma do ktorej sa riadok uloží, predstavuje pole. Dĺžka riadka je limitovaná na 2048 pixelov. Zvyčajne sa takto načítaný riadok delí na bloky, v ktorých prebieha samotné kódovanie. Optimálna veľkosť bloku je definovaná na 16 pixelov. Bloky sa spracúvajú postupne až do ukončenia riadku. Proces získavania jedného riadku je rýchlejší ako doba chodu kódovacieho algoritmu. Preto sa nový riadok ukladá do pamäti, kým sa nedokončí proces kódovania predchádzajúceho riadka [11], [7].

Kódovací algoritmus. Takto načítaný riadok je potrebné upraviť tak, aby vyhovoval potrebám kódovacieho algoritmu. Načítané pixely sa zvyknú v riadku opakovať, preto je nutné nájsť tie, ktoré sa opakujú a zoradiť si ich vzostupne, od najmenšieho po najväčšie. Takto zoradením pixelom priradíme poradové číslo, ktoré je potrebné ku ďalšiemu kroku. V ďalšom kroku vytvoríme dvojice pixelov, v ktorých nahradíme ich hodnotu príslušným poradovým číslom [5].

Napríklad: Blok A_i pozostáva z takýchto hodnôt: $A_i = \{[p_{i1} - p_{i16}], [p_{i17} - p_{i32}], \dots, [p_{i(n-j)} \dots p_{in}]\}$, kde A_i je skupina pixelov riadku obrazu a p_{ij} je hodnota pixelu.

Ak počet pixelov zachyteného riadku nie je násobkom 16, konečný blok je tvorený zostávajúcimi pixelmi posledného bloku. Kódovací algoritmus sa aplikuje na každý blok alebo skupinu 16 hodnôt poľa. Po dokončení bloku sa index poľa zameriava na nasledujúcich 16 hodnôt. Táto úprava v blokovej forme prispieva k nasledujúcim výhodám: zvyšuje rýchlosť vykonávania pri práci s blokmi 16 pixelov, znižuje požiadavky na pamäť, uľahčuje vykonávanie v reálnom čase, generuje štruktúru dát, ktorá zjednodušuje aplikáciu nasledujúceho procesu kódovacieho algoritmu.

Na názornom príklade môžeme vidieť ako tento proces usporiadania pixelov prebieha a ako vyzerá finálna usporiadaná podoba blokov. Nech blok B_i je tvorený 16 pixelmi obrazu idúcimi za sebou: $B_i = \{p_{i1}, p_{i2}, p_{i3}, p_{i4}, \dots, p_{i16}\}$, ktorého hodnoty sú napr.:

$$B_i = \{84, 5, 98, 5, 16, 128, 84, 5, 98, 195, 5, 16, 84, 128, 98, 195\}.$$

V takomto bloku, ako aj v každom ďalšom je nutné nájsť pixely, ktoré sa opakujú a zoradiť ich vzostupne podľa tejto podmienky: $p_1 < p_2 < p_3 < \dots < p_n$, kde p označuje pixel a n označuje poradie pixelov. V tomto prípade by to vyzeralo nasledovne: $5 < 16 < 84 < 98 < 128 < 195$ a $n_{max} = 6$.

Takto zoradené pixely označíme poradovým číslom (Tab. 1).

Tabuľka 1 Nahradenie pixelov ich poradovým číslom

Hodnota pixelu	5	16	84	98	128	195
n(p)	1	2	3	4	5	6

Pixely každého bloku sú zoradené do skupín usporiadaných dvojíc. Hodnota pixelu určuje polohu oboch pixelov v usporiadanej dvojici. Prvý prvok v každej dvojici zodpovedá pixelu s menšou hodnotou (p_l) a druhý prvok zodpovedá pixelu s väčšou hodnotou (p_h) (Tab. 2).

Tabuľka 2 Princíp usporiadania pixelov vo dvojiciach

(p_{i1}, p_{i2})	(p_{i3}, p_{i4})	(p_{i5}, p_{i6})	...	(p_{i15}, p_{i16})
(p_l, p_h)	(p_l, p_h)	(p_l, p_h)	...	(p_l, p_h)

V našom prípade by dvojice bloku B_i boli zoradené do týchto 8 dvojíc:

$$(84, 5), (98, 5), (16, 128), (84, 5), (98, 195), (5, 16), (84, 128), (98, 195)$$

a vyzerali by podľa Tab.2 tak, ako je to v Tab. 3.

Tabuľka 3 Usporiadanie dvojíc pixelov z bloku B_i

(p_l, p_h)	(p_l, p_h)	(p_l, p_h)	(p_l, p_h)	...	(p_l, p_h)
(5, 84)	(5, 98)	(16, 128)	(5, 84)	...	(98, 195)

Na kódovanie dvojice (p_l, p_h) sa používa niekoľko rádo $n_j(p_l)$ a $n_j(p_h)$. Štruktúra kódového slova každého usporiadaného páru je tvorená postupnosťou po sebe idúcich jednotiek a núl, pričom:

- $n_j(p_l)$ definuje počet jednotiek (ak $n(p_l) = 2$, potom kódové slovo začína 11),
- $[n_j(p_h) - n_j(p_l)]$ definuje počet núl (ak $n(p_l) = 2$ a $n(p_h) = 5$, tak $5 - 2 = 3$, pridáme 000).

Kód obsahuje doplnok, tzn. pridanie jednej nuly na pravej strane na konci kódového slova.

Úplné kódovanie bloku je sekvencia kódových slov vygenerovaných každým z dvojíc pixelov. Každé kódové slovo je urobené zo série po sebe idúcich jednotiek a núl, čo nám umožňuje výrazne zmenšiť veľkosť kódového slova. V každom kódovom slove je dvojica bitov, ktorých tvar je 10, nahradená bitom 0. Blok, ktorý pozostáva zo šiestich rôznych pixelov, preto potrebuje 3 bity na pixel, čo znamená, že veľkosť kódového slova bude 48 bitov. Skúmaná metóda kódovania nám umožňuje zmenšiť túto veľkosť na 42 bitov. Cieľom aj tejto opísanej metódy je dosiahnuť čo najlepší kompresný pomer podobne, ako ho dosahujú stratové kompresné techniky kódovania.

Potom kódovacia tabuľka jednotlivých dvojíc vyzerá nasledovne (Tab. 4). Kompresný pomer bude samozrejme závisieť od kombinácie dvojíc pixelov v každom obrázku. Dosiahnuté výsledky ukazujú, že požadovaný počet bitov na kódovanie bloku 16 pixelov je 42 bitov, čo predstavuje hodnotu 2,625 bitov na pixel.

Tabuľka 4 Výsledná kódovacia tabuľka bloku B_1

Dvojica	$n(p_l)$	$n(p_h)$	Kódové slovo	Dvojica	$n(p_l)$	$n(p_h)$	Kódové slovo
(5,84)	1	3	1000	(98,195)	4	6	1111000
(5,98)	1	4	10000	(5,16)	1	2	100
(16,128)	2	5	110000	(84,128)	3	5	111000
(5,84)	1	3	1000	(98,195)	4	6	1111000
Počet bitov= 42							

Na základe týchto skutočností môžeme tvrdiť, že skúmaná metóda skutočne dosahuje úsporu bitov. Kompresný pomer tejto metódy samozrejme stále bude závisieť od jednotlivého usporiadania pixelov v načítanom obraze, no môžeme s istotou tvrdiť, že bude dosahovať lepšie výsledky ako Huffmanov kód či aritmetické kódovanie.

III. MODIFIKOVANÉ HUFFMANOVE KÓDOVANIE OBRAZU

Huffmanov kód je pomerne jednoduchá a zároveň efektívna metóda kompresie obrazov, pričom sa považuje za jednu z najčastejšie používaných metód bezstratovej kompresie. Modifikované Huffmanove kódovanie garantuje zlepšenie kompresného pomeru o približne 20% voči klasickému spôsobu Huffmanovho kódovania. Ak binárny strom je príliš veľký, tak aj samotná veľkosť samotného Huffmanovho kódu digitálneho obrazu pre každý pixel bude veľká. Špeciálne ak sa bude jednať o obrázky, v ktorých je veľký výskyt symbolov s nízkou pravdepodobnosťou výskytu. Princíp tejto modifikovanej metódy spočíva v znížení počtu bitov na pixel, čo sa dá dosiahnuť rozkladom binárneho stromu na menšie časti a následnou aplikáciou pôvodného algoritmu na takéto bloky.

Algoritmus, akým táto metóda kóduje obrázky sa dá veľmi ľahko zhrnúť do nasledujúcich krokov [8], [10]: 1) Načítaj obraz v odtieňoch šedej; 2) Rozdeľ zotriedenú maticu koeficientov (M) do 4 menších submatic (A, B, C, D); 3) Zostupne zorad' koeficienty v závislosti od frekvencie ich výskytu v obraze; 4) Aplikuj Huffmanove kódovanie na každú zo 4 submatic (A, B, C, D) samostatne; 5) Prirad' každému kódovému slovu označenie skupiny bez toho, aby narástol počet bitov v kódovom slove; 6) Nahrad' kódové slovo, ktoré zaberá viac ako 8 bitov pôvodným koeficientom; 7) Nahrad' každý koeficient pôvodného obrazu príslušným kódovým slovom zo submatic A, B, C, D. Výsledok: Komprimovaný obraz.

Každý krok algoritmu naznačuje, že ide o jednoduché operácie, ktoré sa budú vykonávať v určitom poradí. Už na prvý pohľad je zjavné, že nejde o žiadnu radikálnu zmenu či transformáciu pôvodného algoritmu Huffmanovho kódovania. Hlavný prínos, ktorý prináša táto opísaná metóda je dosiahnutá rozkladom binárneho stromu, ako už bolo spomenuté vyššie. Druhý aspekt, ktorý táto metóda prináša je nahradenie kódových slov, ktorých veľkosť by mohla presiahnuť 8 bitov za pôvodnú hodnotu bitového kódu (hodnota načítaného pixelu zakódovaná klasickým Huffmanovým kódom).

Algoritmus pracuje výhradne s obrazmi, ktoré sú v odtieňoch šedej. Ďalej sa takto načítaný obraz spracuje do matice, ktorá sa následne rozdelí do 4 submatic zostupne, podľa početnosti výskytu jednotlivých pixelov, ktorú je potrebné vypočítať. Na takto rozdelené pixely sa aplikuje klasický Huffmanov algoritmus. Po zakódovaní pixelov sa v jednotlivých submaticiach nahradia prvé dva bity v každom kódovom slove podľa schémy:

- kódové slová v submatici A začínajú bitmi „00“,
- kódové slová v submatici B začínajú bitmi „01“,
- kódové slová v submatici C začínajú bitmi „10“,

- kódové slová v submatici D začínajú bitmi „11“.

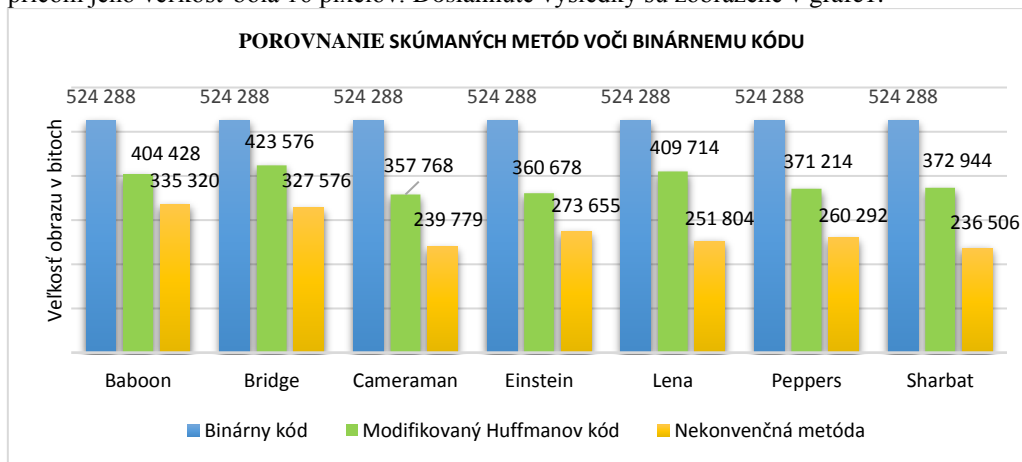
Tento postup zabezpečí, že sa kódové slová v jednotlivých submaticiach nebudú opakovať ako aj v prípade využitia klasického postupu kódovania. V snahe zvýšiť kompresný pomer dochádza ku zámene kódového slova, ktorého veľkosť presahuje 8 bitov. Tak, ako bolo uvedené vyššie, nahrádza sa pôvodnou bitovou hodnotou získanou Huffmanovým kódovaním. Posledným krokom algoritmu je nahradenie pôvodných pixelov príslušným kódovým slovom z jednotlivých submatic. Takto vzniknuté submatice sa neskôr využívajú aj v procese dekódovania [8].

IV. DOSIAHNUTÉ VÝSLEDKY A ICH ZHODNOTENIE

Softvérová aplikácia na testovanie nekonvenčnej metódy kódovania a modifikovanej metódy Huffmanovho kódovania bola vytvorená v programovom prostredí MATLAB. Aplikácia umožňuje výber ľubovoľného vstupného obrazu. Ak bol načítaný farebný obraz, následne bol prekonvertovaný na viacúrovňový šedý obraz. Ďalej aplikácia umožňuje výber z piatich metód entropického kódovania – okrem uvedených dvoch skúmaných metód aj možnosť výberu Huffmanovho, aritmetického a binárneho kódovania a to kvôli možnosti porovnania [9]. Čo sa týka možnosti výberu veľkosti kódového bloku, tie sú: 4, 16, 64 a 256. Vstupnými obrazmi boli štandardné testovacie viacúrovňové šedé obrázky veľkosti 256 x 256 pixelov typu bmp a to: Baboon, Lena, Einstein, Bridge, Cameraman, Peppers, Sharbat. Výsledky sme porovnávali z viacerých hľadísk. Faktory, pomocou ktorých sme metódy porovnávali:

- celková veľkosť zakódovaného obrazu v bitoch,
- vplyv veľkosti kódovacieho bloku na celkovú veľkosť zakódovaného obrazu,
- počet bitov na pixel,
- kompresný pomer,

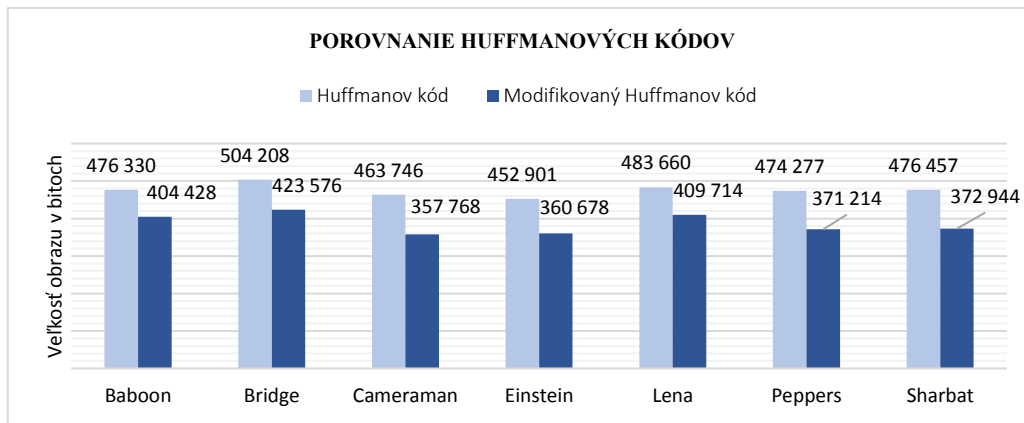
Prvým skúmaným parametrom bola celková veľkosť zakódovaného súboru v bitoch. Je to jeden z podstatných parametrov, ktorý bude v budúcnosti kľúčový pri uchovávaní obrazov, ktoré sa snímajú stále novšími a výkonnejšími snímačmi. V rámci tohto parametra sme sa zamerali na porovnanie známych metód voči metódam, ktoré v tejto práci skúmame. Vychádzali sme z údajov, ktoré sme získali z vykonaných simulácií. Spoločný parameter, ktoré všetky kódovania obsahovali bol počet pixelov v kódovom bloku, ktorý sme nastavili rovnaký pre všetky metódy, pričom jeho veľkosť bola 16 pixelov. Dosiahnuté výsledky sú zobrazené v grafe 1.



Graf 1. Porovnanie skúmaných metód voči binárnemu kódu.

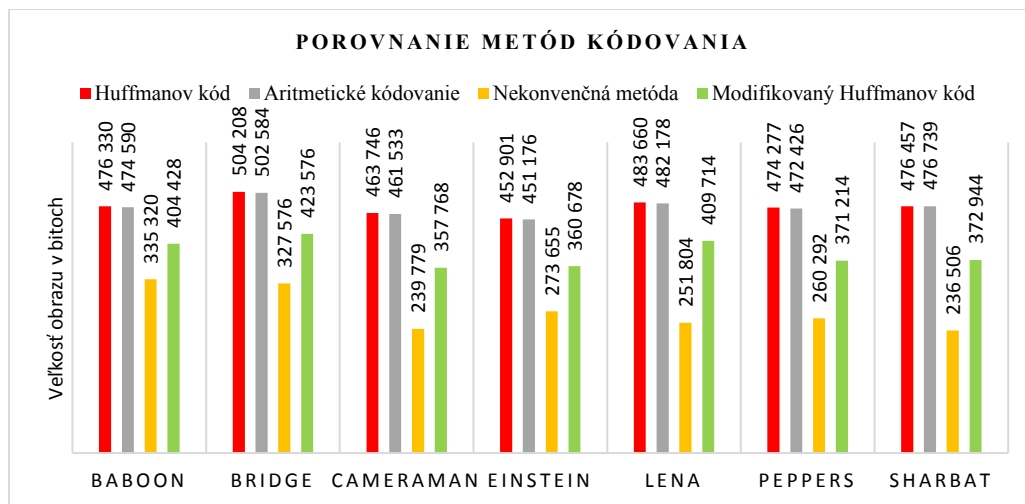
Po pohľade na tento Graf 1. môžeme tvrdiť, že úspora bitov dosiahnutá v metódach, ktoré sú predmetom skúmania v rámci tohto článku je zrejmalá. Percentuálne dochádza k úspore na úrovni 47,55 % pri využití nekonvenčnej metódy kódovania a pri využití modifikovaného Huffmanovho kódu dochádza k úspore vo výške 26,42% voči binárnemu kódu. Ak porovnáme dve skúmané metódy navzájom, tak nekonvenčná metóda dosahovala o 21,13% väčšiu úsporu dát ako modifikovaný Huffmanov kód.

Ak by sme chceli porovnať Huffmanov kód s jeho vylepšenou verziou nazvanou Modifikovaný Huffmanov kód (kap.III.), tak na grafe 2. vidíme, že modifikáciou algoritmu sa dosiahol želaný výsledok. Úspora bitov medzi starým spôsobom kódovania a skúmaným sa pohybovala na úrovni 19,01%. To znamená, že využívaním vylepšeného algoritmu dosiahneme rovnaký výsledok, čo sa týka kvality výsledného zakódovaného obrazu a zároveň dokážeme ušetriť úložný priestor.



Graf 2. Porovnanie Huffmanových kódov.

Ďalšie porovnanie bolo medzi metódami, ktoré sa už v praxi používajú a tými, ktoré sme skúmali. Z grafu 3. vyplýva, že metódy entropického kódovania (Huffmanovho a aritmetického), ktoré sa v praxi bežne používajú, dosahujú navzájom porovnateľné výsledky, pričom o čosi úspornejšia bola aritmetická metóda. Ak ich porovnáваме skúmanými metódami, tak vidíme, že vývoj nových metód kódovania mal zmysel. Dochádza k úspore dát na úrovni 20-40 %, v závislosti na dvojiciach metód, ktoré by sme medzi sebou porovnávali.



Graf 3. Porovnanie metód kódovania.

Pri porovnávaní metód kódovania bol ďalším skúmaným parametrom počet bitov na pixel (BPP), ktorý sa využíva pri vyjadrovaní veľkosti zakódovania grafického súboru a je úzko spätý s rozmerom obrazu. Zhrnuté výsledky sú v Tab.5.

Tabuľka 5 Priemerné hodnoty počtov bitov na pixel

Metóda kódovania	Binárny kód	Huffmanov kód	Aritmetické kódovanie	Nekonvenčná metóda kódovania	Modifikovaný Huffmanov kód
BPP	8	7,261	7,234	4,294	5,919

Ako môžeme vidieť, tak hodnoty v Tab.5. potvrdili predpoklady, pretože nižšia hodnota BPP priamo súvisí s celkovou veľkosťou obrazu v bitoch. Nekonvenčná metóda opäť dosiahla najlepší výsledok na úrovni 4,294 bit/pixel a modifikovaný Huffmanov kód skončil hneď za nekonvenčnou metódou s hodnotou 5,919 bit/pixel.

Pomer kompresie údajov je mierou relatívneho zmenšenia veľkosti reprezentácie údajov produkovanej daným kompresným algoritmom. Zvyčajne sa kompresný pomer vyjadruje ako pomer medzi nekomprimovanou a komprimovanou veľkosťou súboru. Skúmali sme aj veličinu, ktorá bola odvodená z hodnoty kompresného pomeru. Šetrenie miesta je zadaná veličina, ktorá skúma zmenšenie veľkosti súboru v bitoch, vzhľadom na jej nekomprimovanú časť. Vyjadruje sa v percentách (rov.1).

$$\text{Ušetrené miesto} = 1 - \frac{\text{Komprimovaná veľkosť súboru}}{\text{Nekomprimovaná veľkosť súboru}} * 100\% \quad (1)$$

Tabuľka 6 Priemerné kompresné pomery

Metóda kódovania	Binárny kód	Huffmanov kód	Aritmetické kódovanie	Nekonvenčná metóda kódovania	Modifikovaný Huffmanov kód
Kompresný pomer	1	1,103	1,106	1,94	1,364

Hodnoty v Tab.6 odzrkadľovali všetky doterajšie zistenia a napĺňali predpoklady o tom, že skúmané metódy skutočne dosahovali deklarované predpoklady. Obrazy po zakódovaní nekonvenčnou metódou mali celkový kompresný pomer na úrovni 1,94:1, čo predstavovalo zmenšenie veľkosti obrazu v bitoch na približnú polovicu veľkosti, ktorá by bola získaná po zakódovaní binárnym kódom. Výsledná veľkosť obrazov v bitoch po zakódovaní Modifikovaným Huffmanovým kódom 1,364 krát menšia, ako hodnota, ktorá by bola získaná po zakódovaní binárnym kódom.

Tabuľka 7 Ušetrené miesto

Metóda kódovania	Binárny kód	Huffmanov kód	Aritmetické kódovanie	Nekonvenčná metóda kódovania	Modifikovaný Huffmanov kód
Ušetrené miesto	0 %	9,22 %	9,50 %	47,55 %	26,42 %

Ak sa pozrieme na hodnoty v Tab.7, ktoré poukazujú na percentuálne ušetrené pamäťové miesto na disku, tak vidíme, že vývoj a výskum v oblasti nových metód kódovania dosahuje prijateľné výsledky, ktoré sa odrážajú v účinnosti daných metód.

V. ZÁVER

Článok bol zameraný na opis princípu a implementácie nekonvenčnej metódy kódovania pri bezstratovej kompresie statických obrazov a modifikovaného Huffmanovho kódovania. Výsledkom bola vytvorená simulácia, pomocou ktorej môžeme porovnávať dané metódy kódovania s inými známymi metódami. Simulácie dokázali, že skúmanými metódami skutočne dosiahneme výraznú úsporu bitov na pamäťovom úložisku. Nekonvenčná metóda kódovania je využiteľná pre akýkoľvek typ obrazu, pričom sa zachováva výrazná úspora bitov. Modifikovaný Huffmanov kód je takisto úspornejší oproti klasickému Huffmanovmu kódu. Celková úspora bitov nekonvenčnej metódy voči binárnemu kódu sa pohybovala na úrovni 47,55%, pričom úsporu 26,42 % voči binárnemu kódu dosiahla aj modifikovaná verzia Huffmanovho kódu. Medzi skúmanými metódami dosiahla nekonvenčná metóda o 21,13% vyššiu úsporu ako modifikovaný Huffmanov kód. Ak porovnáme metódy Huffmanových kódov, tak modifikáciou algoritmu sme dosiahli úsporu 19,01% voči klasickému algoritmu. Samotné porovnanie skúmaných metód nám ukázalo, že nekonvenčná metóda kódovania dosahuje vyššiu úsporu bitov pri dodržaní predpísaného počtu bitov v kódovom bloku.

Na koniec záveru patrí poďakovanie Ing. Jakobovi Lukáčovi, ktorý sa podieľal na riešení úloh z uvedenej problematiky.

LITERATÚRA

- [1] D.D. Panda – D. Swain, "Competency assessment of image compression in the lossy and lossless domain." *Proc. of Fourth International Conference on Image Information Processing (ICIIP 2017)*, 2017, pp.1-4.
- [2] M.A Mydin et al., "A Study on Medical Image Compression Techniques based on Huffman Coding and Discrete Wavelet Transform (DWT). *IEEE 9th Conf. on Systems, Process and Control (ICSPPC 2021)*, 2021, pp.86-91.
- [3] S. Jancy – C. Jayakumar "Various lossless compression techniques surveyed." *Proc. of Third Inter. Conf. on Science Technology Engineering & Management (ICONSTEM)*, 2017, pp.65-68.
- [4] J. Mihalík: *Efektívne kódovanie obrazov. LDIPV, FEI TU Košice*, 2013.
- [5] J. I. Larrauri – E. Kahoraho „A New Method for Image Coding in Computer Vision.“ *Proceedings EC-VIP-MC 2003. 4th EURASIP Conference focused on Video/Image Processing and Multimedia Communications (IEEE Cat. No.03EX667)*, 2003.
- [6] I. El-Feghi et al. „New Lossless Image Compression Technique Using Adaptive Block Size.“ *Int. Jour. of Circuits, Systems and Signal Processing*, vol.8, 2014, pp.131-136.
- [7] K. Sharma – Gupta, K.: Lossless data compression techniques and their performance. *Proc. of Inter. Conf. on Computing, Communication and Automation (ICCCA)*, 2017, pp. 256–261.
- [8] M.A. Al-Khalayleh – M. Otair „A New Lossless Method of Image Compression by Decomposing the Tree of Huffman Technique.“ *Int. Journal of Imaging and Robotics*, Jan. 2015, pp. 79-96.
- [9] E. Bodden – M. Clasen – J. Kneis "Arithmetic Coding Revealed. A guided tour from theory to praxis." McGill University School of Computer Science Sable Research Group, Sable Technical Report, 2007.
- [10] A. Odat – M. Otair – M.A. Al-Khalayleh „Comparative study between LM-DH technique and Huffman coding technique.“ *Int. Journal of Applied Engineering Research*, vol.10, 2015, pp. 36004-36011.
- [11] J.I. Larrauri, „A New Algorithm for Lossless Compression Applied to Two-dimensional Static Images.“ *Int. Conf. on Intelligent Cloud Computing*, March 2012.

Virtual reality environment as a basis for humanoid robot training and testing system

¹Branislav SOBOTA, ²Miriama MATTOVÁ, ³Soňa MATULOVÁ, ⁴Štefan KOREČKO

^{1,2,3,4}Department of Computer and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

¹branislav.sobota@tuke.sk, ⁴miriama.mattova@tuke.sk

Abstract — The aim of this article is to provide more detailed information about a system designed for training and testing humanoid robots in virtual reality. Nowadays, humanoid robots are being developed to perform various tasks, which are often associated with interaction with people (Human-robot interaction, HRI). Standard repeated learning may damage the robot or its components. This can be costly and dangerous. Therefore, an application has been created that enable the simulation of such humanoid robots within virtual reality. With the help of virtual robot models, the required movements are performed. After obtaining the necessary results from these simulations, the data are transferred to a real robotic device that can then be used. In this way, it is possible to effectively train and test even humanoid robots. All in all, virtual reality makes it possible to simulate various situations and scenarios, and thus test the functioning of humanoid robots in various conditions and environments without the risk of damage.

Keywords — humanoid robot, virtual reality, training system, A-frame,

INTRODUCTION

Robots have become an important part of our daily lives in recent years. Their main task is to help a human with activities that are either too difficult or dangerous for a human. So far, the primary deployment of a robots is in an industry. However, there is also a field of a robotics that deals with humanoid robots. They must generally behave like people and they should be part of people's normal life in the future. Because of this use, these robots must be able to successfully interact with people and help them in various situations. Then it is necessary to solve the human-robot interaction (human-robot interaction, HRI). HRI has demonstrated a rapid progress in the last decade [1]. Standard human-human interactions are of different types and in modern society they are gradually being extended to interactions between humans and robots. This interaction covering three “co-actions”: a cooperation, a communication and a collaboration. Transparent HRI supports the capability to read the partner's intentions that leads to the theory of mind, involving the mentalizing network [2]. An equally important aspect is the safety of robot actions for the human partner [3]. The robots with humanoid shape have been proposed as a plausible controllable tool for studying social interactions [4]. At the same time, humanoid robots have been demonstrated to play an increasingly important role in human society, with their applicability across many tasks or contexts [5]. This condition requires processing a large amount of information at once, which can lead to erroneous output.

It is necessary to realize the main difference between robots used in industry and those that we will use in homes. The current industrial robot generally responds to a precisely determined input and generates a precisely determined output. However, humanoid robots must be able to recognize human emotions and make decisions based on them (e.g. by using of artificial intelligence). They will need to know when a human is happy, when he is afraid or in pain. This part is very difficult from a psychological point of view, because not all people express emotions in the same way, and sometimes we don't even know if a given human is happy, in pain or stressed until we know the cause of his expression.

These activities are also part of robot learning. The learning process consists of repeating movements or interactions. However, there is a high probability of damaging the robot's components or performing a movement that could cause harm or even endanger the interacting human. For this reason, simulators were created in which a digital twin of the robot is imported, on which it is possible, for example, to learn movements without the risk of damage. The most used applications for robot simulation are CoppeliaSim [6], Gazebo [7] or Unity engine [8]. When

choosing the right software, it is important to find out if it can work with ROS and formats such as URDF or SDF. ROS [9] is an operating framework for robots that is used in their development. The URDF format defines the appearance and the robot kinematics itself, while the SDF can also describe the overall scene into which the robot is imported. A good knowledge and use of these tools can help in the development of better and more efficient robotic systems.

One such humanoid robot is NICO (Neuro-Inspired Companion) [10]. This robot was designed for neuro-cognitive research and human-robot interaction. The inspiration for its production was the robot iCub [11]. The iCub has the appearance of a three- to four-year-old child to make it more intuitive for humans and easier to teach new activities, just like children. This approach mimics a parent-child or teacher-student relationship. Research at NICO is mainly focused on grasping objects and emotions recognition currently. To learn new activities, the mentioned CoppeliaSim is used, into which its URDF file is imported. The development of the NICO robot tries to overcome barriers in human-robot interaction, and it opens up new possibilities in the field of neuro-cognitive research. In addition to simulators such as CoppeliaSim, virtual reality and related technologies are also one of the interesting way for humanoid robot training.



Figure 1 NICO Robot (leg-less version)

Compared to physical HRI, VR has many advantages [12]: high flexibility in setting the environment, good control of independent variables, safety or the possibility of teleoperation. Then, for the effective cooperation of workers from different locations around the world on one humanoid robot, the use of a collaborative virtual reality system is appropriate. Collaborative virtual environments are open to the possibilities of various trainings, education, simulations or other types of activities. Collaborative virtual environments at this level are implemented using virtual caves or virtual-reality headsets systems. In the case of virtual caves, the applicable virtual cave is LIRKIS CAVE (Figure 2) available in the authors' home laboratory (LIRKIS laboratory at DCI FEEI TU Košice). This laboratory disposes also several devices at the level of VR headsets such as Oculus Quest or MS Hololens. At software level, LIRKIS G-CVE (LIRKIS Global - Collaborative Virtual Environments) system [13] [14] also developed in the authors' home laboratory can be advantageously used. This system is a web-based system, built on top of the A-Frame web VR framework and its extension, called Networked-Aframe (NAF). Thanks to A-Frame, it is possible to run LIRKIS G-CVE on various user devices, from desktop computers and notebooks through tablets and smartphones to virtual and mixed reality headsets. In the following will be short description of the pilot application parts using LIRKIS-GCVE for visualization simulation movement of a humanoid robot NICO in the virtual environment.



Figure 2 LIRKIS CAVE

THE SYSTEM PROPOSAL AND IMPLEMENTATION

The implementation of the system for the humanoid robot NICO was successfully implemented using the A-Frame framework on the Glitch web platform. The choice of these technologies was motivated by their optimal use, which enabled the efficient implementation of the system. A-Frame is known for its easy implementation of the virtual environment. Such an implementation provides developers with an optimal way to create virtual reality. On the other hand, Glitch provides developers with an editor for creating web applications and enables optimal launch and control of their behavior. The system was designed based on the officially defined humanoid robot NICO, which was described in the URDF file. The *gITF* robot model was used in the implementation. This made it possible to work efficiently with the robot model. In order for the system to be able to work like already created programs, it was necessary to incorporate requirements for the rotation of robot parts, for recording and playback of this action. Other tasks included changing the robot's head to affect its emotional expression and changing its field of vision and the space it was in. Next, it was necessary to transfer the values from the URDF file to JSON format so that the rotations of the robot parts take place according to its official kinematics and so that they can be used in the created web application.

The user can perform these processed requests using the control menu located within the virtual scene (Figure 3). The menu contains buttons that are used to perform a certain action. Each of the groups of buttons represents a specific functionality. For each of them, an A-Frame component is defined, determining what action should be performed after clicking the button. Additionally, the user can interact with objects in the scene using head or hand movements, providing a more intuitive experience using the app.

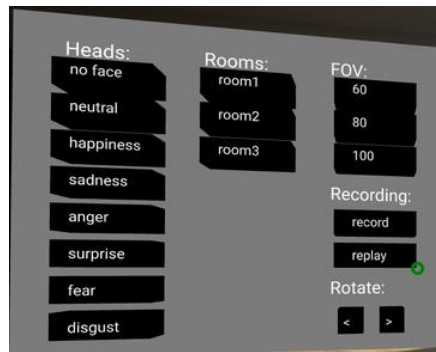


Figure 3 The application main menu

The following diagram (Figure 4) shows use-case that demonstrate the above functions and allow interaction with the virtual robot and its virtual environment. These use-cases are important to verify the functionality and proper functioning of the system in real time.

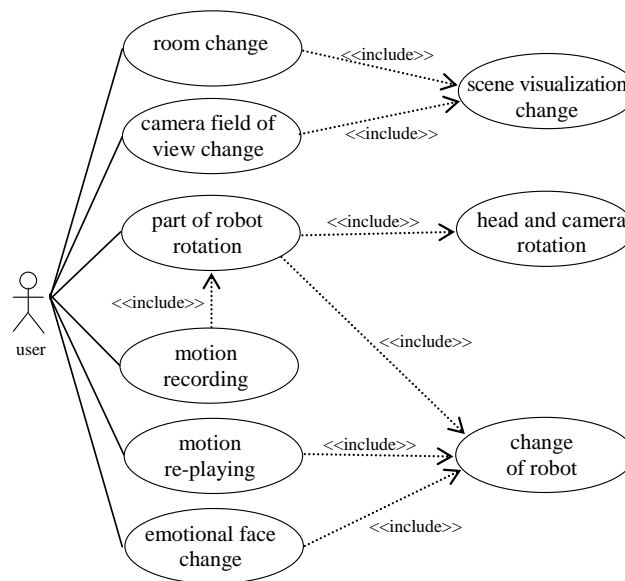


Figure 4 The use case of the application

If the user performs the first or second use case, he can change the visualization of the scene. This means he can change the model of the room he's in or change the camera's field of view. The camera is placed in NICO's head, so the user can see the scenery from its point of view. These changes can help the user better navigate the environment and interact more effectively with the robot.

Next cases describe the robot updates that allow you to rotate some part of its body or change its texture. When the head rotates, the camera placed in its place also rotates automatically. Rotation is performed by clicking on a robot body part and using the arrow keys in the control menu. Another option is to select a limb and use the "p" and "l" keys. An important part of this system is the ability to record the robot's movements and then playback them.

A. Robot movement recording and playback

The robot movement can be recorded using the "record" button. The first click starts recording the movement of the robot. In this way, the positions of individual parts of the robot's body are captured after each change made. After pressing the "record" button a second time, a window for saving the JSON file to the selected folder will appear. After opening it, it will be possible to see the name of the robot part, its position and the time in seconds since the start of the recording (Figure 5). This file is then used during playback.

```
{
  "boneName": "head_z",
  "boneposx": -0.00593295640955391,
  "boneposy": -0.0655391169731177,
  "boneposz": -0.09034472209643907,
  "timepos": 1
},
{
  "boneName": "head_y",
  "boneposx": -0.004478377240783857,
  "boneposy": 0.10107277833624664,
  "boneposz": 0.12607077024313987,
  "timepos": 1
},
```

Figure 5 The example of record file

The movement will be replayed if the user clicks on the "replay" button. In this case, the user should see the movement of the robot captured using the record button. After performing the overall recording, the button must be clicked once more to return the NICO to its original state, to the initial position after import. These three functions create the core of simulation program and they are an important part of this system.

B. Robot emotional state implementation

The last function that can change the model of the robot imported into the system is changing its head. As already mentioned, NICO has to change its emotions because it is also used for research on human-robot interaction. NICO can display seven different emotional states: neutral, happy, sad, angry, surprised, terrified and disgusted. It is possible to switch between them using the created web application. To implement this requirement in the system, seven new heads were created, which are part of the NICO model. These heads were created using Blender. The implementation of this functionality consists in displaying only the head that represents the selected emotional state from the control menu.

C. Collaborative virtual reality implementation

The Socket.IO library was used for the implementation of collaborative virtual reality. Communication was established using it. Thanks to it, users connected to the virtual environment can work together with the given robot and see in real time the same changes made by a certain user. The basis is the use of a client-server architecture. For example, if one of the collaborate clients changes the coordinates by rotating the NICO's hand, the new positions are sent to the server. Then it sends this change to all clients connected to the system. In this way, all important data that needs to be sent to other clients are implemented. Avatar technology is used so that users can see each other. They use them to recognize themselves and, thanks to it, they can interact socially with others. This principle works in web applications thanks to the Networked A-frame library. Users can thus observe the movement and rotation of another participating user moving around the room.

SYSTEM EVALUATION

After each requirement was implemented, the application was tested. During this evaluation, the lacks in collaborative virtual reality were identified. Those lacks were related to the sharing of data between clients if a larger number of users entered the virtual environment. This was the only functionality where problems were encountered.

After evaluation the necessary functions, there were only comments regarding the appearance of the user interface. The main complaint was that the web app did not include any instructions or explanations when using more complex functions such as recording and playback. It was also found that it is not stated anywhere that you need to click the "p" and "l" keys on the keyboard to rotate the robot part. As a last comments, it was impossible to choose your own avatar with your own name. To improve the user experience, the explanations and instructions have been added to functionalities that were considered more difficult. The option to choose your own avatar and add your own name was also added. This will allow users to better enjoy collaborative virtual reality.

CONCLUSION

The presented system is intended for training and testing humanoid robots in a virtual reality environment. It focuses on a specific NICO robot model. This system allows the rotation of individual parts of the robot in accordance with the official kinematics stored in its URDF file. The system has also the function of recording and playing back the simulation. In addition, it also allows changing the appearance of the robot e.g. the field of vision and the space in which the robot is located. Its purpose is not only to improve and optimize the robot's movements, but also to explore its interaction with different environments. The system is capable of simulating various situations and environments (including collaborative environments), allowing the robot to gain significant experience without the need for physical contact with the environment.

This system can only be used as desktop VR currently, but it is expected to be used on other devices in the future. This would make it possible to get closer to real conditions and increase the effectiveness of robot training. In addition, it would be advisable to focus on the improvement of collaborative virtual reality, because with a larger number of users, the system cannot handle sharing the necessary data between clients. The system will also be expanded with other options, such as training and testing different humanoid robots (different URDF), creating custom environments for simulations, or the possibility of running the system on different devices and platforms.

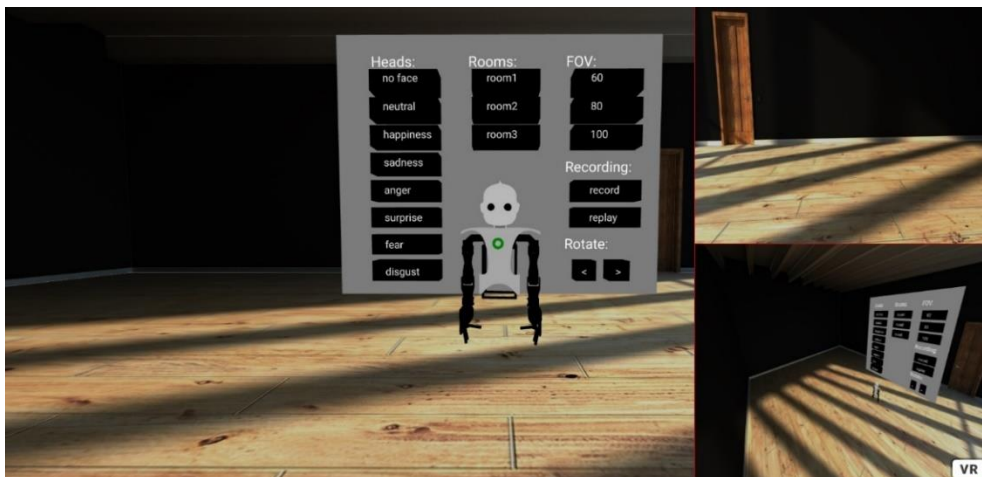


Figure 6 The example of the application including robot's camera and overview camera views

Overall, such a system is an excellent tool for training and testing humanoid robots in a virtual reality environment. Its potential could be used in various fields, therefore it is important to pay attention to its development and improvement to make it even more effective and increase its impact on modern technological solutions both in the field of robotics and virtual reality.

ACKNOWLEDGMENT

This work has been supported by the APVV grant no. APVV-21-0105 "Trustworthy human-robot and therapist-patient interaction in virtual reality".

REFERENCES

- [1] Goodrich M.A., Schultz A.C. (2008). Human-Robot Interaction: A Survey. *Foundations and Trends in Human-Computer Interaction*, 1(3), pp. 203-275.
- [2] Vinanzi S., Cangelosi A., Goerick C. (2021). The collaborative mind: intention reading and trust in human-robot interaction. *Iscience*, 24, <https://doi.org/10.1016/j.isci.2021.102130>
- [3] Lasota P.A., Song T., Shah J.A. (2017). A Survey of Methods for Safe Human-Robot Interaction. *Now Foundations and Trends*, <https://doi.org/10.1561/23000000052>
- [4] Sciutti A., Sandini G. (2017). Interacting with robots to investigate the bases of social interaction. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(12), 2295-2304.
- [5] Bahishti A.A. (2017). Humanoid Robots and Human Society. *Advanced Journal of Social Science*, 1(1), pp. 60-63
- [6] Rooban S., Suraj S. D., Vali S. B., and Dhanush N., "Coppeliassim: Adaptable modular robot and its different locomotions simulation framework," *Materials Today: Proceedings*, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S221478532100078X>
- [7] Gazebo project home page, url: <https://gazebo.org/home>.
- [8] UNITY home page, url: <https://unity.com/download>
- [9] Whitney D., Rosen E., Ullman D., Phillips E., and Tellex S., "Ros reality: A virtual reality framework using consumergrade hardware for ros-enabled robots," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1–9.
- [10] Kerzel M., Pekarek-Rosin T., Strahl E., Heinrich S., and Wermter S., "Teaching nico how to grasp: An empirical study on crossmodal social interaction as a key factor for robots learning from humans," *Frontiers in Neurorobotics*, vol. 14, 2020. url: <https://www.frontiersin.org/article/10.3389/fnbot.2020.00028>
- [11] iCUB project home page, url: <https://icub.iit.it/>
- [12] Wijnen L., Lemaignan S., Bremner P. (2020). *Towards using virtual reality for replicating HRI studies*. HRI'2020 Companion, Cambridge. UK.
- [13] Hudák, M.; Sobota, B.; Korečko Š. ; Sivý M.: "Fully Immersive Web-Based Collaborative Virtual Environment for Upper Limb Rehabilitation Purposes," *2020 18th International Conference on Emerging eLearning Technologies and Applications (ICETA)*, 2020, pp. 194-199, doi: 10.1109/ICETA51985.2020.9379239.
- [14] 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, 343–348, doi:10.1109/CogInfoCom50765.2020.9237899.

Virtuálna škola v prostredí kolaboratívnej virtuálnej reality

¹Branislav SOBOTA, ²Miriama MATTOVÁ, ³Sebastián DIMUN, ⁴Štefan KOREČKO

^{1,2,3,4}Katedra počítačov a Informatiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Letná 9, 04200 Košice, Slovenská republika

¹branislav.sobota@tuke.sk, ²miriama.mattova@tuke.sk

Abstrakt — Príspevok popisuje systém virtuálnej školy, ktorá poskytuje možnosť vzdelávania v prostredí kolaboratívnej virtuálnej reality pre študentov základných škôl so špeciálnou formou pedagogiky. Implementovaný systém pozostáva z webového prostredia pre možnosť prípravy jednotlivých scén prezentujúcich rôzne okruhy výučby a prostredia virtuálnej reality pre samotný proces vzdelávania a možnosť rozvoja kognitívnych funkcií. Systém zároveň umožňuje viacero spôsobov interakcie jednotlivých používateľov, ako aj ich spoluprácu pri plnení rozličných úloh

KLúčové slová — virtuálna škola, kolaboratívna virtuálna realita, online vzdelávanie

I. ÚVOD

Informačné technológie čoraz viac vplývajú na súčasný svet a vzdelávanie nie je výnimkou. Aj nedávna pandemická situácia ukázala nové možnosti hlavne online vyučovania. S rozvojom technológií ale aj vzdelanosti používateľov narastajú aj požiadavky na tento druh vyučovania či už napríklad v interaktívnosti alebo „inteligencii“ komunikácie. Podľa Burgessa a Sievertseny [1] cítiť bezprecedentný prechod na online vzdelávanie v mnohých smeroch. Online forma má v súčasnej podobe pomerne dobré výsledky ak sprostredkováva digitálny obsah v štandardnej multimedialnej forme (audio, statické obrázky a video). Mnohé problematiky, ako napríklad fyzika, chémia, história či biológia však môžu mať požiadavky na interaktívnejšie sprostredkovanie vyučovaného obsahu. Toto je možné dosiahnuť spravidla použitím nových, inovatívnych technológií. Jednou z najdynamickejších rastúcich technológií, ktoré majú ambíciu interaktivity zvýšiť je rozšírená realita (XR) a v rámci nej virtuálna realita (VR) a pridružené technológie [2]. Prienik týchto technológií je v poslednom období viditeľný napríklad v oblasti použitia VR simulátorov pre tréning resp. výuku chirurgov [3], hasičov [4] či stavbárov [5].

Na pracovisku autorov príspevku (Laboratórium LIRKIS, KPI FEI TU Košice) sa vykonávajú práce, ktoré by mohli pomocou technológií virtuálnej reality v kontexte interakcie človeka s výpočtovým systémom [6] uľahčiť, zjednodušiť, zrýchliť a sprehľadniť komunikáciu študentov a žiakov s výpočtovou technikou ako aj so svojím okolím. Tento koncept je zaujímavý aj z hľadiska možných dlhodobých cieľov v kontexte rozvoja online vzdelávania.

Používanie virtuálnych prostredí rozširuje možnosti spolupráce skupiny používateľov v reálnom čase. Zdieľané (kolaboratívne) virtuálne prostredia sú otvorené možnostiam rôznych simulácií, tréningov alebo edukácie a iných typov aktivít. Kolaboratívne systémy z pohľadu Flaviána [7] predstavujú špecifickú architektúru a technologické vybavenie pre možnosti vytvárania kolaboratívnych aktivít. Autori [8] uvádzajú potrebu fyzickej prítomnosti používateľov alebo ich virtuálnej podoby vo forme avatarov v kolaboratívnych prostrediach. Najčastejšie kolaboratívne prostredia na tejto úrovni sú systémy založené na virtuálnych jaskyniach a systémy založené na virtuálno-reality/dátových prilbách ako formy kolaboratívnych prostredí dostupných pre širokú skupinu používateľov. Tieto systémy sú aj cenovo dostupné pre nasadenie v školách. XR (eXtended Reality) technológie (aj VR) interaktívne a vizuálne kvalitnejšie dokážu spracovať učivo v rámci výchovno-vzdelávacieho procesu vrátane novej kolaborácie žiakov. Tým sa predpokladá rýchlejšie, dokonalejšie a ľahšie pochopenie učiva. Takýmto nasadením sa zaoberajú aj autori príspevku vo svojom laboratóriu [9] [10] pričom napríklad v nadväznosti na [11] sa snažia dodržiavať základné zásady pre vytváranie vzdelávacích kurzov. Vo všeobecnosti by sa mali dodržiavať nasledovné zásady:

- presne stanovená štruktúra učiva, aby došlo k naplneniu cieľa učiva
- používanie jednoduchého a zrozumiteľného jazyka, pričom jazyk by mal zodpovedať úrovni preberaného učiva a v žiadnom prípade by nemal byť zložitý
- správne spracovanie kurzu po technickej stránke

To bolo niekoľko zamyslení ohľadom potrieb žiaka avšak ďalším dôležitým aspektom je nastavenie systému z pohľadu učiteľa. Je potrebné, aby systém predstavoval pre učiteľa pomôcku pri vyučovaní. Z toho dôvodu by nemal byť príliš zložitý na použitie v triede. Čas vyučovacej hodiny je pre učiteľa obmedzený a v prípade, ak by použitie systému vyžadovalo príliš dlhý čas, systém by bol využitelný iba v špeciálnych prípadoch vyučovania.

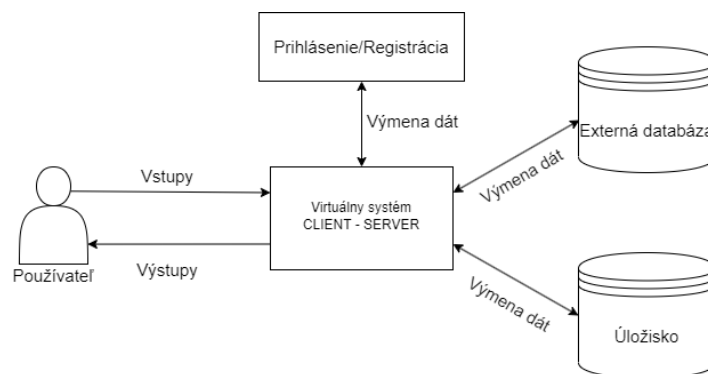
V nasledujúcom texte bude uvedený krátky opis systému vyvíjaného na pracovisku autorov príspevku z hľadiska použitia práve XR/VR technológií vo vyučovacom procese.

II. NÁVRH SYSTÉMU

Prostredie virtuálnej kolaboratívnej reality poskytuje možnosť dištančnej výučby študentov. V porovnaní s klasickou formou výučby v školách by mala táto forma študentov nadchnúť a pomôcť im pri zdokonaľovaní svojich vedomostí vďaka vizuálnemu vnemu, ktoré môže pozitívne ovplyvniť záujem učiť sa. Pomocou analýz jednotlivých možností tvorby virtuálneho prostredia bude vytvorený systém, ktorý umožňuje zábavnou a zaujímavou formou spoznávať a študovať jednotlivé okruhy vzdelávania. Fyzická prítomnosť žiakov v školách nebude potrebná. Pri splnení určitých podmienok bude možné podieľať sa na výučbe aj z pohodlia domova. Navrhovaný systém pozostáva z dvoch základných častí, ktorými sú:

- Webové prostredie
- Prostredie virtuálnej reality

Chod celého systému je založený na nakonfigurovaní počiatočných nastavení vo webovom prostredí a následnej forme výučby zvoleného okruhu v zdieľanom virtuálnom prostredí. Pre systém virtuálnej zdieľanej školy bol navrhnutý konceptuálny model (Obr. 1), ktorý znázorňuje komunikačný proces medzi jednotlivými časťami systému.



Obr. 1 Konceptuálny model

Používateľom v tomto prípade je učiteľ alebo študent. Pomocou vstupov môže používateľ interagovať s virtuálnym systémom, či už sa jedná o webové prostredie, alebo virtuálne zdieľané prostredie. Výstupy zabezpečujú spätnú väzbu pre používateľa vo forme obrazu a sú závislé od jeho vstupov. Prihlásenie a registrácia používateľa je počiatočným krokom pred vstupom do webového prostredia. Každý študent a učiteľ musí byť registrovaný. Jedinečné prihlasovacie meno s heslom zabezpečujú používateľovi prístup do menu virtuálneho systému.

Navrhovaný virtuálny systém pozostávajúci z webového prostredia a prostredia virtuálnej reality potom zabezpečuje

- chod celého systému,
- proces vytvárania a prihlásenia používateľa,
- komunikáciu s externými komponentami, ktorými sú úložisko a externá databáza.

Základnou časťou webového prostredia je menu virtuálneho systému, ktoré obsahuje rôzne funkcie, nie len na nakonfigurovanie prostredia virtuálnej reality, ale aj klasické možnosti, akými sú návrat domov, zmena hesla, či odhlásenie zo systému.

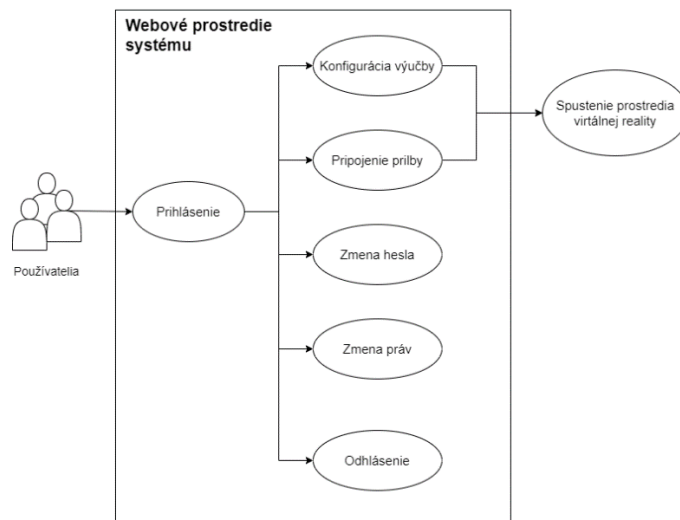
Podstatnou súčasťou je aj podpora pre paralelné pripojenie VR prilby so systémom. Proces pripojenia spočíva vo vytvorení objektu prilby v databáze a následné pripojenie konkrétnej prilby (použitá Oculus Quest II). Spustenie virtuálneho zdieľaného prostredia je možné dvomi spôsobmi:

- v datovej prilbe (HMD VR), ktorú používateľ musí mať spárovanú, alebo
- v osobnom počítači (Desktop VR)) alebo
- na mobilných zariadeniach na báze Android a iOS nakoľko virtuálny systém možno spustiť pomocou webového prehliadača
- prípadne v všetkých zariadeniach súčasne.

Na úložisku sa nachádzajú jednotlivé modely, obrázky, logá a iné zdrojové súbory, je podporované v platforme Glitch, ktoré je použité pre tento kolaboratívny virtuálny systém. Virtuálny systém potrebuje externú databázu pre ukladanie dát. Z viacerých databáz bola práve pre tento systém zvolená cloudová real-time databáza, ktorú poskytuje platforma Firebase. Firebase real-time databáza umožňuje ukladanie, synchronizáciu a poskytovanie dát pre systém v reálnom čase. Je navrhnutá tak, aby bola jednoduchá na použitie a zároveň umožňovala vývojárom vytvárať aplikácie s rýchlym a efektívnym načítaním dát.

III. WEBOVÉ PROSTREDIE

Webové prostredie je dôležitou súčasťou celého systému. Skladá sa z viacerých častí, ktoré poskytujú rôzne možnosti použitia systému (Obr. 2).



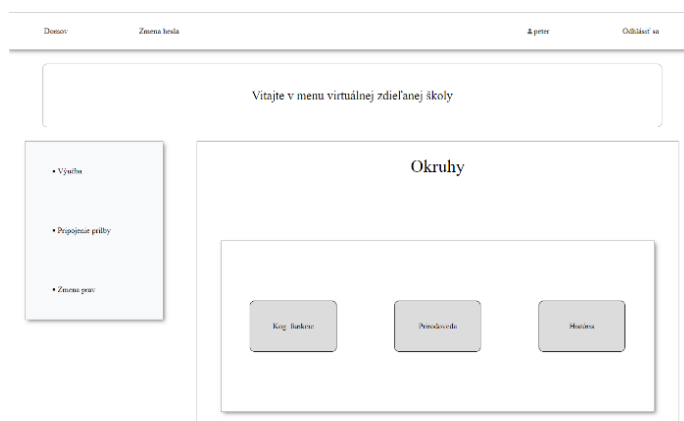
Obr. 2 Webové prostredie systému

Ako bolo uvedené, pred vstupom do systému je potrebné prihlásiť sa. Systém umožňuje jednoduchý proces prihlásenia pomocou používateľského mena a hesla. V prípade ak používateľ nebol ešte registrovaný, môže sa jednoducho zaregistrovať pred prihlásením do systému (Obr. 3).

Obr. 3 Úvodná webová stránka

Okrem registrácie a prihlásenia sa do systému je možné pomocou webového prostredia paralelne pripojiť VR datovú prilbu. Vďaka tomu môže používateľ spustiť systém na počítači, laptope, či smartfóne, nakonfigurovať virtuálne prostredie a následne ho spustiť v paralelne pripojenej VR datovej prilbe.

Po prihlásení sa používateľ nachádza v menu webového prostredia. V ňom si môže zvoliť okruh výučby a jednu z jeho úrovní (Obr. 4). Následne je možné prejsť do prostredia kolaboratívnej virtuálnej reality, kde je umožnený vzdelávací proces používateľom na základe konfiguračných nastavení.



Obr. 4 Webové prostredie s okruhmi výučby

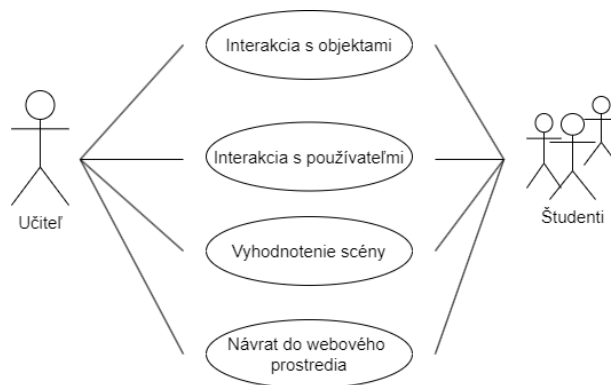
V prípade záujmu o spustenie procesu výučby v paralelne pripojenej VR datovej prilbe je potrebné pripojiť vytvorený objekt prilby ku konkrétnemu používateľovi. Na to slúži sekcia Pripojenie prilby v ovládacom paneli na ľavej časti.

System na základe práv odlišuje 2 typy používateľov - Učiteľ a Študent. Tie je možné meniť vo webovom prostredí v sekcii Zmena práv.

Súčasťou sú aj štandardné webové prvky ako tlačidlo domov, možnosť zmeny hesla a odhlásenie zo systému.

IV. PROSTREDIE VIRTUÁLNEJ REALITY

Pre tvorbu virtuálnej reality v tomto systéme bol použitý webový rámec A-frame. Systém umožňuje viacero spôsobov interakcie so zdieľanými objektami v scénach, interakciu s používateľmi a vyhodnotenie jednotlivých scén. Use-case diagram práce v prostredí virtuálnej reality je zobrazený na Obr. 5.



Obr. 5 Use-case diagram práce v prostredí virtuálnej reality

Keďže prechod do virtuálneho prostredia je implementovaný formou prechodu do nového webového okna s konkrétnou virtuálnou scénou, návrat do webového menu je možný jednoduchým zavretím tohto okna. To platí pre prípad ak používateľ používa práve 1 zariadenie. V prípade ak je scéna virtuálnej reality spustená v paralelne pripojenej VR datovej prilbe, menu webového prostredia používateľ neopúšťa so zariadením, s ktorým doňho vstúpil. Zmena scén môže byť vykonaná nezávisle od pripojenej prilby a po zmene okruhu a úrovne je spustená správna scéna v prilbe.



Obr. 6 Ukážka scény v prostredí VR

Vo virtuálnej scéne počas plnenia úloh využívajú používatelia rôzne interakcie so zdieľanými objektami. Môžu ich zodvihnúť, položiť na zem, uložiť na správne/nesprávne miesto alebo označiť/zrušiť označenie. To všetko je možné viacerým používateľom v rovnakom čase v rámci jednej scény. Ukážka scény v prostredí VR je na Obr. 6.

Systém obsahuje implementáciu fyzikálneho podsystemu, ktorý zabezpečuje základne fyzikálne vlastnosti, ktoré platia aj v reálnom svete. V tomto prípade sa jedná o gravitáciu a kolízie medzi jednotlivými objektami/používateľmi.

V. ZÁVER

Cieľom práce bolo vytvoriť systém, ktorý poskytne vzdelávací proces žiakom vo virtuálnom prostredí. V aktuálnom stave dokáže tento systém simulovať virtuálnu realitu, v ktorej jeho používatelia môžu aj spolupracovať v rovnakom čase a rovnakom zdieľanom virtuálnom prostredí pri riešení jednotlivých úloh. To môže pozitívne prispieť k rozvoju kognitívnych schopností, motoriky a úrovni vzdelania, ale aj k vzájomnej empatii používateľov. Ďalšími prínosmi vytvoreného systému virtuálnej zdieľanej školy je aj poskytovanie dištančnej formy výučby z pohodlia domova, využívanie informačných technológií pre proces vzdelávania, ale aj možnosť vzdelávania sa vo virtuálnom prostredí, ktoré v skutočnosti nemusí byť žiakovi dostupné a môže ho spoznať vďaka virtuálnej realite.

POĎAKOVANIE

Tento príspevok vznikol s podporou projektu KEGA 048TUKE-4/2022 „Technológie zdieľanej virtuálnej reality vo vzdelávacom procese“.

REFERENCIE

- [1] Burgess, S.; Sievertsen, H.H.: Schools, Skills, and Learning: The Impact of COVID-19 on Education. In: VoxEU.org (2020), dostupné na: <https://voxeu.org/article/impact-covid-19-education>
- [2] Sobota, B.; Korečko, Š.; Hrozek, F.: On building an object-oriented parallel virtual reality system; In: Central European Journal of Computer Science. Vol. 2, no. 3 (2012), p. 261-271. - ISSN 2081-9935
- [3] Mishra, K.; Boland, M.V.; Woreta, F.A. Incorporating a virtual curriculum into ophthalmology education in the coronavirus disease-2019 era. In: Curr. Opin. Ophthalmol. 2020, 31, 380-385.
- [4] Wijkmark, C.H.; Fankvist, S.; Heldal, I.; Metallinou, M.M. Remote Virtual Simulation for Incident Commanders: Opportunities and Possibilities. In: Proceedings of the 2020 11th IEEE International Conference on Cognitive Infocommunications (2020); pp. 445-452.
- [5] Luimula, M.; Linder, M.; Pieskä, S.; Laimio, E.; Lähde, T.; Porramo, P. Unlimited Safety Productivity—A Finnish Perspective Using Virtual Learning Methods to Improve Quality and Productivity in the Construction Industry. In: Proceedings of the 2020 11th IEEE International Conference on Cognitive Infocommunications (2020); pp. 263-266.
- [6] Sinha, G.; Shahi, R.; Shankar, M.: Human Computer Interaction. In: Emerging Trends in Engineering and Technology (ICETET), 2010 3rd International Conference on. IEEE, 2010. p. 1-4
- [7] Flavián, C.; Ibáñez-Sánchez, S., & Orús, C. (2018). The impact of virtual, augmented and mixed reality technologies on the customer experience. Journal of Business Research.
- [8] Gonzalez-Franco, M., Hall, M., Hansen, D., Jones, K., Hannah, P., & Bermell-Garcia, P. (2015, March). Framework for remote collaborative interaction in virtual environments based on proximity. In: 3D User Interfaces (3DUI), 2015 IEEE Symposium on (pp. 153-154). IEEE.
- [9] 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.
- [10] Hudák, M., Korečko, Š. and Sobota, B. : Enhancing Team Interaction and Cross-platform Access in Web-based Collaborative Virtual Environments, In: Proceedings of 2019 IEEE 15th International Scientific Conference on Informatics, IEEE, 2019, pp. 160-16.
- [11] Glozar, Jiri; Kastnerova, Lucie; Necas, Ondrej; Ondra, Svatoslav; Penaz, Petr. Prístupnosť e-learningu pro studenty s postizením. Masarykova univerzita, 2007. Stredisko pro pomoc studentum se specifickými naroky.

Rozpoznávanie hoaxov modelovaním slovníka ich autorov

¹Kristína MACHOVÁ, ²Miroslav PLAVEC

^{1,2} Katedra Kybernetiky and Umelej Inteligencie, Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská Republika

¹kristina.machova@tuke.sk, ²miroslav.plavec@student.tuke.sk

Abstrakt — Článok sa zameriava na rozpoznávanie hoaxov v online komunikácii o COVID-19. Koronakríza vyvolaná ochorením COVID-19, ktorá sa týkala celého sveta, bola a je zároveň aj dezinformačnou krízou. Je viac ako zrejmé, že manipulácia s informáciami týkajúcimi sa pandémie sa stala pre ľudí veľkým problémom. Článok popisuje oblasť rozpoznávania hoaxov a onúka návrh funkčnej aplikácie na rozpoznanie nejakého textu, z hľadiska toho či je alebo nie je hoax. Pomocou algoritmov strojového učenia boli natréňované a vyhodnotené viaceré modely a najlepší z nich bol použitý v grafickom používateľskom rozhraní. Každý z modelov bol ohodnotený pomocou viacerých mier efektívnosti.

Kľúčové slová — detekcia hoaxov, spracovanie textových dokumentov, dolovanie webu, strojové učenie, Naive Bayes, Random Forest, SVM – Support Vector Machine

I. ÚVOD

Koncom roka 2019 sa v čínskom meste Wu-chan prvý raz identifikovalo infekčné ochorenie COVID-19. Odtiaľ sa rozšírilo do celého sveta a spôsobilo celosvetovú pandémiu, ktorá sa dostala aj na Slovensko začiatkom roka 2020. Pandémia koronavírusu spôsobila najväčšiu globálnu spoločenskú aj ekonomickú recesiu od čias veľkej hospodárskej krízy. Postupne sa zavádzali opatrenia, ktoré rušili kultúrne, športové a náboženské udalosti a prinútili školy prejsť na distančné vzdelávanie. Ľudia začali viac tráviť svoj voľný čas na internete a sociálnych sieťach, kde na nich čakalo obrovské množstvo dezinformácií a hoaxov o ochorení, pandémie alebo očkovaní. V dnešnej dobe môže prakticky ktokoľvek publikovať rôzne mediálne výstupy v online podobe. Dokonca aj osoba, ktorá nemá žiadne znalosti o určitej problematike, sa k nej môže vyjadrovať a dávať ľuďom rady, ako ju riešiť. Tým pádom sa do online priestoru dostanú nepravdivé informácie, ktoré sa veľmi rýchlo šíria a mnohí ľudia im veria aj keď ide o hoaxy. Šírenie hoaxov a manipulácia s informáciami týkajúcimi sa pandémie sa stali pre spoločnosť veľkou hrozbou. Dostali sme sa do doby, kedy používatelia webu nevedia, ktorým informáciám môžeme veriť. Aj preto je cieľom práce prezentovanej v tomto článku navrhnúť prístup ku detekcii hoaxov a vytvoriť funkčný program s grafickým rozhraním, ktorý bude na základe vopred natréňovaných modelov detegovať hoaxy v zadanom texte. Modely boli natréňované pomocou troch algoritmov strojového učenia. Taktiež bol skúmaný vplyv dĺžky n-gramov a spôsoby ohodnotenia textu na výsledky klasifikácie.

II. DEFINÍCIA HOAXU

Slovo hoax má svoj pôvod v anglickom jazyku. Môžeme ho preložiť ako podvod, fáma, falošná správa, podfuk alebo švindel. Podľa druhého vydania Oxfordského anglického slovníka sa toto slovo prvýkrát objavilo v roku 1808. Britský jazykový vedec Robert Nares však tvrdil, že korene tohto slova môžeme vystopovať asi o dvesto rokov skôr, a to ku fráze hocus-pocus. Túto frázu väčšinou využívajú kúzelníci pri svojich predstaveniach, kde dochádza k podvodu divákov. Aj keď sa nám slovo hoax môže zdať ako moderný pojem, tak nie je ničím novým. Takýto druh informácie existoval už v minulosti. Vtedy sa namiesto pomenovania hoax používali pomenovania ako povera, klamstvo alebo mestská legenda. V dávnejších dobách sa ale hoaxy mohli šíriť len obmedzene, napríklad ústne alebo obrázkami. Práve vývoj internetu a sociálnych sietí im poskytol potenciál šíriť sa naplno. Na jednej strane nám internet prináša spôsob ako si rozširovať naše znalosti, no na druhej strane nadmerné trávenie času na internete a sociálnych sieťach škodí nášmu zdraviu a mysli. Rozpoznať pravdivú informáciu od tej nepravdivej je v dnešnej dobe čoraz zložitejšie. Najväčší problém s takýmto rozpoznávaním majú starší ľudia.

Podľa štúdie vedcov z Princetonu a New York Univerzity ľudia starší ako 65 rokov zdieľajú takmer 7-krát viac hoaxov ako dospelí do 30 rokov [1].

V odbornej literatúre alebo na internete dokážeme nájsť veľa rôznych definícií hoaxu, všetky však majú určité charakteristiky rovnaké. V počítačovom svete hoax najčastejšie predstavuje poplašnú správu, ktorá varuje pred neexistujúcim nebezpečenstvom. Môže to byť aj správa, ktorá obsahuje nepresné, skresľujúce informácie, účelovo upravené polopravdy alebo zmes polopráv a klamstiev. Na to, aby bol hoax úspešný, potrebuje dosiahnuť zdieľanie s čo najväčším množstvom čitateľov a na to používa lákavé alebo emotívne témy. Niektoré hoaxy sa zameriavajú na zdravotnú stránku, pracujú so strachom z nákazy a ponúkajú tip, ako sa jej zbaviť. Iné zase pracujú so strachom z nedostatku potravín. Pri pandémie sme sa mohli presvedčiť, že hoaxy sú natoľko nebezpečné, že na nich mnohí ľudia doplatili svojim životom. Ľudia, ktorí nemali dostatočný prístup k relevantným informáciám a preberali správy z nedôveryhodných médií uverili antivakcinačnej propagande. Tak sa stalo, že nielen nemocnice, ale aj cintoríny začali plniť samotní antirúškarmi a antivaxermi.

Základné znaky hoaxu sú:

- žiadosť o zdieľanie
- emotívna téma a titulok
- predstieranie varovania od dôveryhodného zdroja
- nepresné údaje.

Aké možnosti máme rozlišovať hoax a bojovať s ním? V prvom rade by sme si mali klásť nasledovné otázky:

- KTO: Kto je autorom danej informácie alebo kto je zodpovedný za jej šírenie?
- ČO: Čo je obsahom danej správy?
- KOMU: Komu je správa určená?
- AKO: Ako sa správa snaží upútať pozornosť čitateľa?
- PREČO: Prečo správa vznikla alebo prečo sa šíri?

Skutočná reálna správa je charakteristická určitými znakmi. Autor je v správe uvedený, text býva vecný a zrozumiteľný a zdroje sú uvedené. Pričom hoaxy majú autora väčšinou neznámeho, poživajú údaje, ktoré sú nepresné a zdroje nie sú uvedené [2]. Takéto zodpovedanie si vyššie uvedených otázok alebo overenie informácie nie je v dnešnej dobe komplikované. Preto môže byť nástroj na automatickú detekciu hoaxov, založený na strojovom učení veľmi nápomocný.

III. PRÍSTUP K BUDOVANIU DETEKČNÝCH MODELOV

A. Analýza Stavů Problematiky

Práca [3] bola vytvorená na univerzite Zagreb a autori túto prácu nazvali ako „Inteligentný automatický systém pre detekciu hoaxov“. Pomocou kontrolovaného a nekontrolovaného učenia sa snažili rozpoznať hoaxové správy v e-mailoch. Kontrolované učenie použili na rozlíšenie hoaxov od bežných e-mailov a nekontrolované učenie použili pre klasifikáciu hoaxov do rôznych skupín. Celý ich detekčný systém pozostáva z viacerých moduloch, ktoré sú medzi sebou prepojené. Prvým krokom autorov bolo predspracovanie textu, ktoré im pomohlo znížiť počet vstupných neurónov v obidvoch sieťach, ktoré použili. Keďže e-mail, s ktorými pracovali, boli písané vo viacerých jazykoch, nemohli pri predspracovaní použiť metódy ako stemming alebo lematizácia, teda upravenie slova na jeho základ. Namiesto týchto metód sa rozhodli použiť n-gramy o veľkosti $n=3, \dots, 8$. Pre každý dôležitý n-gram vypočítali hodnotu TF-IDF. Vektor takýchto hodnôt potom predstavoval vstup do neurónovej siete. Na rozlišovanie medzi hoaxom a obvyčajným e-mailom použili doprednú neurónovú sieť, ktorá pozostávala z 1369 vstupných neurónov, zo 100 skrytých neurónov a z 2 výstupných neurónov. Po úspešnej detekcii hoaxu autori tejto práce použili metódu SOM (samorganizujúca sa sieť), aby dokázali zaradiť do akej skupiny daný hoax patrí. Následne natrénovaná neurónová sieť dokázala správne určiť hoax s presnosťou 74%.

Dalšia práca z univerzity Victoria v Kanade bola zameraná na detekciu online fake news pomocou analýzy n-gramov a techník strojového učenia. V tejto práci bol prezentovaný klasifikátor založený na slovných n-gramoch, ktoré reprezentujú kontext dokumentov, a taktiež bol skúmaný vplyv dĺžky n-gramov na výsledok detekcie. Pracovalo sa s datasetom pravdivých a nepravdivých správ, ktoré boli dostupné na internetových stránkach (12 600 pravdivých a 12 600 nepravdivých správ). Po predspracovaní a vytvorení n-gramov bola pre každý n-gram vypočítaná hodnota TF a TF-IDF. Pri trénovaní klasifikátorov bolo použitých šesť rôznych algoritmov strojového učenia a to: Stochastic Gradient Descent (SGD), metóda podporných vektorov (SVM), lineárna metóda podporných vektorov (LSVM), k-najbližší susedia (KNN) a

rozhodovací strom (DT). Pri vyhodnocovaní sa ukázalo, že lineárne klasifikátory dosahovali lepšie výsledky ako tie nelineárne. Najvyššiu presnosť bola dosiahnutá pomocou unigramu a LSVM a to 92%. Ukázalo sa, že predlžovaním n-gramov sa presnosť znižovala až na 47,2% [4].

Článok [5] prezentuje testovanie 8 rôznych algoritmov strojového učenia a to: Naivný Bayes, Adaboost, k-NN, náhodný les, logistická regresia, rozhodovací strom, viacvrstvový perceptrón a SVM a taktiež štyri algoritmy hlbokého učenia a to: konvulučná neurónová sieť (CNN), rekurentná neurónová sieť (RNN), long short-term memory (LSTM) a gated recurrent unit (GRU). Modely strojového učenia boli tréňované na fake news datasete, ktorý obsahoval fake news o COVID-19. Pri porovnávaní jednotlivých algoritmov zistili, že najlepšie výsledky mali modely konvolučnej neurónovej siete a BiLSTM. Tieto algoritmy mali 97% presnosť. Avšak ostatné algoritmy mali tiež vysokú presnosť, napríklad Naivný Bayes mal 95% presnosť. Aj v práci [6] boli testované metódy strojového učenia tréňované na datasete zavádzajúcich správ a najlepšie výsledky boli dosiahnuté použitím neurónovej siete a word embeddingu – 99,31%.

B. Použité Metódy

Na základe analýzy súčasného stavu boli zvolené nasledovné algoritmy: Naivný Bayes, metóda podporných vektorov (SVM) a náhodný les (RF).

V porovnaní s inými metódami strojového učenia, je naivná bayesová klasifikácia jednoduchá metóda na implementáciu, pričom dosahuje relatívne nízku chybovosť. Pri vytváraní modelu, teda pri tréňovaní nepotrebuje veľké množstvo dát. Klasifikátor vychádza z predpokladu, že efekt, ktorý ma hodnota každého atribútu na danú triedu, nie je ovplyvnený hodnotami ostatných atribútov. Čiže požaduje aby predikátory boli medzi sebou nezávislé. Kvôli tomuto zjednodušeniu sa tento klasifikátor nazýva aj ako „naivný“. Naivný bayesov klasifikátor našiel svoje uplatnenie pri analýze sentimentu, filtrovanie spamu, klasifikácii textu alebo doručovacích systémov. Medzi používané naivné bayesové klasifikátory patrí napríklad Gaussovský Naivný Bayes, Multinomický Naivný Bayes alebo Bernoulliho Naivný Bayes [7].

Metóda podporných vektorov (SVM) klasifikuje dáta spravidla do dvoch tried. V prípade, že sú triedy lineárne separovateľné, hovoríme o lineárnej SVM, v prípade, že nie sú, hovoríme o nelineárnej SVM alebo kernelovej SVM. Na správne fungovanie algoritmu potrebujeme labelovanú tréňovaciu množinu. Cieľom SVM algoritmu je nájsť nadrovinu (ang. Hyperplane) v N-rozmernom priestore, kde N je počet príznakov, ktorá rozdeľuje jednotlivé triedy a určuje ktoré body patria do ktorej triedy. Nadrovina môže byť tiež označená ako rozhodovacia hranica (ang. Decision boundary). Vo všeobecnosti na oddelenie tried môže existovať nekonečne veľa oddeľovacích nadrovin, avšak pri metóde SVM hľadáme takú hranicu, ktorá vymedzuje po celej dĺžke čo najširšiu vzdialenosť medzi hraničnými bodmi jednotlivých tried. Pri tréňovaní SVM, body, ktoré sa nachádzajú najbližšie k nadrovine sa nazývajú podporné vektory (ang. Support vectors) a pomocou týchto bodov maximalizujeme šírku hranice, tieto body nám pomáhajú budovať SVM. Šírka hranice reprezentuje kritériálnu funkciu [8].

Náhodný les (angl. Random forest) je metóda založená na kombinácii viacerých dekorelovaných rozhodovacích stromov, kde každý strom poskytne unikátny pohľad na riešenie problematiky. Každý strom závisí na hodnotách náhodného vektora hodnôt vzorkovaného nezávisle a s rovnakým rozložením pre všetky stromy v tzv. lese stromov. Používa sa ak jeden rozhodovací strom nie je dostatočne presný. Ide teda o metódu, ktorá vytvára viacero rozhodovacích stromov, pomocou ktorých získavame presnejší výsledok [9].

IV. NÁVRH RIEŠENIA

Prvým a taktiež veľmi dôležitým krokom pri riešení hocijakého problému pomocou strojového učenia je zber dát s príkladmi hoaxov a pravdivých správ o CoViD-19. Pracovali sme so správami v slovenskom jazyku a dataset sme sami vytvorili. V našom datasete je 100 hoaxov a 100 pravdivých správ, ktoré sa týkali situácie okolo koronakrízy. Hoaxy boli zozbierané z facebookovskej stránky „Hoaxy a podvody – Polícia SR“. Pravdivé informácie boli zozbierané z oficiálnych stránok zdravotníckych organizácií a z oficiálnej stránky ministerstva zdravotníctva. Dáta boli následne predspracované nasledovne:

- odstránenie interpunkcie a symbolov
- prevod na malé písmená
- eliminácia neplnovýznamových slov
- lematizácia
- vytvorenie vektorovej reprezentácie pomocou TF-IDF.

Na týchto dátach boli tréňované modely použitím Naivnej Bayesovej metódy (NB), metódy podporných vektorov (SVM) a metódy generovania náhodného lesa (RF). Výsledky testovania týchto modelov sú uvedené v Tabuľke 1.

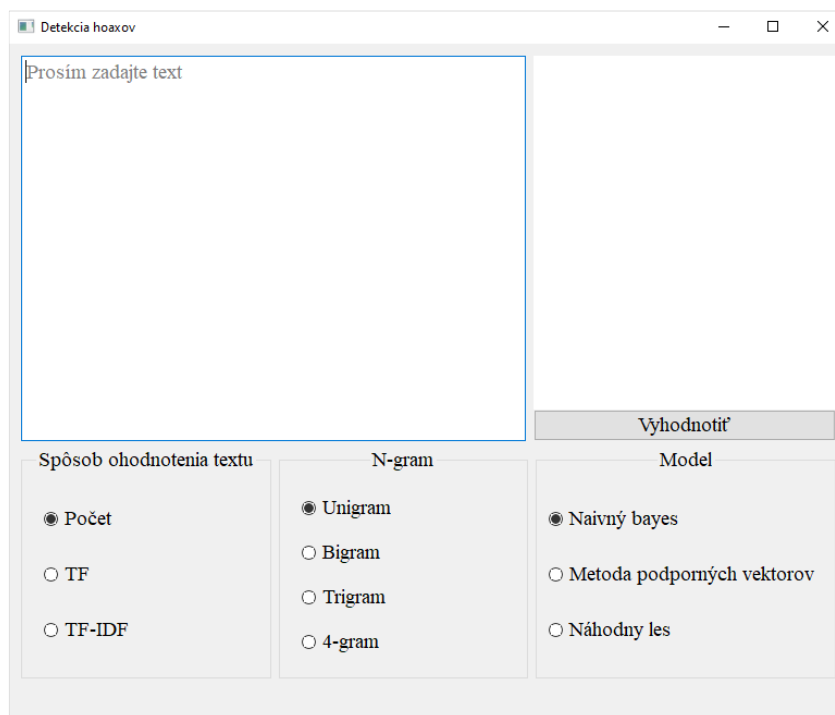
Tabuľka 1

Výsledky testov modelov generovaných metódami strojového učenia NB, SVM a RF nad dátami s použitím TF-IDF reprezentácie pomocou mier efektívnosti: Správnosti (Accuracy), Presnosti, Návratnosti a F1-skóre.

	NB	SVM	RF
Správnosť	0.972	0.975	0.925
Presnosť	0.951	0.952	0.800
Návratnosť	0.988	1.000	1.000
F1 Rate	0.974	0.975	0.888

Výsledky testovania ukázali, že najvhodnejšia metóda na generovanie detekčného modelu pre rozpoznávanie hoaxov je SVM. Všetky tri použité modely boli najlepšie v návratnosti, teda vedeli nájsť skoro všetky problematické hoaxové správy v testovacej množine.

Bola vytvorená aj aplikácia, ktorá umožňuje detegovať prípadný hoax v skúmanom texte, ktorý je potrebné zadať do okna aplikácie. Po zvolení možnosti spracovania sa ukáže výsledok analýzy. Táto aplikácia je ilustrovaná na Obrázku 1.



Obr. 1 Grafické rozhranie aplikácie na detekciu hoaxu v zadanom texte.

V. ZÁVER

Článok prezentuje návrh prístupu k budovaniu detekčných modelov pre rozpoznávanie hoaxov, a taktiež aplikáciu, v ktorej si používateľ môže overiť nejakú podozrivú správu, či nie je hoaxom. Najlepšie výsledky boli dosiahnuté metódou SVM použitím TF-IDF. Vyhodnotenie však v našom prípade závisí aj od charakteru vkladanych textov. Môže sa stať, že správa bude klasifikovaná nesprávne vzhľadom na slová, ktoré sa v texte nachádzajú. Aplikácia môže vyhodnotiť pravdivú správu ako hoax, ak sa v správe nachádzajú slová, ktoré boli vo veľkej miere využívané na hoaxových portáloch. Výhodou nášho riešenia je, že detekcia pri správnom predspracovaní textu môže byť použitá pre akýkoľvek jazyk. V našom prípade sme boli pri experimentoch obmedzení slabou dostupnosťou textov hoaxov v slovenskom jazyku.

POĎAKOVANIE

Táto práca vznikla s podporou Slovenskej grantovej agentúry Ministerstva školstva a SAV Slovenskej republiky v rámci projektu VEGA 1/0685/21.

LITERATÚRA

- [1] GUESS, Andrew – NAGLER, Jonathan – TUCKER, Joshua: Less than you think: Prevalence and predictors of fake news dissemination on Facebook, Science Advances, 2019, vol.5, ISSN: 2375-2548. [cit. 2023-05-15]. Dostupné na internete: <<https://www.science.org/doi/epdf/10.1126/sciadv.aau4586>>.

- [2] VAJNORÁKOVÁ, Veronika: Hoax a jeho jazykovo-štylistická analýza. Nitra: Katedra slovenského jazyka a literatúry FF UKF, [cit. 2023-05-15]. s. 15-22. Dostupné na internete: <<https://opac.czsp.sk>>.
- [3] VUKOVIC, Marin – PRIPUŽIC, Krešimir – BELANI, Hrvoje. An Intelligent Automatic Hoax Detection System. [cit. 2023-05-15]. Dostupné na internete: <https://www.fer.unizg.hr/_download/repository/kes09.pdf>.
- [4] AHMED, Hadeer – TRAORE, Issa – SAAD, Sherif. 2017. Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques, Springer, vol.10618, 2017, ISBN: 978-3-319-69155-8.
- [5] HAIDER BANGYAL, Waqas – QASIM, RUKHMA – REHMAN, Najeeb – AHMAD, Zeeshan – DAR, Hafsa – RUKHSAR, Laiqa – AMAN, Zahra – AHMAD, Jamil. 2021. Detectio of Fake News Text Classification on COVID-19 Using Deep Learning Approaches. Computational and Mathematical Methods in Medicine. [cit. 2023-05-15]. Dostupné na internete: <<https://downloads.hindawi.com/journals/cmmm/2021/5514220.pdf>>.
- [6] ELHADAD, Mohamed – FUN LI, Kin – FAYEZ GEBALI. Detecting Misleading information on COVID-19, IEEEAccess. vol.8., 2020, ISBN: 165201-165215.
- [7] HAN, Jiawei – KAMBER, Micheline – PEI, Jian. Data Mining Concepts and Techniques. Third Edition. ISBN: 978-0-12-381479-1. [cit. 2023-05-15]. Dostupné na internete: <<http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>>.
- [8] GANDHI, Rohith. 2018. Support Vector Machine — Introduction to Machine Learning Algorithms, towards data science. [cit. 2023-05-15]. Dostupné na internete: <<https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>>.
- [9] DÚBRAVA, Juraj. Využití strojového učení pro predikci vlivu mutací na stabilitu proteinů. Brno: Fakulta Informačních technologií, Vysoké Učení Technické v Brne, 2018, [cit. 2023-05-15]. Dostupné na internete: <https://www.vut.cz/www_base/zav_prace_soubor_verejne.php?file_id=181194>.

Metódy dátovej analytiky pre predikciu platov v oblasti športu

¹Anna Biceková, ²Dávid Ujhelyi

^{1,2} Katedra kybernetiky a umelej inteligencie, FEI, Technická univerzita v Košiciach,
Letná 9, 042 00 Košice, Slovenská republika

¹anna.bicekova@tuke.sk, ²david.ujhelyi@student.tuke.sk

Abstrakt — Cieľom predkladaného príspevku bolo na základe dát získaných z hokejových zápasov nastaviť pravdepodobný plat hokejových hráčov. V platovom ohodnotení hráčov častokrát dochádza k rozdielom, kedy niektorí hráči sú nedostatočne platovo ohodnotení, alebo naopak majú príliš vysoký plat. Analýza dát a ich predikcia v oblasti športu, ako aj iných sférach dokáže priniesť tímom konkurenčnú výhodu. Výsledky plynúce z analýz dokážu poskytnúť cenné poznatky pre rozhodovanie, optimalizáciu výkonu a zlepšenie strategického plánovania. Súčasťou predkladaného článku bola aj analýza prác venujúcich sa danej problematike, kde autori venujú pozornosť predikcii platov hráčov NHL a ich vzťahu k výkonom na ľade. Naším cieľom bolo na základe štatistík predpovedať plat útočníkov, obrancov a vyhodnotiť finančnú štruktúru konkrétneho hokejového tímu. V analyzovaných štúdiách sa využívala väčšinou len regresná analýza. Naša pozornosť sa zameriavala práve na využitie viacerých predikčných modelov, ako je Lineárna regresia, K-najbližších susedov, Rozhodovací strom, LightGBM, Random forest, Naive Bayes, na základe ktorých sme vybrali najlepší predikčný model a predikovali plat čo najpresnejšie.

Kľúčové slová — dátová analytika, CRISP-DM, predikcia, NHL.

I. ÚVOD

Dátová analytika a predikcia sú v oblasti hokeja čoraz dôležitejšie. Tímy analyzujú výkonnostné metriky, štatistiky hráčov a ďalšie relevantné údaje, aby identifikovali potenciálny talent, zhodnotili vývoj hráča a stanovili presnejšie predpovede o budúcom výkone hráčov. Dátovú analýzu možno použiť na vývoj prediktívnych modelov, ktoré predpovedajú herné výsledky, výkon hráčov resp. platové podmienky. Využitím historických údajov a začlenením rôznych faktorov, ako je sila tímu, výkon hráča, zranenia a ďalšie kontextové informácie, môžu prediktívne modely odhadnúť pravdepodobnosť výhry tímu, gólové rozdiely, platobné podmienky a prínos jednotlivých hráčov. V príspevku sa zameriavame na predikciu platov útočníkov a obrancov. Tieto dve pozície sú dôležitými piliermi tímu a majú v zápasoch odlišné úlohy a zodpovednosti. Pomocou údajov o výkonnosti jednotlivých hráčov a platov zo sezóny 2016/2017 budeme predikovať budúci vývoj platov útočníkov a obrancov na základe rôznych štatistík. Platy hokejových hráčov najvyššej hokejovej ligy na svete (NHL- national hockey league) boli prvýkrát zverejnené v sezóne 1989/1990. Táto nová politika vytvorila nekonečný priestor na možné špekulácie medzi hráčmi a tímami, ktorú sledujeme dodnes. Výsledky mnohých analýz ukazujú, že medzi hráčmi existujú výrazne rozdiely v priemerných platoch a odmenách na základe výkonu. Čiastočne sa to dá vysvetliť aj rozdielmi v tímových príjmoch, ale čoraz väčší dôraz sa kladie na výkonnosť jednotlivých hráčov. Na určenie predikcie platu sme využili rôzne modely a metódy. Platy v hokeji sa môžu značne líšiť v závislosti od úrovne hry, ligy, zručností hráča, skúseností a ďalších faktorov. Ako sme spomenuli, NHL predstavuje najvyššiu úroveň profesionálneho hokeja a hráči v tejto lige zvyčajne zarábajú najvyššie platy. NHL funguje podľa systému platových stropov, ktorý stanovuje limit na celkovú sumu, ktorú môžu tímy minúť na platy hráčov. Minimálny plat pre hráčov NHL sa každý rok líši, ale vo všeobecnosti sa pohybuje od približne 700 000 USD do viac ako 1 milióna USD. Špičkoví hráči však môžu zarábať výrazne vyššie platy, pričom niektorí zarábajú desiatky miliónov dolárov ročne.

II. ANALÝZA SÚČASNÉHO STAVU

Na základe štatistík zozbieraných z hokejovo najvýkonnejšej súťaže sa viacerí autori vo svojich príspevkoch snažia využitím rôznych technológií a modelov určiť čo najpresnejšiu predikciu platov jednotlivých hráčov.

Kimberly Woo sa vo svojej práci „An analysis of NHL defencemen salary in relation to individual on-ice performance“ [1] venoval predikcii platu hokejového obrancu v závislosti od času stráveného na ľadovej ploche. Skúmaný dataset obsahoval informácie o výkone 733 hráčov NHL v sezóne 2017/2018. Dáta obsahovali tri výkonnostné štatistiky a jednu z troch herných situácií. Tieto premenné reprezentovali individuálny výkon hráčov počas 60 minút vo všetkých herných situáciách: hra 5 na 5, presilovky (PP) a oslabenia (PK). Všetky tieto podmnožiny boli spojené do jedného súboru údajov, pričom boli odstránení hráči, ktorí nehrali na pozícii obrancu, a hráči, ktorí boli vymenení do iného tímu počas sezóny. V konečných dátach sa nachádzalo 195 hráčov. Autori dospeli k záveru, že hodnota hráčov na základe plusových/mínusových bodov môže niektorých hráčov nadhodnotiť, zatiaľ čo iných podhodnotiť. Súčasná literatúra týkajúca sa hodnotenia miezd na základe individuálnych výkonnostných štatistík je veľmi obmedzená a zameriava sa najmä na štatistiky skórovania gólov, ktoré nemusia nevyhnutne zdôrazňovať alebo oceňovať úlohu a schopnosti obrancov. Model má 50% predikčnú silu, čo ukazuje, že výkony súvisiace s presilovou hrou, a najmä pri oslabeniach, sú dosť významné pre hodnotu platu.

Autori Mathieu Parisien a prof. Luois- Philippe Morin sa v práci „Reaction to the NHL Salary Cap“ [2] zameriavali na nové obmedzenia platových stropov a rozdelenie platov hráčov v celej NHL. Cieľom tejto štúdie bolo analyzovať ligové štatistiky hráčov v období od sezóny 2000/2001 do sezóny 2011/2012 a skúmať, ako sa tímy prispôbili platovému stropu. Na preskúmanie účinkov platového stropu v NHL autori zhromaždili verejné údaje o platoch každého hráča v lige od roku 2000 do roku 2011 z USA Today. Celkovo mali autori k dispozícii 8 654 pozorovaní. Platy tímov sa so zavedením platových stropov znížili o 30%. Pri preferovanej štatistike sa zistilo, že tímové platy sa znížili o 56% pre tímy s celkovými výdavkami v roku 2003 nad úrovňou v roku 2005. Tímový plat sa tiež zvýšil o 19,6 % pre tímy s celkovými výdavkami v roku 2003 pod úrovňou 2005. Je evidentné, že zavedenie platových stropov v NHL ovplyvnilo spôsob, akým generálni manažéri pridelujú hráčom mzdy. Ukázalo sa, že zavedením maximálneho stropu a minimálneho dna sa tímy mimo týchto limitov museli drasticky prispôbiť novým pravidlám. Taktiež bolo spozorované, že zmeny v tímových výdavkoch boli spôsobené tým, že samotní generálni manažéri zmenili svoje správanie a rozhodovanie o pridelovaní miezd.

V ďalšej štúdií s názvom „A Comparative Analysis of NHL Player Contracts using Linear Weights“ sa autor Mathieu Parisien [3] venoval odhaleniu objektívneho spôsobu, ako vyhodnotiť prínos hráčov pomocou regresnej analýzy, ktorá zohľadňuje význam určitých štatistík pre úspech tímu. Model tvorila viacnásobná lineárna regresia s cieľom identifikovať ako rôzne štatistiky ovplyvňujú počet strelených gólov a obdržaných gólov pre každý tím. Model hráča bol vytváraný v súlade s Winstonovou regresiou lineárnych váh [4], ktorá je využívaná pri hodnotení hráčov bejzbalu. Prístup spočíval v regresii štatistík o góloch a následnej aplikácii výsledkov regresie na štatistiky jednotlivých hráčov. Model odhalil, že najvyšší odhadovaný objektívny plat pre útočníkov bol tesne pod 2,5 milióna dolárov, zatiaľ čo najväčší spravodlivý plat pre obrancu bol viac ako 8,5 milióna dolárov. Model odhalil, že najväčšie platové nezrovnalosti sa vyskytujú pri útočníkoch s priemerným preplatkom približne 2,5 milióna dolárov ročne, zatiaľ čo pri obrancoch to je približne 800 000 dolárov.

III. EXPERIMENTY A VÝSLEDKY

Pri analýze dát sme využili metodológiu CRISP-DM, ktorá zahŕňa komplexný proces pre dolovanie dát v projektoch. Proces pozostáva z jednotlivých fáz. Postupnosť týchto fáz je adaptívna, čo znamená, že nasledujúca fáza je ovplyvnená výstupom z predošlej fázy. Vo všeobecnosti má každá fáza CRIPS-DM svoje vlastné časové nároky, ktoré sa môžu líšiť od veľkosti konkrétneho projektu a jeho charakteristík [5].

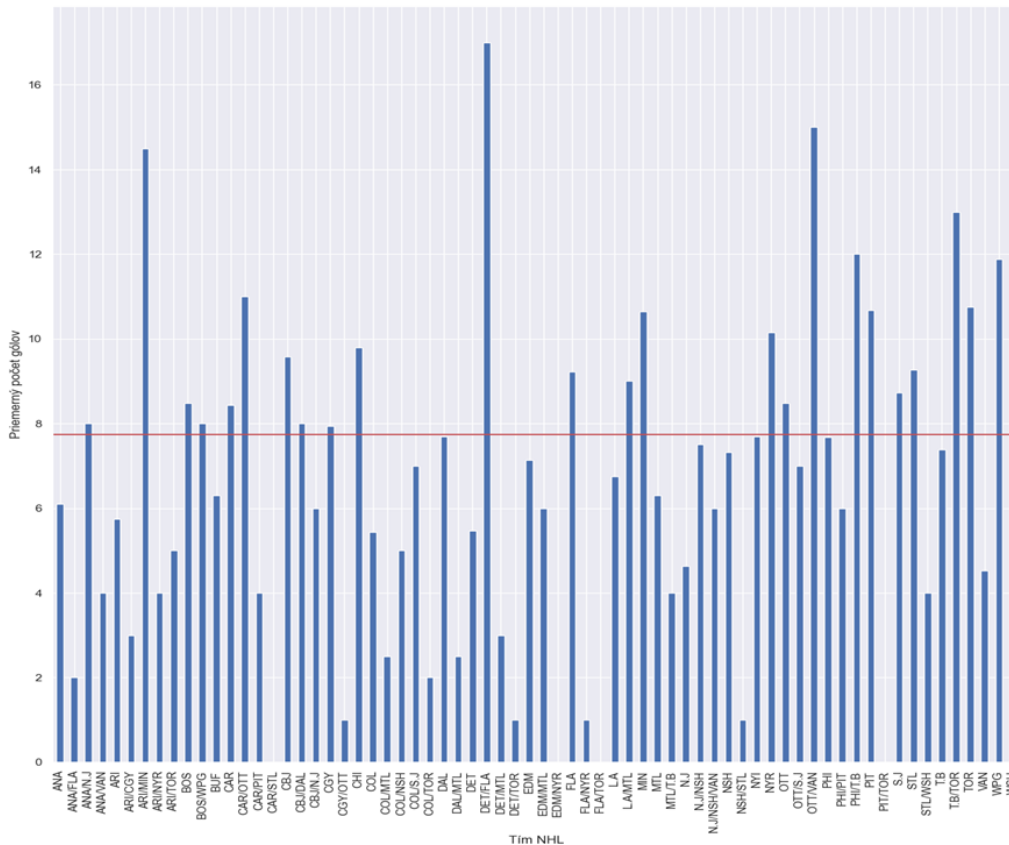
A. Pochopenie cieľa

Cieľom práce bolo na základe dát získaných z hokejových zápasov nastaviť pravdepodobný plat hokejových hráčov. V platovom ohodnotení hráčov častokrát dochádza k rozdielom, kedy niektorí hráči sú nedostatočne platovo ohodnotení, alebo naopak majú príliš vysoký plat. Pri našej analýze sme brali do úvahy iba hráčov na pozícií útočníkov a obrancov. Cieľom bolo dosiahnuť, aby hodnota platov jednotlivých hráčov bola adekvátna ich výkonom podaných v zápasoch a aby hráči, ktorí podávajú nadpriemerné výkony boli dostatočne platovo ohodnotení. Maximálny platový strop pre sezónu 2022/2023 je 82,5 milióna dolárov a minimálny platový strop je 61

miliónov dolárov. V sezóne 2016/2017, ku ktorej sa viažu dáta, bol platový strop 73 miliónov dolárov a minimálna hranica, ktorú mohli jednotlivé tímy minúť bola 54 miliónov dolárov. Cieľom bolo na základe pravdepodobnostných modelov predpovedať plat s čo najvyššou presnosťou, teda s čo najmenšou odchýlkou medzi použitými metódami. Pri predikcii platov sme využili modely lineárnej regresie, rozhodovacieho stromu, k-najbližších susedov, LightGBM, Random Forest a Naive Bayes.

B. Pochopenie dát

Skúmané dáta sme získali z voľne dostupného datasetu¹, ktorý obsahoval 262 riadkov a 154 stĺpcov údajov. Pomocou príkazu `data.describe()` sme z načítaných dát získali základné štatistiky pre každý stĺpec, ako sú priemerná hodnota, štandardná odchýlka, minimálna hodnota, maximálna hodnota, percentily. Na Obr.1 vidíme priemerný počet gólov tímov NHL v sezóne 2016/2017.



Obr. 1 Priemerný počet gólov tímov NHL

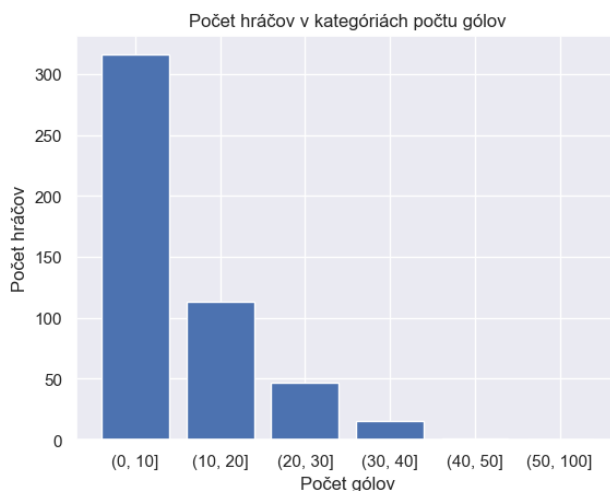
Červená vodorovná čiara znamená celkový priemer gólov. Na základe toho vidíme, že 25 tímov NHL dosiahlo v tejto sezóne viac gólov ako bol stanovený priemer, ktorého hodnota bola po zaokrúhlení 7 gólov. V rámci skúmanej fázy boli taktiež vytvorené korelácie vo vzťahu k cieľovému atribútu, ktorým bol práve plat hráča. Definíciu závislosti medzi atribútmi sme realizovali pomocou korelácie, ktorá môže byť medzi dvoma alebo viacerými premennými. Výsledný korelačný koeficient reprezentuje lineárnu závislosť medzi premennými. Korelácia sa používa ako základná informácia o jej reliabilite. Pomocou funkcie `corr()` sme získali korelačnú maticu na neupravených dátach medzi numerickými atribútmi. Najvyššiu koreláciu s cieľovým atribútom dosahovali atribúty počet zápasov (GP), asistencie na gól (A, A2), čas strávený na ľade (TOI), priemerný čas na ľade za zápas (TOI/GP), strelecké pokusy (S.Bckn, S.Slap.), nezblokované strely (FA) a góly (G.Slap, G.Wrap).

C. Príprava dát

Vo fáze prípravy dát sme si stanovili stĺpce, s ktorými sme pracovali a nepotrebné atribúty dát sme vymazali. Diskretizáciou hodnôt gólov a hodnôt platu sme zoskupili podobné hodnoty do kategórií, ktoré nám umožnili lepšie porovnávať a analyzovať dáta. Ako môžeme vidieť na Obr.2, diskretizáciou počtu gólov strelených hráčmi sme si rozdelili množinu na šesť častí. Zistili sme, že 316 hráčov strelilo od 0 – 10 gólov, 113 hráčov strelilo 10 – 20 gólov, 47 hráčov strelilo 20 –

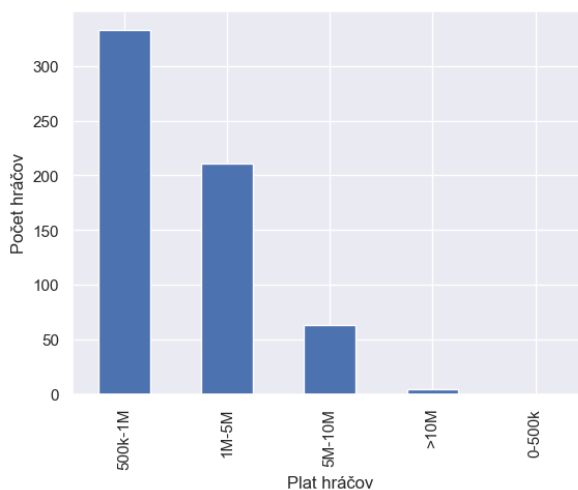
¹ Link na dáta : <https://www.kaggle.com/datasets/camnugent/predict-nhl-player-salaries>

30 gólov, 30 – 40 gólov strelilo 15 hráčov, v poslednej hranici 40 – 50 gólov sa nachádzal jeden hráč a viac ako 50 gólov nestrelil nikto.



Obr. 2 Počet gólov hráčov

Rovnakou metódou sme si do jednotlivých kategórií rozdelili aj štatistický atribút plat hráčov a zistili sme v akých rozpoch sumy sa hráči nachádzajú. Zistili sme, že žiaden hráč nemá plat nižší ako je 500 000 USD. Najpočetnejší stĺpec bol s platom v rozmedzí 500 000 USD až 1M USD, takýto plat má 333 hráčov, 211 hráčov má plat v rozmedzí hodnôt 1M USD až 5M USD, 63 hráčov má plat 5M USD až 10M USD a iba 5 hráčov má plat vyšší ako 10M USD. Priemerný plat hráča bol stanovený vo výške 2 264 509 USD.



Obr. 3 Diskretizácia platu hráčov

Pri konštrukcii dát sme vygenerovali nový stĺpec na základe už preskúmaných dát. Vytvorili sme stĺpec priemerný počet gólov hráča na zápas s názvom AVG_G_Game. Pre predikovanie nášho atribútu plat hráča sme si rozdelili dáta na tréningovú a testovaciu množinu v pomere 80:20.

D. Modelovanie a vyhodnotenie

V nasledujúcej fáze sme sa venovali modelovaniu, ktoré patrí medzi jednu z hlavných fáz procesného modelu CRISP – DM. Táto fáza je dôležitá, pretože sa v nej vytvárajú a testujú rôzne modely, ktoré nám pomáhajú porozumieť vzťahom a vzorom v dátach. V rámci tejto fázy sme taktiež zhodnotili výsledky použitých modelov z pozície útočníkov, ako aj obrancov. Na vyhodnotenie chýb predikovania platu sme použili R- kvadrátový koeficient, Mean Squad Error a Mean Absolute Error. V Tab.1 môžeme vidieť vyhodnotenie chýb merania pri predikovaní platu útočníkov.

Tab. 1 Vyhodnotenie chýb predikovania útočníkov

	R- kvadrátový koeficient	Mean squared error (MSE)	Mean absolute error
Lineárna regresia	0,216	2,8075.10 ¹²	1 152 871
K- najbližších susedov	0,013	5,3728.10 ¹²	1 355 551
Rozhodovací strom	0,113	4,1799.10 ¹²	1 140 468
LightGB Model	0,539	2,2797.10 ¹²	977 384
Random Forest	0,470	2,6115.10 ¹²	1 062 506
Naive Bayes	0,000	6,5863.10 ¹²	1 564 728

Najhoršie výsledky predikovania platu sme dostali pri použití metódy Naive Bayes. Najlepšie výsledky sme dostali pri metóde Random Forest modelu a pri modelovaní pomocou LightGBM. V prípade Mean Absolute Error dosahoval model LightGBM je priemernú absolútnu chybu 977 384 čo znamená, že tento model robí chybu $\pm 977\,384$ USD pri predpovedaní platu hráča vzhľadom na jeho štatistiky. V Tab. 2 môžeme vidieť vyhodnotenie chýb merania pri pedikovaní platu obrancov.

Tab. 2 Vyhodnotenie chýb predikovania obrancov

	R- kvadrátový koeficient	Mean squared error (MSE)	Mean absolute error
Lineárna regresia	0,444	4,6653.10 ¹²	1 545 448
K- najbližších susedov	0,025	4,6752.10 ¹²	1 428 959
Rozhodovací strom	0,100	7,0822.10 ¹²	1 802 646
LightGB Model	0,622	1,4369.10 ¹²	866 402
Random Forest	0,650	1,3281.10 ¹²	836 288
Naive Bayes	0,170	4,2107.10 ¹²	1 373 125

Môžeme zhodnotiť, že podľa R – kvadrátového koeficientu je najlepší model na predikciu platu Random Forest s hodnotou 0.650. Druhý v poradí je LightGBM s hodnotou 0.622. Priemerná absolútna chyba (Mean Absolute Error) 866 402 znamená, že model robí chybu $\pm 866\,402$ USD pri predpovedaní platu hráča pomocou LightGBM. Presnejší model Random Forest robí priemernú absolútnu chybu $\pm 836\,402$ USD. Tieto dva modely sa podľa R – kvadrátového koeficientu líšia len o 3 stotiny v prípade priemernej absolútnej chyby (Mean Absolute Error) sa tieto dva modely líšia v predikcii platu o $\pm 30\,114$ USD.

Pri vyhodnotení sme sa zamerali na tím Washington Capitals, ktorý má značnú históriu. Náš dataset obsahoval údaje o 17 hráčov z tímu Washington Capitals, 5 obrancov a 12 útočníkov. Najvyššie plateným hráčom podľa dostupných údajov bol Alexander Ovechkin s 10 miliónmi dolárov na rok, najmenší plat má Paul Carey a Aron Ness so sumou 575 000 dolárov na rok. Priemerný plat týchto 17 hráčov je 3 018 823,52 dolárov.

Tab. 3 Porovnanie skutočného a odhadovaného platu (LightGBM)

Hráči	Skutočný plat	Odhadovaný plat	Rozdiel
Tom Wilson	1 750 000	2 280 000	530 000
Riley Barber	667 500	1 640 000	972 500
Evgeny Kuznetsov	3 400 000	4 350 000	950 000
Brooks Orpik	5 500 000	2 810 000	-2 690 000
Aron Ness	575 000	1 680 000	1 105 000
Dmitry Orlov	2 570 000	3 730 000	1 160 000
Andre Burakovsky	832 500	3 430 000	2 597 500
Alexander Ovechkin	10 000 000	3 590 000	-6 410 000
Taylor Chorney	750 000	1 820 000	1 070 000
Daniel Winnik	1 800 000	3 510 000	1 710 000
Paul Carey	575 000	1 700 000	1 125 000
Justin Williams	3 250 000	3 980 000	730 000
Matt Niskanen	5 750 000	3 430 000	-2 320 000
Nicklas Backstrom	7 000 000	4 530 000	-2 470 000
Lars Eller	4 250 000	3 920 000	-330 000
Nate Schmidt	875 000	3 130 000	2 255 000
Jay Beagle	1 750 000	4 240 000	2 490 000
Suma	50 620 000	53 770 000	3 150 000

Môžeme konštatovať, že piati hráči si podľa predikcie platovo pohoršilo a dvanástim hráčom by sa plat vylepšil, čo ma za následok aj zvýšenie celkovej sumy platov. Najvýraznejšiu zmenu možno pozorovať u hráča Alexandra Ovechkina, ktorý bol najlepšie plateným hráčom s platom 10 000 000 USD, po našej predikcii by si pohoršil o 6 410 000 USD. Celková suma platov týchto sedemnástich hráčov je 53 770 000 USD po našej predikcii došlo k zvýšeniu celkovej sume platu o 3 150 000 USD.

IV. ZÁVER

Predikčné modely nám vedú poskytnúť odhady, avšak nie sú neomylné, pretože platy hráčov môžu byť ovplyvnené rôznymi subjektívnymi faktormi, dynamikou vyjednávania a výkyvmi na

trhu. Využitím historických údajov a relevantných faktorov však predikčné modely môžu poskytnúť cenné informácie o odhadovaní platov hráčov v hokeji. Predikcia platov je užitočným nástrojom, ktorý môže pomôcť NHL klubom pri hodnotení a rozhodovaní o ich finančnom ohodnotení. Na základe predpovede, ktorá bola získaná, môžu tímy spravovať svoje rozpočty efektívnejšie a rozhodovať o tom, koľko peňazí by mali investovať do jednotlivých hráčov. Nasadenie modelu môže byť rozhodujúcim faktorom pri posúdení, či hráč je hodný ponúknutia nového kontraktu alebo predĺženia. Manažéri tímov môžu využívať predikcie na hráčov, ktorí majú vysoký potenciál a nízku cenu. Kluby môžu použiť tieto predikcie na predpoveď budúcich nákladov na hráčov a plánovania zdrojov vzhľadom na platový strop NHL. Úžitok v tom nájdú aj agenti hráčov, ktorí predikciou platov získajú lepšie porozumenie trhových trendov a budú môcť pomôcť svojim klientom pri rozhodnutiach o budúcich zmluvách. Na vytvorenie predikcií a ich následné porovnanie sme použili metódu lineárnej regresie, LightGBM, Random Forest, Naive Bayes, metóda najbližších susedov a rozhodovacie stromy. Najlepšie hodnoty spomínaných použitých modelov dosahoval práve LightGBM, ktorý vďaka svojej schopnosti riešiť problém s nevyváženými dátami a rýchlemu gradientovému zosilňovaniu, dosiahol najlepšiu presnosť predikcií platov hráčov NHL.

REFERENCIE

- [1] Kimberly Woo.: An analysis of NHL defensemen salary in relation to individual on-ice performance, 2019, [online], [cit. 2022-10-15] Dostupné na internete: <https://ckmsports.com/wp-content/uploads/2019/04/CKM-UBC-NHL-DefenseMen-Salary-Study.pdf>
- [2] Parisien, M., Morin, Louis-Philippe, Reaction to the NHL Salary Cap. University of Ottawa [online], [cit. 2022-10-15], Dostupné na internete: https://ruor.uottawa.ca/bitstream/10393/36326/1/Parisien%20Mathieu_Final%20Poster.pdf
- [3] Parisien, M. : A Comparative Analysis of NHL Player Contracts using Linear Weights, 2019, [online], [cit. 2022-10-20] Dostupné na internete: https://www.econ.queensu.ca/sites/econ.queensu.ca/files/student_papers/573.pdf
- [4] Winston, Wayne L. *Mathletics: How Gamblers, Managers, and Sports Enthusiasts Use Mathematics in Baseball, Basketball, and Football* (New in Paper). Princeton University Press, 2012.
- [5] Saltz, Jeff, Nicholas Hotz, David Wild, and Kyle Stirling. "Exploring Project Management Methodologies Used Within Data Science Teams Orleans, LA, USA, August 16-18, 2018." 24th Americas Conference on Information Systems, AMCIS 2018, New Orleans, LA, USA, August 16-18, 2018: Association for Information Systems. [cit. 22.12.2022]

Vplyv elektromobility na životné prostredie

¹Jozef HUMENÍK, ²Jaroslav DŽMURA

¹ Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

² Katedra elektroenergetiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹jozef.humenik@student.tuke.sk, ²jaroslav.dzmura@tuke.sk

Abstract — Tento článok hodnotí vplyv elektrických vozidiel na životné prostredie a hospodársku politiku, ktorá podporuje rozvoj a širšie využitie elektromobility, a taktiež hodnotí energetickú efektívnosť a environmentálny vplyv batériových elektrických vozidiel, ktoré sú jedinou technickou alternatívou na dnešnom trhu pre vozidlá so spaľovacími motormi.

Keywords — elektromobilita, emisie skleníkových plynov, energetika, ochrana životného prostredia

I. ÚVOD

Elektromobilita ako koncept trvalo udržateľnej dopravy a vozidla na elektrický pohon sú vo svete uznané za kľúčovú technológiu, ktorá ma pomôcť k znižovaniu úrovne emisií a spotreby energie v sektore mobility. Elektrickú energiu vo vozidle môže zabezpečiť batéria (angl. BEV - Battery Electric Vehicle) alebo palivový článok. V článku sa popisuje technická štruktúra, ktorá je relatívne jednoduchá v porovnaní s automobilmi so spaľovacími motormi. Potenciálne hroziaca energetická kríza je jedným z kľúčových dôvodov rozvoja technológií obnoviteľných a alternatívnych palív. Životné prostredie trpí najmä v krajinách s rozvinutým priemyslom, kde množstvo vyprodukovaných emisií skleníkových plynov má zvyšujúcu sa tendenciu, a každoročne dosahuje novú, vyššiu úroveň. V tejto oblasti sa stále skúmajú novšie, efektívnejšie a modernejšie filtračné systémy a iné spôsoby detoxikácie ovzdušia. Najväčším podielom k tomu prispieva najmä výrazné zníženie produkcie skleníkových plynov, čo je významný krok k dosiahnutiu tohto cieľa. O plne elektrických vozidlách sa tvrdí, že neprodukujú žiadne výfukové plyny, ale to nie je úplne správne tvrdenie a záleží od uhla pohľadu, lebo to nezohľadňuje emisie, ktoré sú výsledkom výroby elektrickej energie použitej na nabíjanie práve týchto vozidiel, alebo na budovanie nabíjacej infraštruktúry.

II. TRENDY, VÍZIE, ROZVOJ A AKTUÁLNY STAV

Globálny automobilový trh s elektrickými vozidlami za posledných 10 rokov zaznamenal výrazný pokrok. Podľa najnovšie dostupných údajov je počet ľahkých elektrických vozidiel v rokoch 2019 - 2020 vyše 2,2 milióna kusov, čo predstavuje nárast asi 9 % oproti roku 2018.

V súčasnosti, v čase spamätávania sa z globálnych, hospodárskych a iných následkov pandémie ochorenia COVID-19 a prebiehajúceho vojenského konfliktu na Ukrajine, vidia napriek tomu odborníci z automobilového priemyslu v nasledujúcich rokoch svetlú budúcnosť. Za posledných niekoľko rokov sa odpozorovalo, že v budúcnosti bude asi 60 % medzinárodného predaja osobných áut pochádzať kategórie elektrických vozidiel, avšak elektromobily budú tvoriť necelú 1/3 z celkového počtu všetkých vozidiel na cestách. Elektrické vozidlá budú aj naďalej tzv. „vozidlom budúcnosti“. Revolúcia elektrických vozidiel sa neustále vyvíja a napreduje míľovými krokmi. Keďže technológia elektrických vozidiel nie je vo svojej podstate zložitá, napr. okrem batérie je možné ojazdené vozidlá prestavať zo spaľovacích na elektrické. Takto je možné znížiť emisie ekvivalentu oxidu uhličitého (CO₂) o 80 %, čo je 5-krát efektívnejšie.

Lítium-iónové batérie uchovávajú energiu pre úplne elektrické vozidlo. Lítium-iónové batérie tiež umožňujú vozidlu prejsť vyše 500 km na jedno nabitie, ale skutočný počet najjazdených kilometrov sa líši podľa modelu vozidla a štýlu jazdy. V súčasnosti obsahuje tento typ batérie dvojnásobné množstvo energie pri rovnakej hmotnosť batériových článkov. Dôvodom preferovania tohto typu batérií, je ich:

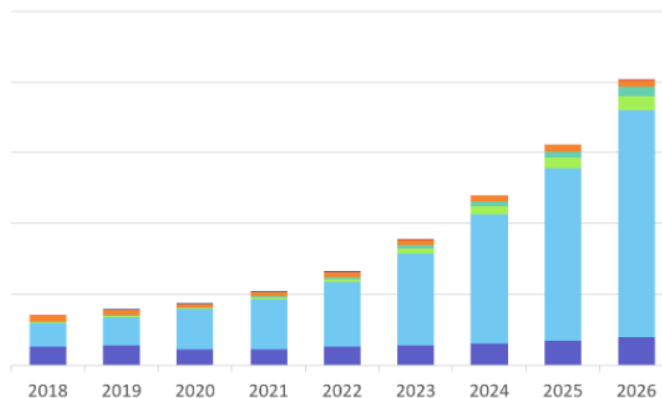
- vysoký pomer výkonu k hmotnosti,
- vysoká energetická účinnosť,
- dobrá výkonnosť pri vysokých teplotách,
- nízke samovybíjanie.

Tieto vyššie uvedené faktory robia z lítium-iónových batérií (Li-Ion) najslubnejšiu formu uchovávaní elektrickej energie vo vozidlách. V súčasnosti sa testujú a skúmajú rôzne chemické zloženia lítium-iónových batérií. Napr. vozidla Nissan Leaf a BMW i3 používajú lítium (Li)-nikel (Ni)-mangánovo (Mn)-kobaltovú (Co) batériu, ktorá má energetickú hustotu 140 – 180 Wh/kg. Problémom je najmä také chemické zloženie, ktoré umožní zvýšiť energetickú náročnosť a zároveň zachovať stabilitu pri vysokých teplotách a opakovaných nabíjaciach cykloch. Model Tesla S využíva chemické zloženie batérií obsahujúce oxid lítový (Li_2O), oxid nikelnatý (NiO), oxid kobaltu a oxid hliníty (Al_2O_3) [10][11].

V súčasnosti sa používajú najmä 3 druhy Li-ion batérií, ktorých spoločným ukazovateľom je omnoho vyššia životnosť v porovnaní s olovenými, NiMH a NiCD, vysoká hustota energie a vysoký nabíjaci prúd. Súčasným trendom je zníženie nákladov na výrobu a zvyšovanie energetickej hustoty. Medzi nich patria:

- NCA (nikel-kobalt-hliník) – vysoká hustota energie a vyššia cena, používajú sa väčšinou pre vozidlá vyšších tried s dlhým dojazdom;
- NCM (nikel-mangán-kobalt) – podobné zložením ako NCA – podobné vlastnosti;
- LiFePO_4 (lítium-železo-fosfát) – majú nižšiu cenu než NCA a NCM a sú bezpečnejšie, ale majú nižšiu energetickú hustotu. Použitie si nachádzajú najmä v stacionárnych úložiskách (napr. fotovoltika), lebo tam až tak nezáleží na hmotnosti a rozmeroch. Ale svoje miesto si nájdu aj v automobiloch s kratším dojazdom.

Medzi inovácie výrobcov a ich vývojárov sa začína skúmať kombinácia lítium-síra. Tento typ batérie sa vyznačuje jednoduchšou výrobou – bez zložitých zlúčenín, porovnateľnou veľkosťou a tretinovou hmotnosťou.



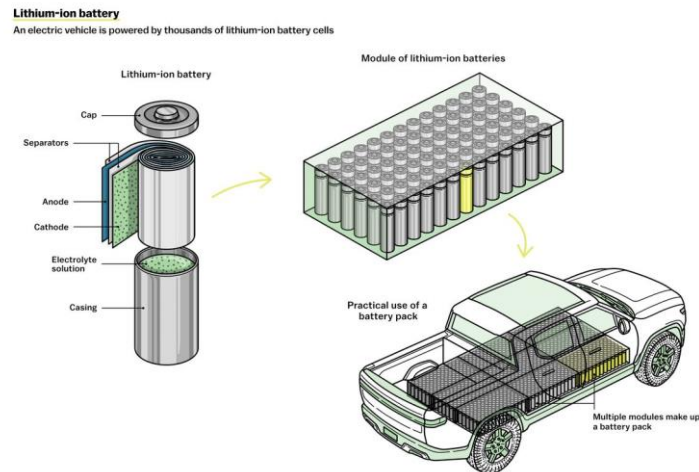
Obf. 1 Globálny aktuálny a predpokladaný dopyt po Li-Ion batériách pre elektrické vozidlá, pomerovo v GWh [2]

Odhaduje sa, že trh s lítium-iónovými batériovými článkami v elektrických vozidlách bude mať do roku 2026 hodnotu takmer 70 miliárd \$ (asi 65 miliárd €; kurz 1 USD = 0,925 €), čo je predpoveď pre elektrické autá, autobusy, dodávky a nákladné autá [2]. Vo všeobecnosti používanie lítium-iónových batérií má ďaleko širšie využitie, než len pre účely použitia v elektrických vozidlách. Tento trh vzrástol medzi rokmi 2018 a 2020 až o 73 %. Predpokladá sa, že cena bude časom klesať, čo by mohlo prispieť k spusteniu nového nárastu výroby elektrických vozidiel, podobne ako tomu bolo v roku 2008. Tu platí lineárna súvislosť, že čím bude jazdiť viac elektrických vozidiel na našich cestách, tým to bude mať väčší vplyv na redukciiu produkcie emisií. Je technicky nespochybniteľné, že pri vysokom zaťažení batérií sa zároveň zvyšuje opotrebovanie absolútne všetkých batérií. Ale prvky, ktoré sú optimalizované pre takýto výkon, sú odolnejšie a spoľahlivejšie pri vysokom zaťažení. Stále prebiehajú výskumy a testovania nových materiálov pre lítium-iónové batérie s cieľom zvýšenia ich energetickej hustoty pre použitie v elektrických vozidlách, aby sa zvýšila vzdialenosť dojazdu vozidla na jedno nabitie, vďaka čomu sa stanú praktickejšie pre svojich majiteľov, ktorí napr. dochádzajú každý deň do práce [7][9][11][12][13]. Autá zohrávajú v spoločnosti osobitnú úlohu z troch dôvodov v porovnaní s inými druhmi prepravy:

- a) autá dominujú v premávke vo väčšine krajín sveta,
- b) najvyššia miera rastu predaja vo svete je v oblasti predaja aut,
- c) existujú alternatívne technológie transmisie, napr. na rozdiel od nákladných áut.

Zatiaľ čo menšie nákladné autá môžu v obmedzenej miere jazdiť aj na elektrinu, ale pri veľkých nákladných autách to už úplne neplatí, a sú závislé od fosílnych kvapalných palív - spaľovacie motory (benzín, nafta). Tieto paliva možno v budúcnosti premeniť na zmes 80 % (fosilny alebo biogénny/ekologický), ale je nutné zabezpečiť, aby zvýšenie podielu biozložiek nespôsobovalo technické problémy s palivovým systémom alebo samotným motorom. Čo sa týka autobusov, tak tie môžu taktiež na obmedzené vzdialenosti jazdiť aj na elektrinu. Už v súčasnosti sa vo svete bežne používajú autobusy na stlačený zemný plyn (angl. CNG - Compressed Natural Gas), čo je alternatívne palivo pre spaľovacie motory. V uliciach taktiež jazdia autobusy poháňané palivovými článkami, zatiaľ čo malé nákladné vozidlá poháňané palivovými článkami a molekulárnym vodíkom (H_2) sú stále na úrovni konceptov [1][3][7].

V budúcnosti možno očakávať obrovský nárast dopravy po celom svete, najmä v rozvojových krajinách. Z globálneho hľadiska je v uliciach cca. 630 miliónov áut, ktorých počet by mohol do roku 2030 stúpnuť na vyše jednu miliardu. Očakáva sa, že do roku 2030 sa výroba áut zvýši z aktuálnych 63 miliónov na 100 miliónov áut ročne [2]. Moderné vozidlá so spaľovacími motormi, okrem tvorby emisií CO_2 produkujú aj iné nebezpečné toxické emisie. Podľa Svetovej zdravotníckej organizácie (angl. WHO - World Health Organization) možno považovať znečistenie ovzdušia medzi jedno zo zdravotných rizík pre zdravie ľudí, a odhaduje sa, že spôsobuje približne dva milióny predčasných úmrtí na celom svete ročne. Je to aj z dôvodu, že ozón, jemný prach, oxid dusičitý (NO_2) a oxid siričitý (SO_2) boli Svetovou zdravotníckou organizáciou označené za najnebezpečnejšie druhy spojené hlavne alebo výrazne s cestnou dopravou. Cestná doprava bude tvoriť približne polovicu týchto kvantitatívnych nákladov na život a zdravie. Toxické emisie spôsobujú vysoké zdravotné náklady dokonca aj v priemyselne rozvinutých krajinách. Napr. v Číne, Indii, Spojených štátoch alebo Ruskej federácii množstvo obyvateľov žije menej ako 500 metrov od ciest, po ktorých každý rok prejde viac než tri milióny vozidiel. Aj v dôsledku vysokej úrovne znečistenia ovzdušia zomrú vo svete každý rok takmer štyri milióny ľudí. Za účelom naplnenia potrieb mobility, zdravého a bezpečného udržateľného rozvoja je potrebné znížiť emisie súvisiace s ovzduším a zdravím. Postupne je potrebné optimalizovať závislosť na rope, lebo dnes zažívame tzv. „ropný vrchol“. Z tohto pohľadu by mali byť terajšie druhy pohonov nahradené efektívnejšími, ekologickejšími a bezpečnými alternatívami; najmä pri súčasnom takmer celosvetovom trende prechodu na udržateľnú spoločnosť a udržateľný spôsob života. Pritom je potrebné túto problematiku riešiť komplexne, napr. z právneho hľadiska, technického (napr. efektívne mobilné a IT technológie), finančného, hospodárskeho a podobne. Viaceré krajiny sveta, napr. Nemecko, Švédsko a Dánsko sa rozhodli zmeniť svoju výrobu elektrickej energie z fosílnych palív na výrobu z obnoviteľných zdrojov energie, čo ešte viac prispieva k zlepšeniu ochrany životného prostredia a menšiemu vplyvu elektrických vozidiel v porovnaní s vozidlami so spaľovacími motormi. Elektrické vozidlá z toho pohľadu omnoho menej negatívne vplyvajú na životné prostredie a ľudské zdravie. Keďže doprava vo všeobecnosti (cestná, vodná, dráhová a letecká) je v súčasnosti najväčším zdrojom znečistenia životného prostredia v mnohých častiach sveta. Keďže je potrebné zabezpečiť kvalitu života, mobilitu ľudí (napr. cesta do práce, za rodinou, ...) a zachovať funkčnosť všetkých odvetví národného hospodárstva (napr. zachovanie základných funkcií štátu – ochrana obyvateľstva a obrana územia štátu, a pod.), tak na koncepčné vyriešenie tohto problému neprichádza do úvahy veľké obmedzovanie využívania vozidiel, ale efektívnym riešením je, aby vozidlá na cestách produkovali čo najnižšie emisie. Keďže emisie vozidiel poškodzujú nielen planétu, ale aj ľudské zdravie. Látky znečisťujúce ovzdušie pri spaľovaní benzínu/nafty môžu spôsobiť zdravotné problémy, ako je astma, rakovina a problémy dýchacieho traktu. Na druhej strane elektrické vozidlá neprodukujú škodlivé výfukové plyny [1][3][4][5][6][7][12][14].



Obr. 2 Zloženie Li-Ion batérií a ilustračná vizualizácia vo vozidle [18]

III. ZHODNOTENIE

Nové výskumy odhaľujú nové zistenia a prinášajú dôkazy, že prechod na elektrické vozidlá má veľký význam pre globálne životné prostredie. Pod tým sa rozumie zlepšenie celkovej kvality ovzdušia a znižovanie uhlíkových emisií, čo je aj cieľom opatrení a politikou EÚ na dosiahnutie cieľov pri znižovaní emisií uhlíka v rámci balíka „Fit for 55“ do roku 2030 [15].

Elektrické vozidlá umožňujú dopravu bez priamych emisií v porovnaní s vozidlami na benzínový/naftový pohon, a celkovo produkujú počas životného cyklu menej emisií než konvenčné vozidlá. Jedným z ukazovateľov je omnoho nižšie množstvo emisií ako pri benzíne a naftě. Majitelia elektrických vozidiel môžu ďalej znižovať emisie elektrinou vyrobenou z neznečisťujúcich, efektívnych a obnoviteľných zdrojov energie, ako je vietor alebo slnko. Elektrické vozidlá na trhu postupne naberajú na sile. Čoraz viac ľudí aj na Slovensku začína siahť po elektromobiloch, čo potvrdzujú aj štatistiky Polície Slovenskej republiky. Len v roku 2017 tvorili vozidlá na alternatívne palivá, najmä elektrické vozidlá, nemalo značnú časť nových registrácií [16][17]. Medzitým tieto nové vozidlá dostali nové možnosti využitia v sektore dopravy v dôsledku zvýšeného dopytu po ľahkých a palivovo úsporných vozidlách. Elektrické vozidlá obsahujú menej pohyblivých komponentov alebo častí ako benzínové a naftové vozidlá. Z technického hľadiska tieto vozidlá neobsahujú mnohokrát finančne nákladné výfukové systémy, a na ich prevádzku je potrebná len malá údržba. Taktiež majitelia takých elektrických vozidiel nemusia kupovať rôzne systémy vstrekovania paliva, štartéry a množstvo súčiastok, ktoré si údržba spaľovacích motorov vyžaduje. Jedinou potrebnou pohyblivou časťou je rotor s minimálnou údržbou pneumatík, porúch a zavesenia. Väčšina výrobcov automobilov ponúka mnohoročnú záruku na batérie elektrických vozidiel.

V neposlednom rade, z motoristického hľadiska možno uviesť tri jasné výhody jazdy na elektrických vozidlách:

- okamžitý krútiaci moment alebo rotácia elektromotora poskytuje oveľa vyššie a rýchlejšie zrýchlenie,
- rekuperačný brzdový systém vracia energiu do batérie, čím sa zníži zrýchlenie a zvyšuje sa jeho účinnosť,
- hmotnosť a nízke ťažisko uľahčujú a zabezpečujú bezpečnú a pohodlnú jazdu.

V konečnom dôsledku platí, že aj napriek pretrvávajúcemu skepticizmu ľudí voči elektrickým vozidlám stále viac ľudí, ktorí majú finančné možnosti, viac uvažuje nad kúpou elektrického vozidla. Okrem pozitívneho vplyvu na životné prostredie, možno vyzdvihnúť dostupnosť a nenákladnú údržbu. Vzhľadom aj na tieto faktory môžeme v rámci „nového normálu“ očakávať prudký rast tohto priemyselného odvetvia, ktorý je už teraz badateľný.

Podľa údajov Slovenskej asociácie pre Elektromobilitu (SEVA) je možné potvrdiť, že počet elektrických vozidiel (plug-in hybrid a batériové elektromobily) registrovaných na Slovensku má stúpajúcu tendenciu, obdobne je to aj pri budovaní nabíjacej infraštruktúry. Na Slovensku je registrovaných asi 6 000 batériových elektrických vozidiel a viac ako 8 000 vozidiel plug-in hybrid. Čo sa týka nabíjacích staníc, tak ich počet je vyše 1 300 kusov, nachádzajú sa asi v 560 miestach. Viac ako 70 % predmetných nabíjacích staníc je typu AC s výkonom do 22 kW.

- [1] RAJAKARUNA, S., SHAHNIA, F., GHOSH, A., 2015. *Plug in electric vehicles in smart grids: charging strategies*. Springer, Heidelberg
- [2] HOLLAND, A., EDMONDSON, J., GEAR, L. *Lithium-ion Batteries for Electric Vehicles 2021-2031*. [online] Dostupné na: <https://www.idtechex.com/en/research-report/lithium-ion-batteries-for-electric-vehicles-2021-2031/814>
- [3] DAS, H.S., RAHMAN, M.M., LI, S., TAN, C.W., 2020. *Electric vehicles standards, charging infrastructure, and impact on grid integration: a technological review*. *Renew Sustain Energy Rev* 120:109618. <https://doi.org/10.1016/j.rser.2019.109618>
- [4] YAN, Q., KEZUNOVIC, M., 2012. „Impact analysis of electric vehicle charging on distribution system“. In: *North American power symposium (NAPS)*, 2012, s. 1–6. ISBN:978-1-4673-2308-6. <https://doi.org/10.1109/NAPS.2012.6336413>.
- [5] SHAREEF, H., ISLAM, Md.M., MOHAMED, A., 2016. *A review of the stageof- the-art charging technologies, placement methodologies, and impacts of electric vehicles*. *Renew Sustain Energy Rev* 64:403–420. <https://doi.org/10.1016/j.rser.2016.06.033>
- [6] MORRISSEY, P., WELDON, P., O'MAHONY, M., 2016. *Future standard and fast charging infrastructure planning: an analysis of electric vehicle charging behavior*. *Energy Policy* 89: 257–270. <https://doi.org/10.1016/j.enpol.2015.12.001>
- [7] DUVALL, M., KNIPPING, E., ALEXANDER, M., 2007. *Environmental assessment of plug-in hybrid electric vehicles, vol 1. Nationwide Greenhouse Gas Emissions*, EPRI, Palo Alto, CA. [online] Dostupné na: https://www.energy.gov/sites/prod/files/oeprod/DocumentsandMedia/EPRINRDC_PHEV_GHG_report.pdf
- [8] Regulation (EU) 2020/22
- [9] MORGAN, J.P., *Driving into 2025: the future of electric vehicles*, October 2018. [online] Dostupné na: https://www.jpmorgan.com/global/research/electric_vehicles
- [10] Nissan Leaf manual, nepublikovaný
- [11] EVlink Catalog 2019 Electric vehicle charging solutions, nepublikovaný
- [12] FOTOUHI, A., AUGER, D.J., PROPP, K., LONGO, S., WILD, M., 2016. *A review on electric vehicle battery modelling: from lithium-ion toward lithium-sulphur*. *Renew Sustain Energy Rev* 56:1008–1021. <https://doi.org/10.1016/j.rser.2015.12.009>
- [13] Bloomberg NEF, *Electric Vehicle Outlook 2020*
- [14] PEREZ-LINKENHEIL, C., *EU energy outlook 2050 - how will Europe evolve over the next 30 years?*, nepublikovaný
- [15] Spravodajstvo Európskeho parlamentu. 2023. *Znižovanie emisii uhlíka: Ciele a politiky EÚ*. [online] Dostupné na: <https://www.europarl.europa.eu/news/sk/headlines/society/20180305STO99003/zniczovanie-uhlikovych-emisii-ciele-a-politiky-eu>
- [16] Polícia Slovenskej republiky. *Celkový počet elektromobilov a hybridov evidovaných v SR k 31.12.2022*. [online] Dostupné na: <https://www.minv.sk/?pocet-evidovanych-elektromobilov-a-hybridov&subor=466881>
- [17] Polícia Slovenskej republiky. *Počty evidovaných osobných vozidiel, osobných elektromobilov a osobných hybridných vozidiel v rokoch 2016 až 2022*. [online] Dostupné na: <https://www.minv.sk/?pocet-evidovanych-elektromobilov-a-hybridov&subor=466883>
- [18] HEILWEIL, R. *How to build a better battery*. [online] Dostupné na: <https://www.vox.com/recode/23027110/solid-state-lithium-battery-tesla-gm-ford>

Návrh systému na podporu výučby

¹Anna Biceková, Alena Čubiňáková²

¹ Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

¹anna.bicekova@tuke.sk, ²alena.cubinakova@student.tuke.sk

Abstrakt — Cieľom predkladaného príspevku bolo navrhnúť a vytvoriť elektronickú pomôcku pre študentov a napomôcť im tak pri upevňovaní vedomostí a príprave na priebežné, ako aj záverečné hodnotenie z predmetu ekonomického zamerania „Účtovníctvo“. Naše riešenie prináša študentom možnosť zoznámiť sa so základnými účtovnými pojmami, účtovnou osnovou, prejsť si vzorové účtovné prípady a v závere sa otestovať v jednotlivých účtovných prípadoch a teoretických otázkach z jednotlivých oblastí základov účtovníctva.

Kľúčové slová — účtovníctvo, podpora výučby, webová aplikácia, elektronická pomôcka

I. ÚVOD

Informačno-komunikačné technológie sú dnes každodennou súčasťou našich životov. Ovplyvňujú našu prácu, štúdium, voľný čas, vzťahy, komunikáciu a taktiež vplyvajú na naše myslenie. V tomto príspevku sa zameriavame na prostriedky pre podporu výučby, ktoré slúžia nielen učiteľom, ale pomáhajú hlavne študentom. Cieľom týchto nástrojov je zlepšiť vyučovanie a učenie pomocou interaktívnych pomôcok a tým zjednodušiť a zlepšiť procesy nášho bežného života. Hlavným cieľom predkladaného článku bolo navrhnúť elektronickú pomôcku, ktorá by slúžila študentom pri vzdelávaní sa a príprave na priebežné alebo záverečné hodnotenia z predmetu Účtovníctvo. Samotné účtovníctvo môžeme chápať ako proces zaznamenávania, analýzy, interpretácie a oznamovania finančných informácií o organizácii alebo podnikateľskom subjekte. Zahŕňa systematické zaznamenávanie finančných transakcií, ich zhrnutie do finančných výkazov a interpretáciu výsledkov s cieľom poskytnúť prehľad a podporiť rozhodovanie. Účtovníctvo ako vyučovací predmet, patrí medzi zložitejšie ekonomické predmety. Pri tradičnom vyučovaní hodín tohto predmetu sa využíva najmä práca s učebnicami, rôznou literatúrou a dokumentáciou. Samozrejme je veľmi dôležité uplatňovať aj rozličnú literatúru a osvojiť si teóriu, avšak pri využívaní informačno-komunikačných technológií študenti lepšie pochopia základné súvislosti, systém účtovania a nebudú účtovať mechanicky, ale zamyslia sa aj nad dôvodom, prečo to je potrebné zaúčtovať práve takto. Z tohto dôvodu je dôležité vybrať pre jeho výučbu správne výučbové metódy a postupy, aby študenti jednoducho pochopili všetky jeho súvislosti. Na základe vyššie uvedeného sme sa rozhodli vypracovať učebnú pomôcku, v ktorej sa študenti budú môcť otestovať v jednotlivých účtovných prípadoch a teoretických otázkach z jednotlivých oblastí účtovníctva. Navrhovaný systém bol vytvorený v podobe elektronického pracovného zošita pre predmet účtovníctvo a jeho hlavným cieľom bolo podporiť a zefektívniť výučbový proces.

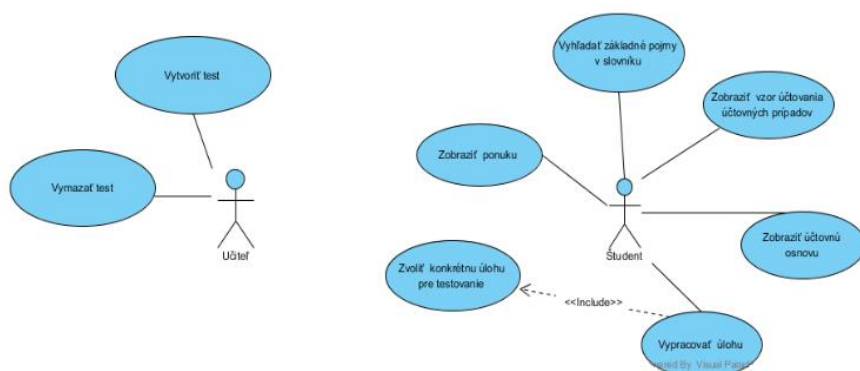
A. Analýza súčasného stavu

V súvislosti s podporou výučby úzko súvisí pojem Technology Enhanced Learning alebo technologicky rozšírené učenie v skratke TEL sa používa ako opis technológií pre aplikácie vo vyučovaní a učení. Môže to byť akákoľvek technológia, ktorá zlepšuje zážitok z učenia. Takto môžeme označiť digitálne, ale i analógové technológie [1]. V dnešnej dobe je to hlavne digitálne TEL, ktoré preberá vzdelávanie vo forme rôznych softvérových riešení. Úlohou TEL je zlepšovať vzdelávanie a vzdelávacie inštitúcie. V dnešnej dobe technológie nie sú iba zdrojom cenných informácií, ale i nástrojom napomáhajúcim k zlepšeniu a uľahčeniu, či už pre študentov alebo učiteľov. Pomocou rôznych technológií pedagógovia zlepšujú interakciu, zapojenie a porozumenie u študentov. Využívajú pri tom rôzne pomôcky na spretrenie a zjednodušenie výuky. Ide hlavne o softvérové riešenia, dáta projektory, interaktívne tabule ale i rôzne techniky,

ktoré častokrát vyvolávajú vyšší záujem o danú problematiku. So zlepšovaním TEL prichádza vzostup diaľkového a hybridného učenia, deti sa učia vlastným tempom, majú prístup k viacerým zdrojom. Tieto technológie ich motivujú, ale i častokrát udržiavajú v kontakte. V reálnom svete sú technológie a ich používanie vo všetkých oblastiach nevyhnutné pre úspech.

II. NÁVRH VLASTNÉHO RIEŠENIA

Pri návrhu a vývoji nášho riešenia sme využili UML diagramy, ktoré slúžia na zjednodušenie implementácie systému. V našom prípade sme ich využili na zobrazenie určitého pohľadu na daný systém [2]. Jednotlivé prípady použitia vychádzajú z požiadaviek našich cieľových používateľov. V nami vytvorenom UC diagrame, ktorý môžeme vidieť na Obr. 1, je uvedený hlavný používateľ systému – študent a taktiež aj používateľ systému administrátor - učiteľ. Títo používatelia majú pridelené jednotlivé prípady použitia, ktoré im systém poskytne.



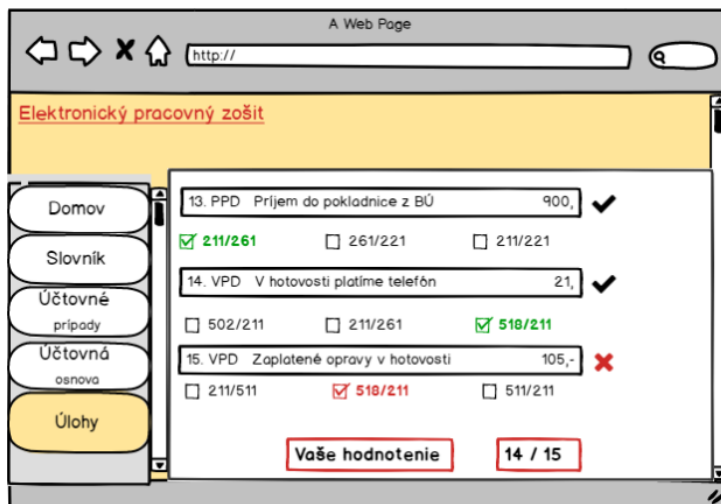
Obr. 1 Diagram prípadov použitia

Ak si vezmeme napr. scenár prípadu- „Vypracovať úlohu“, kde aktérom je študent. Študent si vyberie z ponuky úloh konkrétnu úlohu pre testovanie, zobrazí sa mu prvý test danej úlohy, ktorý vypracuje. Po vypracovaní všetkých otázok testu sa študentovi zobrazí výsledné hodnotenie. Ak je jeho hodnotenie bezchybné, pokračuje na ďalší test tejto úlohy. V prípade, že bezchybne zvolí všetky správne možnosti vo všetkých testoch, ktoré sa v danej úlohe nachádzajú, splní cieľ vybranej úlohy a tým ju celú vypracuje. V prípade ak študent nevypracuje test bezchybne, systém ponúkne študentovi možnosť „Opakuj znova“ alebo „Úlohy“. Ak študent zvolí možnosť „Úlohy“, systém vráti študenta do ponuky úloh, kde má študent možnosť zvoliť inú úlohu a začať test. V prípade ak si študent zvolí možnosť, „Opakuj znova“, systém vráti študenta na otázky prvého testu vybranej úlohy. Predpoklady vykonania tohto scenára je, že systém je on-line dostupný prostredníctvom internetu, študent je prihlásený v systéme, v ktorom existuje minimálne jeden test, ktorý je možné vypracovať.

B. Architektúra riešenia

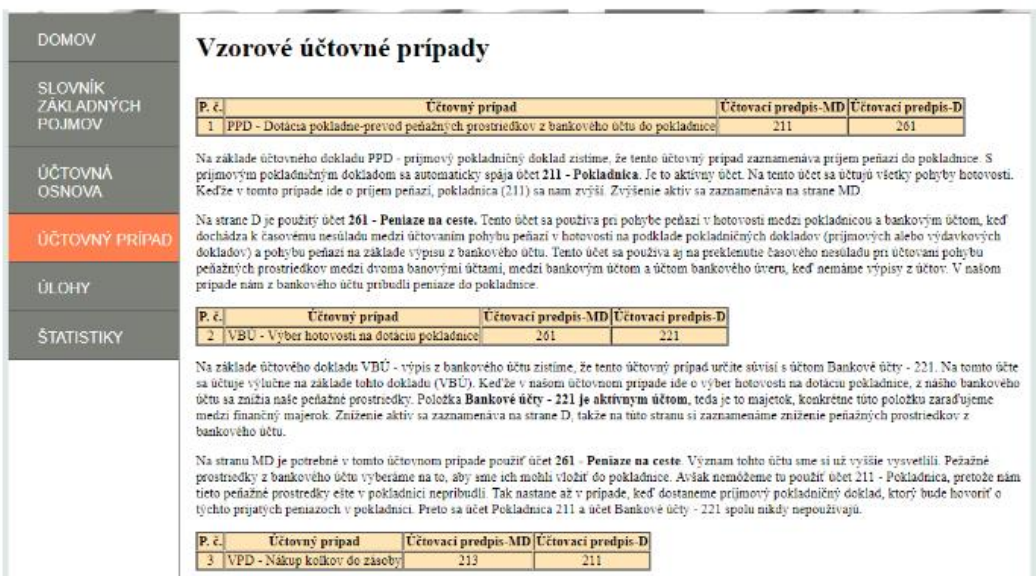
Pri implementácii nášho riešenia sme použili skriptovací jazyk PHP, ktorý sa používa najmä na vytváranie webových aplikácií. Na vývoj našej webovej aplikácie sme používali platformu WampServer, ktorá nám umožnila vytvárať dynamickú webovú stránku s webovým serverom Apache, PHP verziou 7.2.14 a MySQL databázou, kde sme pomocou PhpMyAdmin jednoducho našu databázu spravovali. Ako vývojové prostredie sme používali PhpStorm, ktoré je naozaj prehľadné a obsahuje rôzne nástroj a funkcie. Pri programovaní úloh a testov našej webovej aplikácie sme použili framework CodeIgniter, ktorý je založený na architektúre MVC. Pri tejto architektúre je aplikačná logika oddelená od prezentačnej. Model predstavuje dátovú štruktúru. View sa zobrazuje používateľovi. A Controller je určený ako sprostredkovateľ medzi Modelom a View [2]. V rámci návrhu vlastného riešenia sme sa zameriavali na vytvorenie prvotného návrhu webovej aplikácie vo forme mock-upov. Všetky požiadavky sme počas celej doby vývoja konzultovali s cieľovými používateľmi. Na Obr. 2 môžeme vidieť prvotný návrh vytvorený prostredníctvom jednoduchého a prehľadného nástroja Balsamiq Mockup. Náš systém na podporu výučby účtovníctva sa skladá z viacerých častí. V prvom rade obsahuje slovník základných pojmov, v ktorom sú zaznamenané základné a dôležité pojmy a ich charakteristiky z oblasti základov účtovníctva. Nachádza sa v ňom aj rámcová účtovná osnova, ktorá v takomto systéme pre účtovníctvo určite nesmie chýbať. V účtovníctve je účtovná osnova neustálym pomocníkom a tak aj v tomto systéme slúži ako pomôcka pri jednotlivých úlohách, v ktorých ju

študenti môžu využiť.



Obr. 2 Ukážka Mock-upu vyhodnotenie úlohy

Veľmi nápomocná položka tohto systému pre študentov sú Účtovné prípady. V tejto časti študenti nájdu 8 zaúčtovaných vzorových účtovných prípadov, ktoré sú veľmi podrobne vysvetlené. Je naozaj dôležité pochopiť systém účtovania, ako účtovníctvo vlastne funguje, aké účty sa spájajú s jednotlivými účtovnými dokladmi, aké účty sa spolu používajú, alebo ako je pri účtovaní potrebné postupovať. V tejto časti sú tieto konkrétne vzorové účtovné prípady vybrané na základe hodín predmetu Základy účtovníctva. Počas absolvovania tohto predmetu malo veľa študentov problém pochopiť účtovanie takýchto typov účtovných prípadov. Napríklad zložité a nepochopiteľné bolo dlhú dobu najmä používanie účtu 261 – Peniaze na ceste, takže vzorový účtovný prípad s týmto účtom sa tam určite musel nachádzať. Výhodou tejto časti je aj to, že pri vysvetlení daného účtovného prípadu sú spomenuté aj súvislosti, ktoré sa netýkajú len konkrétneho účtovného prípadu ale aj všeobecne daného účtu a to môže študentom naozaj veľmi dopomôcť k riešeniu úloh a ich jednotlivých testov. Vzorové účtovné prípady, boli čerpané z cvičebnice, ktorá je výstupom projektu Kvalita vo vzdelávaní – úspech v živote, realizovaného v Obchodnej akadémii vo Vranove nad Topľou v rámci operačného programu Vzdelávanie [3].



Obr. 3 Ukážka webovej aplikácie- účtovný prípad

Najdôležitejšou časťou systému je položka s názvom Úlohy. Nachádza sa tu spolu 6 úloh. Každá jednotlivá úloha obsahuje buď 2 alebo 3 testy. Začiatkové úlohy zahŕňajú základné teoretické otázky, avšak postupne ich náročnosť stúpa. Začínajú sa v nich objavovať aj otázky, v ktorých je nutné zvoliť možnosť so správnym zaúčtovaním alebo so správnym použitím účtu k danému účtovnému prípadu. V testoch je použitý princíp gamifikácie. Vzdelávanie sa študentom viac zatriktívni, stáva sa pre nich zaujímavejšie a určite aj motivujúce. Princíp danej úlohy je založený

na tom, že ak sa z jedného testu chceme dostať k ďalšiemu, ktorý je ako keby vyšším stupňom alebo levelom našej úlohy, musíme úspešne, bezchybne zvládnuť vyriešiť prvý test. Následne sa nám otvorí ďalší, náročnejší test našej úlohy. Po jeho bezchybnom zvládnutí, v prípade, že ide o úlohu, ktorá obsahuje 2 testy, sme zvládli vypracovať celú prvú úlohu s dvoma testami a splnili sme jej cieľ. Všetky otázky jednotlivých testov boli čerpané a tvorené z nasledujúcich zdrojov [4] [5]. Nasledujúci Obr.4 zobrazuje bezchybne vyriešený test.

1. Ako oceňujeme hmotný majetok, ktorý je vytvorený vlastnou činnosťou ?
vlastnými nákladmi

2. Čo sú nepôsobiace hospodárske operácie ?
operácie, ktoré spôsobujú zmeny v majetku a zdrojoch krytia, čiže v súvahe

3. Aký typ nepôsobiacej hospodárskej operácie nastane pri účtovnom prípade "FD. Prijemka - Nákup materiálu na sklad od dodávateľov" ?
zvýšenie A, zvýšenie P

4. Aký typ nepôsobiacej hospodárskej operácie nastane pri účtovnom prípade "VPD - Nákup kolkov a poštových známok" ?
zvýšenie A, zníženie A

5. Aký typ pôsoiacej hospodárskej operácie sa spája s účtovným prípadom "VPD - Nákup kancelárskych potrieb a ich okamžitá spotreba" ?
zvýšenie nákladov, zníženie aktív

Počet Vašich bodov je:
5/5

Úlohy

Grafulujem! pokračuj na ďalší test

Obr.4 Ukážka webovej aplikácie- bezchybne vyriešený test

Naša webová aplikácia bola testovaná pomocou stránky www.hostinger.sk, kde sme si zriadili hosting. Cieľoví používatelia, čiže študenti, mali za úlohu sa do našej webovej aplikácie prihlásiť a vyskúšať si ju. Následne vyplnili dotazník, ktorý obsahoval niekoľko otázok a tvrdení. Na základe nášho testovania môžeme zhodnotiť, že:

- 100% opýtaných cieľových používateľov prostredníctvom dotazníka považuje aplikáciu za čitateľnú a nekomplikovanú,
- 75% zo všetkých opýtaných cieľových používateľov prostredníctvom dotazníka tvrdí, že sa v našej webovej aplikácii rozhodne dá jednoducho orientovať, pretože je prehľadná,
- 75% zo všetkých opýtaných cieľových používateľov prostredníctvom dotazníka hodnotí aplikáciu ako rozhodne motivujúcu pri riešení úloh a ich testov s vedomím, že sa po správnom vypracovaní všetkých otázok testu posunú na ďalší test,
- 75% zo všetkých opýtaných cieľových používateľov prostredníctvom dotazníka zhodnotilo, že sú nepochybne spokojní s celkovou funkcionalitou aplikácie.

III. ZÁVER

Cieľom predkladaného príspevku bolo použitím vhodných technológií navrhnuť a vytvoriť elektronickú pomôcku pre študentov a napomôcť im pri príprave a upevňovaní vedomostí pri príprave na priebežné a záverečné hodnotenie z predmetu. Cieľovým používateľom bola prostredníctvom nášho riešenia predstavená inovatívna forma výučby a samovzdelávania sa. Používatelia aplikáciu kladne ohodnotili a ocenili túto formu podpory výučby. Samotnú aplikáciu považujú za prehľadnú, jednoducho ovládateľnú a motivujúcu pri riešení úloh a dosahovaní lepších študijných výsledkov.

REFERENCIE

- [1] Adrian Kirkwood & Linda Price (2014) Technology-enhanced learning and teaching in higher education: what is 'enhanced' and how do we know? A critical literature review, *Learning, Media and Technology*, 39:1, 6-36, DOI: 10.1080/17439884.2013.770404
- [2] Čápka, D. Lekce 1 – Úvod do UML. [online]. Dostupné na inernete: <<https://www.itnetwork.cz/navrh/uml/uml-uvod-historie-vyznam-a-diagramy/>>
- [3] British Columbia Institute of Technology. 2014. Model-View-Controller. [online]. 2019. Dostupné na internete: https://www.codeigniter.com/user_guide/overview/mvc.html
- [4] Chrapeková, M. – Ocilková, A. – Strmeňová, M. 2013. Účtovníctvo pre 2. ročník : Cvičebnica. Obchodná akadémia Vranov nad Topľou. 33 s. ITMS 26110130257
- [5] Andrejovská, A. – Tušan, R. 2015. Základy podvojného účtovníctva : Zbierka príkladov. Košice : elfa, s. r. o. 148 s. ISBN 978-80-8086-213-8.

Možnosti použitia postupností OEIS pri štatistickom testovaní obrazových šifrovacích algoritmov

Jakub ORAVEC, Ľuboš OVSENÍK, Samuel ANDREJČÍK

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

{jakub.oravec, lubos.ovsenik, samuel.andrejcik}@tuke.sk

Abstrakt – Tento príspevok stručne predstavuje databázu celočíselných postupností OEIS a skúma možné použitie postupností z tejto databázy v obrazových šifrovacích algoritmoch. Pri týchto algoritmoch je problematická distribúcia šifrovacích kľúčov pre zaručenie opakovateľnosti experimentov inými výskumníkmi, pričom momentálne sa v publikáciách uvádzajú väčšinou celé kľúče. V prípade, ak by bolo možné niektoré zo záznamov databázy OEIS použiť ako šifrovací kľúč, pre distribúciu kľúčov by stačilo uvádzať iba identifikátor záznamu v databáze OEIS. Okrem zjednodušenia práce výskumníkov by to časom mohlo viesť aj ku ustáleným databázam experimentálnych kľúčov, ktoré by sa používali podobne ako databázy experimentálnych obrazov.

Kľúčové slová – binárny rozvoj, matematická konštanta, obrazový šifrovací algoritmus, OEIS, postupnosť

I. ÚVOD

Prvé obrazové šifrovacie algoritmy (OŠA) boli navrhnuté koncom 90. rokov minulého storočia s myšlienkou vylepšenia štatistických vlastností výsledných zašifrovaných obrazov [1]. Ku týmto vlastnostiam patrí najmä:

- čo najvyrovnanjšie rozloženie intenzít obrazových prvkov v histograme obrazu,
- čo najmenšia miera korelácie medzi intenzitami susedných obrazových prvkov,
- čo najvyššia miera entropie informačného zdroja reprezentovaného intenzitami obrazových prvkov.

Pre objektívne ohodnotenie každej z týchto vlastností existuje viacero numerických parametrov, ktoré využívajú buď zašifrovaný obraz, alebo porovnávajú dvojicu obrazov – otvorený obraz (pôvodný, nešifrovaný obraz) a zašifrovaný obraz. Pri šifrovaní obrazov sa používa kľúč, ktorý má nasledujúce vlastnosti:

- dĺžka kľúča závisí od konkrétneho OŠA, najčastejšie sa používajú dĺžky kľúča 128 alebo 256 bitov,
- zápis kľúča by mal byť čo najjednoduchší pre možnosť opakovania experimentov,
- jednoduchosť zápisu kľúča by nemala negatívne vplyvať na jeho štatistické vlastnosti (napr. kľúč, v ktorom sa opakuje malý počet znakov nie je vhodný pre praktické aplikácie).

V našom predošlom výskume sme pri experimentoch používali kľúče, ktoré vznikli binárnym rozvojom konštánt ako napr. π [2]–[4]. Podobné postupnosti sú súčasťou databázy OEIS (On-Line Encyclopedia of Integer Sequences), ktorá bola vytvorená a je stále rozširovaná komunitou výskumníkov.

V tomto príspevku sa chceme zamerať na výber vhodných kandidátov z databázy postupností OEIS pre účely štatistického testovania OŠA. Tieto kľúče by mali spĺňať všetky vyššie uvedené požiadavky.

II. POSTUPNOSTI Z DATABÁZY OEIS

História dnešnej databázy OEIS sa dá dopátrať do 60. rokov minulého storočia, keď N. Sloane začal zhromažďovať postupnosti celých čísel pôvodne pre jeho záverečnú prácu v oblasti kombinatoriky [5]. V roku 1973 bola vydaná prvá knižka obsahujúca 2 372 postupností [6]. Rozšírené vydanie s 5 488 postupnosťami bolo publikované v roku 1995 [7]. S rozvojom internetu, ktorý

umožnil jednoduchšiu komunikáciu medzi výskumníkmi na celom svete sa N. Sloane rozhodol existujúcu databázu postupností uverejniť aj na internete, čím vznikla OEIS v podobe, v akej ju poznáme dnes [8]. To spôsobilo rýchly nárast počtu postupností v databáze a umožnilo použitie postupností v nových aplikáciách. Už v roku 1996 počet záznamov v OEIS prekonal hodnotu 16 000 [5]. V súčasnosti (máj 2023) počet záznamov v databáze presahuje hodnotu 360 000 [8].

Záznamy v databáze OEIS majú viacero polí, pričom niektoré z nich sú voliteľné. Ku najdôležitejším poliam z pohľadu tohto príspevku patria:

- identifikátor záznamu (aktuálne tvorený písmenom A a šesťmiestnym číslom),
- hodnoty prvkov postupnosti (označené aj ako dáta),
- vzorec na výpočet prvkov postupnosti,
- počet prvkov postupnosti, resp. rozsah vstupných parametrov, ak ich postupnosť vyžaduje,
- komentár ku vlastnostiam postupnosti,
- stav záznamu (platnosť záznamu môže byť zrušená napr. pri zistení chýb pri výpočte prvkov postupnosti).

Pomerne jednoduchým príkladom postupnosti je postupnosť A000045, ktorá obsahuje Fibonacciho čísla. Prvky tejto postupnosti sa počítajú predpisom (1):

$$F_n = F_{n-1} + F_{n-2} \quad \text{pre } n = 2, 3, 4, \dots \quad (1)$$

pričom $F_0 = 0$ a $F_1 = 1$.

Prvých desať prvkov postupnosti A000045 ilustruje Tabuľka 1. Rovnako ako v prípade dokumentácie k tejto postupnosti [9], aj v tomto príspevku sa namiesto označovania prvkov postupnosti zápisom A000045(0), A000045(1), A000045(2), atď. používa substitúcia $a(0)$, $a(1)$, $a(2)$, atď.

Tabuľka 1
Prvých desať prvkov postupnosti A000045 z databázy OEIS [9]

a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
0	1	1	2	3	5	8	13	21	34

Z prvkov postupnosti A000045 v Tabuľke 1 si môžeme všimnúť viacero možných problémov pri použití tejto postupnosti ako šifrovacieho kľúča pre OŠA. V prvom rade by bolo potrebné prvky postupnosti previesť do binárnej sústavy. Ak by vopred nebol určený počet bitov pre jednotlivé prvky postupnosti, mohlo by to spôsobiť rozdielne interpretácie postupnosti A000045 v binárnej sústave. Takisto by bolo potrebné preveriť štatistické vlastnosti výslednej postupnosti.

A. Metodológia výberu kandidátskych postupností

Kvôli možnosti opakovania experimentov, pre ktoré je potrebný stále rovnaký kľúč sme sa rozhodli pri výbere kandidátskych postupností z databázy OEIS postupovať takto:

- 1) prehľadávať iba binárne postupnosti, ktoré nie sú ovplyvnené rôznymi prevodmi, resp. interpretáciami (výhodné je, že priamo v komentári záznamu sú uvedené tiež ich štatistické vlastnosti),
- 2) z nájdených postupností vybrať iba tie s dostatočným počtom prvkov (pre súčasné OŠA aspoň 256 bitov),
- 3) kvôli prehľadnosti zápisu vyradiť postupnosti, ktoré na generovanie hodnôt ich prvkov využívajú iné postupnosti.

Nájdené postupnosti, ktoré vyhovujú uvedeným pravidlám sú zhrnuté v Tabuľke 2. Tieto kandidátske postupnosti na experimentálne kľúče pre OŠA je potrebné ďalej bližšie analyzovať.

B. Analýza kandidátskych postupností

V niektorých prípadoch je pre vytvorenie tzv. najhoršieho scenára (angl. worst-case scenario) pre šifrovanie pomocou OŠA vhodné používať kľúče s výraznými nedostatkami ako napr. periodicitou časti kľúča. Vo všeobecnosti by však mali použité kľúče poskytnúť reálny náhľad na vlastnosti OŠA. Preto sme kandidátske postupnosti ďalej analyzovali.

1) *Binárne rozvoje:* Binárny rozvoj sa dá laicky vysvetliť ako aproximácia čísla pomocou určitého počtu bitov. Týmto bitom prislúchajú váhy vyjadrené mocninami čísla 2. Po vyjadrení celočíselnej časti čísla kladnými mocninami čísla 2 nasleduje vyjadrenie desatinnej časti čísla zápornými mocninami čísla 2. Napr. decimálne číslo $(0,625)_{10} = 0 \cdot 2^0 + 1 \cdot 2^{-1} + 0 \cdot 2^{-2} + 1 \cdot 2^{-3}$ sa dá binárnym rozvojom bez znamienka zapísať ako $(0,101)_2$.

Binárne rozvoje matematických konštánt ako π , e , odmocnín prvočísel, či logaritmov určitých čísel sa kvôli vlastnostiam použitých čísel javia ako vhodné postupnosti pre použitie v úlohe kľúča pre OŠA.

Tabuľka 2
Kandidátske postupnosti z databázy OEIS

OEIS identifikátor	stručný popis postupnosti	hexadecimálny zápis prvých 128 bitov
A004539	binárny rozvoj $\sqrt{2}$	0×B5 04 F3 33 F9 DE 64 84
A004547	binárny rozvoj $\sqrt{3}$	0×DD B3 D7 42 C2 65 53 9D
A004555	binárny rozvoj $\sqrt{5}$	0×8F 1B BC DC BF A5 3E 0A
A004569	binárny rozvoj $\sqrt{7}$	0×A9 53 FD 4E 97 C7 4D BC
A004585	binárny rozvoj $\sqrt{10}$	0×CA 62 C1 D6 D2 DA 94 90
A004593	binárny rozvoj e	0×AD F8 54 58 A2 BB 4A 9A
A004601	binárny rozvoj π	0×C9 0F DA A2 21 68 C2 34
A004609	binárny rozvoj $\sqrt{6}$	0×9C C4 70 A0 49 09 73 E8
A010060	Thueova-Morseova postupnosť, začína bitom 0 a vzniká pridávaním doplnku existujúcej postupnosti na jej záver ($0 \rightarrow 01 \rightarrow 0110 \rightarrow \dots$)	0×69 96 96 69 96 69 69 96
A014577	doplňky bitov, ktoré sa pri binárnom zápise čísel $n = 1, 2, 3, \dots$ nachádzajú o 1 pozíciu vľavo od najmenej významného bitu s hodnotou 1	0×D9 C9 D8 C9 D9 C8 D8 C9
A059448	parita nulových bitov pri binárnom zápise čísel $1, 2, 3, \dots$	0×4D 2C D3 2D 2C D2 D3 2C
A068426	binárny rozvoj $\log_e 2$	0×B1 72 17 F7 D1 CF 79 AB
A068431	binárny rozvoj $\sqrt{\pi}$	0×E2 DF C4 8D A7 7B 55 3C
A068433	binárny rozvoj $\log_e 3$	0×8C 9F 53 D5 68 18 54 BB
A068434	binárny rozvoj $\log_e 5$	0×CE 02 0F BF 6C 69 9B 57
A091862	nadobúda hodnotu 0 ak pri sume binárneho zápisu exponentov prvočíselného rozkladu čísel $1, 2, 3, \dots$ nastane prenos medzi rádmí; inak nadobúda hodnotu 1	0×FB B9 F2 BB 08 BB D8 2B
A104338	binárny rozvoj Catalanovej konštanty $G = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^2}$	0×EA 7C B8 9F 40 9A E8 45
A261178	binárny rozvoj e^γ ; Eulerova-Mascheroniho konštanty $\gamma = \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{k} - \log_e n$	0×E3 FA 2E 55 89 AB 5F 8A
A270742	binárny rozvoj $\prod_{n=1}^{20000} \frac{2^{n-1}}{2^n}$	0×49 EE 04 3F DF 00 20 7F
A341389	prvky postupnosti $a_n, n = 0, 1, 2, 3, \dots$ podľa: $a(n) = \begin{cases} 0 & \text{pre } n = 0 \\ 1 - a(\lfloor \frac{n}{2} \rfloor) & \text{pre nepárne } n \\ a(\lfloor \frac{n}{2} \rfloor) & \text{pre párne } n > 0 \end{cases}$ kde $\lfloor b \rfloor$ zaokrúhli b na najbližšie menšie celé číslo	0×5A 5A A5 A5 5A 5A A5 A5
A342892	doplňky bitov, ktoré sa pri binárnom zápise čísel $n = 0, 1, 2, \dots$ nachádzajú o 2 pozície vľavo od najmenej významného bitu s hodnotou 1	0×FA D8 F2 D0 FA 58 F2 50

2) *Thueova-Morseova postupnosť (A010060) a postupnosť A341389*: Táto dvojica postupností počíta hodnoty nasledujúcich prvkov s využitím hodnôt predošlých prvkov. Aj keď sú v prípade niektorých matematických zobrazení ako napr. logistické zobrazenie tieto vlastnosti výhodné [10], pri týchto dvoch postupnostiach majú hodnoty prvkov postupností veľmi krátku periódu, resp. hexadecimálny zápis postupností (viď Tabuľku 2) obsahuje len dva znaky.

Táto dvojica postupností sa teda javí vhodná iba na otestovanie vlastností OŠA v tzv. najhorších scenároch. Tu by sa dala zaradiť aj tzv. nulová postupnosť obsahujúca samé nuly (v OEIS označená ako A000004).

3) *Postupnosti používajúce doplnky bitov (A014577 a A342892) a paritu bitov (A059448)*: Pri postupnostiach, ktoré obsahujú doplnky bitov pre jednotlivé rády binárneho čísla, ktoré vzniklo prevodom z decimálnych čísel sa môžu hodnoty doplnkov po určitom počte prvkov opakovať. Jednoduchým príkladom sú decimálne čísla $(5)_{10}$, $(9)_{10}$, $(13)_{10}$ a $(17)_{10}$, ktorých binárne reprezentácie sú $(101)_2$, $(1001)_2$, $(1101)_2$ a $(10001)_2$. Doplnky bitov, ktoré sú o jeden rád vľavo od najmenej významnej jednotky (postupnosť A014577) sú v tomto príklade stále rovné jednotke. Postupnosť A342892 je odlišná len tým, že uvádza doplnky bitov, ktoré sú o dva rády vľavo od najmenej významnej jednotky. Podobný problém nastáva aj pri postupnosti A059448, ktorá počíta paritu nulových bitov v binárnej reprezentácii decimálnych čísel.

Takisto ako v prípade postupností A010060 a A341389, aj táto trojica postupností je vhodná iba pre preverenie vlastností OŠA v tzv. najhorších scenároch.

4) *Postupnosť A091862*: Táto postupnosť je medzi kandidátskymi postupnosťami uvedenými v Tabuľke 2 špecifická, na čo poukazuje aj jej pomerne zložitá definícia. Hodnoty prvkov postupnosti sú značne ovplyvnené počtom členov v prvočíselnom rozklade decimálnych čísel. V prípade, ak prvočíselný rozklad obsahuje len jeden člen s jedným exponentom, tak pri sume exponentov nemá ako nastať prenos medzi rádmi a hodnota prvku postupnosti je rovná jednotke. Prvok s hodnotou nula sa prvýkrát vyskytuje pri prvočíselnom rozklade čísla $6 = 2^1 \cdot 3^1$, kde binárny súčet exponentov 1 a 1 vyžaduje prenos medzi rádmi $(1 + 1 = 10)_2$.

Z tohto dôvodu sa zo začiatku postupnosti A091862 vyskytujú pomerne dlhé monotónne sekvencie, kde sa hodnoty prvkov postupnosti nemenia (napr. pre decimálne čísla 1 až 5, 7 až 9, 11 až 13, 16 až 20, atď.). Aj keď na túto skutočnosť nepoukazuje komentár k postupnosti, obmedzuje to použitie postupnosti v úlohe šifrovacieho kľúča pre OŠA. Pre overenie možnosti použitia časti tejto postupnosti by bolo potrebné vykonať hlbšiu analýzu, avšak výber časti postupnosti by skomplikoval jednoduchosť zápisu postupnosti (pri zvyšných kandidátskych postupnostiach stačí len identifikátor postupnosti).

III. ZÁVER

V tomto príspevku bola stručne predstavená databáza celočíselných postupností OEIS so zreteľom na možnosť použitia záznamov databázy ako kľúčov pre OŠA. Po opise základných vlastností záznamov databázy a príklade jedného z nich boli bližšie analyzované záznamy, ktoré sa vďaka ich vhodným vlastnostiam označili ako kandidátske. Tieto záznamy boli ďalej analyzované a následne boli v príspevku uvedené odporúčania ohľadom možného použitia týchto postupností pri šifrovaní pomocou OŠA.

V budúcnosti by sme sa chceli zamerať na ešte podrobnejšiu analýzu postupností, ktoré boli vytvorené z binárnych rozvojov. V databáze OEIS sa nachádzajú najmä binárne rozvoje matematických konštánt alebo prvočísel. Preto je možné uvažovať aj o analýze binárnych rozvojov iných čísel, pre ktoré doteraz neboli vytvorené záznamy v databáze OEIS. Tento výskum by mohol spôsobiť aj vytvorenie databáz experimentálnych kľúčov pre testovanie OŠA, ktoré by mohli využívať viacerí výskumníci, podobne ako je to v prípade databáz experimentálnych obrazov.

POĎAKOVANIE

Tento príspevok bol podporený výskumnými projektami APVV-17-0208 „Odolné mobilné siete na doručovanie obsahu“ a VEGA 1/0260/23 „Umelá inteligencia, robustné MANET a viacpreškové D2D siete integrované do siete 6. generácie“.

LITERATÚRA

- [1] J. Fridrich, Symmetric Ciphers Based on Two-Dimensional Chaotic Maps, *Intl. J. of Bifurcation and Chaos*, 1998, roč. 8, č. 6, s. 1259–1284. ISSN: 0218-1274. DOI: 10.1142/S021812749800098X.
- [2] J. Oravec, L. Ovseník, J. Papaj, An Image Encryption Algorithm with a Plaintext-related Quantisation Scheme, *IET Image Processing*, 2021, roč. 15, č. 9, s. 2039–2055. ISSN: 1751-9667. DOI: 10.1049/ipr2.12174.
- [3] J. Oravec, L. Ovseník, S. Andrejčík, An Image Encryption Algorithm with an Efficient Quantization Scheme, v *Proc. of 32nd Intl. Conf. Radioelektronika 2022*. Košice (Slovensko), 2022, s. 1–6. ISBN: 978-1-7281-8686-3. DOI: 10.1109/RADIOELEKTRONIKA54537.2022.9764931.
- [4] J. Oravec, L. Ovseník, M. Lapčák a kol., An Image Encryption Algorithm Suitable for Medical Images, v *Proc. of 16th Intl. Conf. Informatics 2022*. Poprad (Slovensko), 2022, s. 1–6. ISBN: 979-8-3503-1034-4. DOI: 10.1109/Informatics57926.2022.10083313.
- [5] J. M. Borwein, Adventures with the OEIS, v *Proc. of ALLADI60 2016: Analytic Number Theory, Modular Forms and q-Hypergeometric Series*. Gainesville (USA), 2016, s. 123–138. DOI: 10.1007/978-3-319-68376-8_9.
- [6] N. J. A. Sloane, *A Handbook of Integer Sequences*. New York: Academic Press, 1973, 206 s. ISBN: 978-01-2648-550-9.
- [7] N. J. A. Sloane, S. Plouffe, *The Encyclopedia of Integer Sequences*. New York: Academic Press, 1995, 587 s. ISBN: 978-01-2558-630-6. DOI: 10.1016/C2013-0-11509-7.
- [8] The On-line Encyclopedia of Integer Sequences [online]. Citované 17.5.2023. Dostupné na internete: <https://oeis.org>.
- [9] OEIS A000045: Fibonacci numbers [online]. Citované 17.5.2023. Dostupné na internete: <https://oeis.org/A000045>.
- [10] J. Oravec, L. Ovseník, J. Turán a kol., 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.

Hardvérový návrh a realizácia laboratória sieťovej steganografie

Jakub ORAVEC, Ľuboš OVSENÍK, Samuel ANDREJČÍK

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

{jakub.oravec, lubos.ovsenik, samuel.andrejcik}@tuke.sk

Abstrakt – V tomto príspevku bude opísaná experimentálna zostava, ktorá má slúžiť na výskum v oblasti sieťovej steganografie. V súčasnosti sú vlastnosti takýchto techník testované najmä v simulačných prostrediach, no skúsenosti z minulosti v oblasti obrazovej steganografie ukazujú na nutnosť praktického overenia teoretických záverov. Na takúto úlohu je však potrebné mať pripravený hardvér. Príspevok okrem zoznamu samotného hardvéru obsahuje aj niektoré poznatky získané pri realizácii návrhu, ktoré môžu byť cenné aj pre iné aplikácie podobných zariadení.

KLúčové slová – Internet vecí, Raspberry Pi, relé, sieťová steganografia

I. ÚVOD

Steganografia je vedná disciplína, ktorá sa zaoberá ukrývaním prenosu správ pomocou ich vkladania do zdanlivo nenápadnej komunikácie. O existencii vložených správ vie iba odosielateľ a príjemca, ostatní používatelia by v ideálnom prípade nemali vedieť určiť prítomnosť správ ani analýzou prenášanej komunikácie [1].

Z histórie sú známe príklady steganografických techník zo starovekého Grécka, kde sa na zaznamenávanie odkazov používali drevené dosky pokryté vrstvou vosku. Do vosku boli vyryté jednotlivé znaky správy. Ak sa však vosk pôsobením tepla nechal roztopiť, tak sa správa dala vyryť priamo do drevenej dosky. Doska sa následne prekryla voskom, do ktorého sa vyryla iná správa. Neznalí poslavia si mohli prečítať správu vyrytú do vosku, ale adresát ho mohol roztopiť, odstrániť z dosky a prečítať si správu vyrytú do dreva [2].

Neskôr sa steganografia rozvíjala najmä kvôli rôznym konfliktom a vojnám. Po rozšírení počítačov a digitálnej reprezentácie správ začali vznikať nové steganografické techniky, ktoré využívali najmä redundanciu (nadbytočnosť) pri zápise dát. Preto sa ako krycie médium začali vo veľkom používať obrázky a pre ne boli vyvíjané steganografické algoritmy ako napr. F5, JSteg alebo Outguess [2].

S rozvojom obrazovej steganografie súvisí aj postupný vznik obrazovej stegoanalýzy, ktorá sa snaží o štatistickú analýzu obrazov a posúdenie, či takéto obrazy sú používané na ukrytý prenos správ [3]. Vylepšovanie techník obrazovej stegoanalýzy spôsobilo ďalší vývoj algoritmov obrazovej steganografie, ktoré začali na ukrývanie správ využívať komplikovanejší matematický aparát. Neskôr sa pri niektorých navrhnutých algoritmoch namiesto experimentálneho overenia ich vlastností prešlo na vyjadrenie ich vlastností pomocou matematických dôkazov.

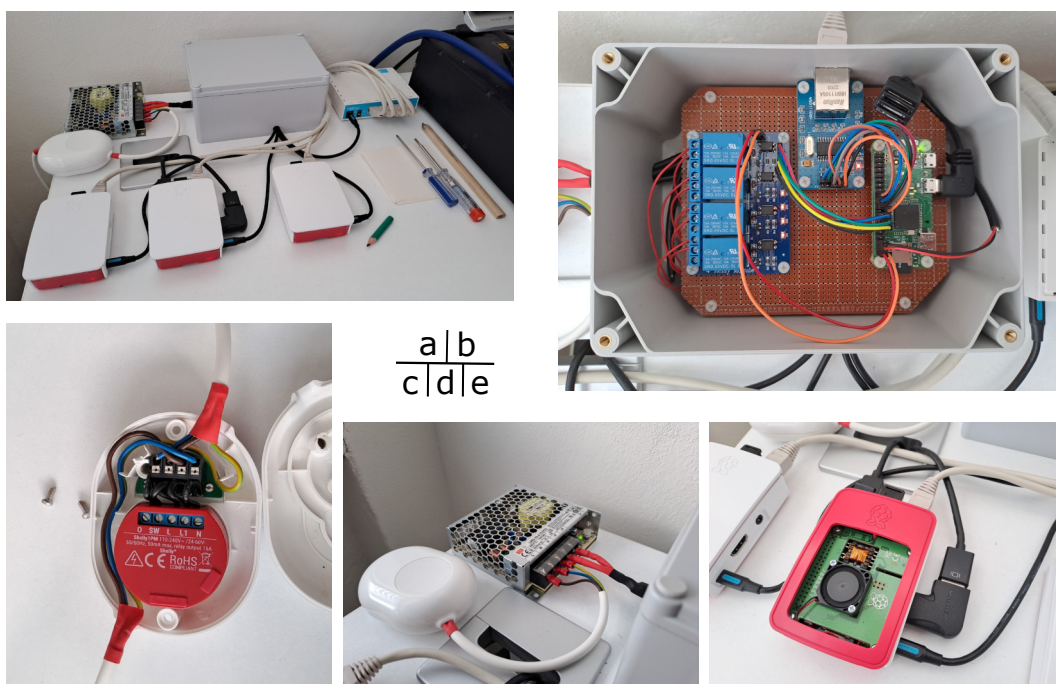
Po roku 2000 začali vznikať aj prvé techniky tzv. sieťovej steganografie [4]–[7]. Tieto techniky sa snažia o využitie prázdnych polí v hlavičkách protokolov, resp. o modifikáciu používaných polí takým spôsobom, aby to negatívne neovplyvnilo používateľov koncových zariadení (napr. zvýšením oneskorenia). Overenie vlastností týchto techník sa však väčšinou robí len teoreticky v rôznych simulačných nástrojoch.

Ako sa ukázalo v minulosti pri algoritmoch obrazovej steganografie, pri praktických aplikáciách sa môžu prejavovať aj iné ako teoreticky predpovedané problémy (napr. nízka kapacita pre ukrývanie správ) [1]. Preto je vhodné navrhnuté riešenia overovať aj prakticky. To je možné vykonať rôznymi spôsobmi – napr. vytvorením viacerých virtuálnych počítačov alebo použitím viacerých reálnych zariadení.

II. HARDVÉROVÝ NÁVRH LABORATÓRIA

Pri hardvérovom návrhu laboratória sieťovej steganografie sme sa snažili do najväčšej možnej miery použiť zariadenia, ktoré boli v našom výskumnom kolektíve uvoľnené po skončení rôznych záverečných prác, resp. projektov. Laboratórium obsahuje nasledujúce zariadenia:

- 3 kusy jednodoskového počítača Raspberry Pi 3 B+ (revízia 1.3),
- dosky s ventilátormi pre napájanie počítačov Raspberry Pi 3 B+ pomocou rozhrania Ethernet (angl. Power over Ethernet, PoE),
- 1 kus jednodoskového počítača Raspberry Pi Zero WH (revízia 1.1),
- doska s čipom ENC28J60 s rozhraniami Ethernet a SPI (angl. Serial Peripheral Interface),
- doska s štvorkanálovým relé spínaná pomocou 3,3 V DC signálov,
- routerboard MikroTik RB941-2nD (hAP lite),
- napájací zdroj Mean Well LRS-75-5 (výstupné napätie 5 V DC, maximálny výstupný prúd 14 A) so svorkovnicou,
- relé Shelly 1PM so spínaním pomocou webového rozhrania s púzdrom a mechanickým tlačidlom,
- pamäťové karty a rôzne druhy káblov.



Obr. 1: a) celkový pohľad na laboratórium, b) krabica s relé doskou, c) relé Shelly 1PM v púzdre, d) napájací zdroj a relé s tlačidlom, e) Raspberry Pi 3 B+ s PoE doskou

V budúcnosti je možné uvažovať o náhrade niektorých použitých zariadení ich novšími verziami. Vo všeobecnosti by to malo byť výhodné z hľadiska vyššieho výkonu novších zariadení, a teda skrátenia času potrebného na experimenty, ale treba hľadiť aj na nevýhody ako napr. na vyššiu spotrebu, nutnosť efektívnejšieho chladenia, či potrebu investícií pri výmene častí zariadení. Pri výmene počítačov za novšiu verziu Raspberry Pi 4 B by bolo nutné vymeniť ich krabičky kvôli výmene pozície USB rozhraní s rozhraním Ethernet a takisto aj použité napájacie káble (počítače Raspberry Pi 4 B používajú rozhranie USB-C).

Návrh laboratória sa v súčasnosti nedá považovať za finálny. Ucelená dokumentácia laboratória zahŕňajúca blokovú schému zapojených zariadení a sieťovú topológiu laboratória spolu s miestnosťou, v ktorej sa laboratórium nachádza bude vypracovaná až pri spustení laboratória do riadnej prevádzky, kedy by mali byť vyriešené už aj rôzne drobné nedostatky, ktoré sa doteraz neprejavili.

III. REALIZÁCIA NÁVRHU A ZISTENÉ POZNATKY

A. Počítače Raspberry Pi 3 B+ a rozširujúce dosky

Pre zjednodušenie ovládania celého laboratória sa ako koncové zariadenia použili počítače Raspberry Pi 3 B+, na ktorých beží operačný systém Raspberry Pi OS. Platforma Raspberry Pi je pre našu aplikáciu výhodná aj z hľadiska veľkej podpory komunitou, rozsiahlej dokumentácie [8], pomernej nízkej spotreby a natívnej podpory nástrojov na vzdialené pripojenie VNC Viewer, resp. VNC Server [9].

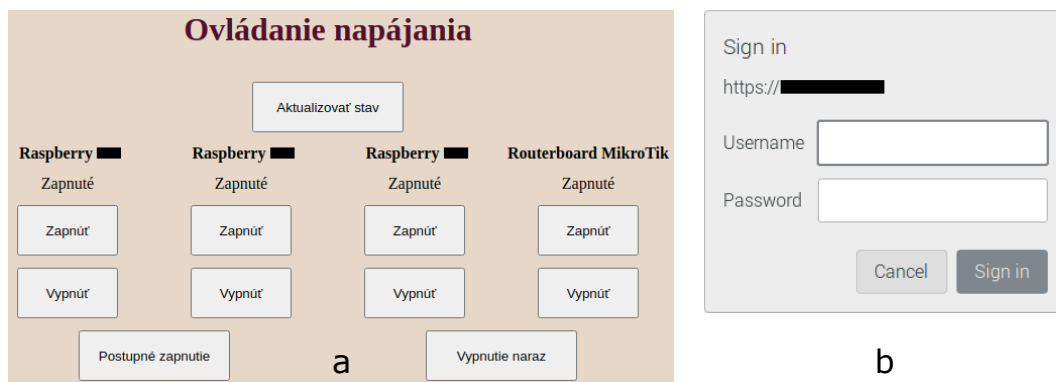
Nevýhodou verzie týchto počítačov, ktoré boli k dispozícii je však pomerne nízka veľkosť operačnej pamäte – 1 GB. V prípade, ak by boli počítačom dávkovo zadané väčšie množstvo úloh (napr. vytvoriť IP pakety a čakať na ich odoslanie), mohlo by dôjsť k vyčerpaniu tejto pamäte. Preto bol na počítače nainštalovaný skript *zram-swap* [10], ktorý priebežne komprimuje aktuálne nevyužívané segmenty operačnej pamäte.

V prípade, ak je potrebné čítať dáta z komprimovaného segmentu pamäte, je potrebné tieto dáta dekomprimovať, čo môže mierne zaťažiť procesor počítača. Pre istotu boli počítače Raspberry Pi 3 B+ vybavené rozširujúcimi doskami, ktoré umožňujú napájanie cez Ethernet, ale v našom prípade sa používajú iba kvôli ich viacrychlostnému ventilátoru (vid' Obr. 1 e)). Aj pri plnej záťaži počítačov sa ich maximálna teplota pohybuje okolo 65 °C, čo je vyhovujúca hodnota pre všetky súčasti počítača (najcitlivejší sa javí byť čip LAN7515 pracujúci s rozhraniami Ethernet a USB, ktorý garantuje správnu funkciu pri teplotách do 70 °C).

B. Počítač Raspberry Pi Zero WH a rozširujúce dosky

Tento počítač sa nachádza spolu s Ethernet a relé doskou v sivej krabici (vid' Obr. 1 a), b)). Hlavnou úlohou tejto zostavy je riadiť napájanie ďalších zariadení v laboratóriu. Pri prvých pokusoch o zostavenie laboratória boli použité rôzne smerovače, z ktorých niektoré pri veľkej záťaži prestali reagovať (laicky povedané zamrzli). Bez použitia skriptu *zram-swap* sa niekedy stalo, že prestal reagovať aj niektorý z počítačov Raspberry Pi 3 B+. Pre umožnenie vypínania a zapínania týchto zariadení na diaľku je ich napájanie riešené pomocou štvorkanálovej relé dosky (3 kanály pre 3 počítače a zvyšný kanál pre routerboard).

Relé doska je pripojená ku počítaču Raspberry Pi Zero WH (vid' Obr. 1 b)). Signály pre jednotlivé relé sú vysielané zo štyroch GPIO (angl. General-Purpose Input/Output) pinov. Používatelia ovládajú napät'ové úrovne pomocou grafického rozhrania – webovej stránky bežiaccej na lokálnom Apache2 servri (vid' Obr. 2 a)) [11]. Táto stránka si pri prihlásení vyžaduje autentifikáciu pomocou používateľského mena a hesla, čo ukazuje aj Obr. 2 b).



Obr. 2: a) grafické rozhranie pre riadenie napájania, b) vyskakovacie okno pre autentifikáciu

S riadením napájania počítačov Raspberry Pi 3 B+ súvisí ešte jedna skutočnosť. Pre zjednodušenie práce s laboratóriom je vhodné vopred si pripraviť skripty, ktoré budú vykonávať rôzne úlohy. Ak by takéto skripty boli na všetkých troch počítačoch, mohlo by to spôsobovať problémy s ich aktualizáciou. Preto sa v laboratóriu určil jeden počítač ako „hlavný“ a na ňom bol spustený Samba server (názov podľa protokolu SMB) na zdieľanie súborov [12].

Aby sa pri spustení mohli zvyšné dva počítače Raspberry Pi 3 B+ pripojiť (angl. mount) na tento Samba server, je potrebné „hlavný“ počítač a routerboard spúšťať po privedení napájania o niečo skôr ako zvyšné dva počítače. Na to slúži Skript 1, ktorý je spúšťaný utilitou *crontab* po spustení počítača Raspberry Pi Zero WH:

Skript 1: Skript spustený po privedení, resp. obnovení napájania

```
#!/bin/bash

#spustenie routerboardu a hlavneho Raspberry
raspi-gpio set 23,27 op dl

#oneskorenie
sleep 30

#spustenie zvyсных Raspberry
raspi-gpio set 17,22 op dl
```

Pre vylepšenie konektivity počítača Raspberry Pi Zero WH, ktorý štandardne umožňuje iba bezdrôtové pripojenie do siete je k nemu rozhraním SPI pripojená doska s čipom ENC28J60, ktorý ovláda rozhranie Ethernet. Pre účel riadenia napájania sa pôvodne uvažovalo aj s počítačom Raspberry Pi 3 B, resp. 3 B+, avšak kvôli priestorovým obmedzeniam, nižšej spotrebe a nižšej prevádzkovej teplote bez potreby prídavného chladenia sa pristúpilo k realizácii s počítačom Raspberry Pi Zero WH.

C. Routerboard MikroTik RB941-2nD

Úlohou tohto zariadenia je poskytnúť pre „hlavný“ počítač Raspberry Pi 3 B+ viacero rozhraní Ethernet, keďže tento počítač disponuje iba jedným takýmto rozhraním, ktoré je použité pre pripojenie všetkých zariadení k vonkajšej sieti. Routerboard MikroTik RB941-2nD hardvérovo umožňuje smerovanie, avšak v laboratóriu je softvérovo nastavený iba ako prepínač (angl. switch). Tým pádom všetko smerovanie závisí od „hlavného“ počítača, čo pri vhodnom nastavení záznamov smerovacej tabuľky umožňuje jednoduchšie odchyťovanie (angl. sniffing) prenášaných dát pri vykonávaných experimentoch. Na rozhraniach tohto routerboardu sú pripojené zvyšné dva počítače Raspberry Pi 3 B+ a posledné rozhranie slúži ako „rezerva“ pri nečakaných udalostiach, keď by sa na routerboard a teoreticky aj do experimentálnej siete bolo potrebné pripojiť zo vzdialenej siete.

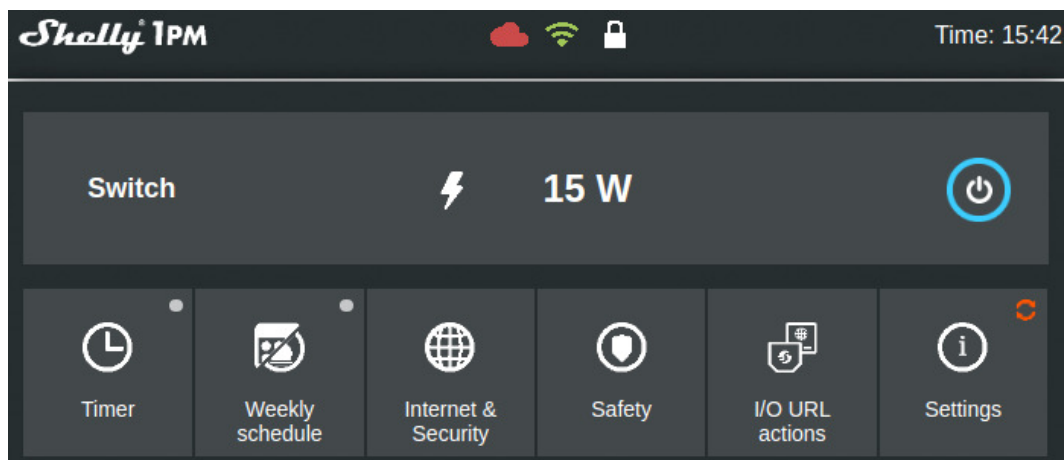
Výhodou routerboardu MikroTik RB941-2nD je aj skutočnosť, že je napájaný napätím 5 V DC pomocou microUSB rozhrania rovnako ako počítače Raspberry Pi. Vďaka tomu môžu byť všetky zariadenia napájané jedným zdrojom a toto napájanie môže byť riadené jednou relé doskou.

D. Napájací zdroj Mean Well LRS-75-5 a relé Shelly 1PM

Pre napájanie počítačov a routerboardu bol použitý napájací zdroj Mean Well LRS-75-5. Výhodou tohto zdroja je jeho tvrdosť, výrobca deklaruje maximálny pokles napätia $\pm 2\%$ aj pri maximálnej záťaži [13], čo znamená minimálne výstupné napätie $5 \cdot 0,98 = 4,9$ V DC. To bolo problémom pri rôznych nabíjačkách pre telefóny, kde počítače Raspberry Pi 3 B+ indikovali podpätie (hranica pre upozornenie je 4,7 V). Pri nevhodnom napájaní podľa výrobcu počítačov Raspberry Pi Foundation môže časom dôjsť ku zníženiu životnosti zariadení [8].

Zdroj sa pripája na vstupné napätie 230 V AC 50 Hz a je schopný poskytnúť maximálny výstupný prúd až 14 A (teda výkon cca $5 \cdot 14 = 70$ W), čo je pre účely laboratória dostatočná hodnota. Pri doterajších experimentoch nebola prekročená hodnota okamžitého výkonu 25 W.

Na elektrickej šnúre, ktorá privádza napätie na zdroj je pripojené púzdro s mechanickým tlačidlom obsahujúce relé Shelly 1PM (vid' Obr. 1 c), d), ktoré spína fázu. Toto riešenie umožňuje v prípade potreby odpojenie napájania celého laboratória. Relé Shelly 1PM navyše disponuje webovým rozhraním s meraním spotreby, zobrazením okamžitej spotreby a ďalšími nastaveniami (vid' Obr. 3). Na otvorenie webového rozhrania je potrebná autentifikácia používateľa prostredníctvom mena a hesla. Na relé sa dá pripojiť z bezdrôtovej siete, ktorá má prístupový bod (angl. access point) v miestnosti, kde sa nachádza predstavené laboratórium. Vďaka iným servrom, ktoré sa nachádzajú v tejto miestnosti sa tak na diaľku dá riadiť napájanie celého predstaveného laboratória.

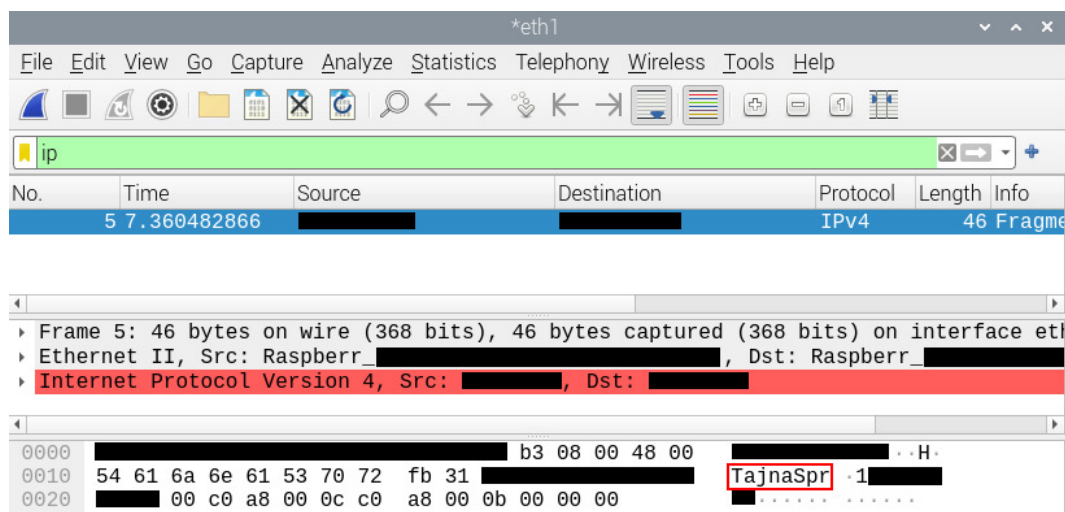


Obr. 3: Ukážka webového rozhrania relé Shelly 1PM

IV. ZÁVER

Tento príspevok bol zameraný na predstavenie experimentálneho laboratória, ktoré má umožňovať praktický výskum v oblasti sieťovej steganografie. V súčasnosti (máj 2023) je laboratórium po hardvérovej stránke funkčné a pripravené na použitie, avšak je potrebné vytvoriť vhodné programové prostriedky pre zjednodušenie experimentov. S tým súvisí viacero problémov, ako napr. správne nastavenie práv pre skupiny používateľov (napr. administrátor a bežní používatelia – študenti). Po vyladení všetkých detailov by malo laboratórium výrazne zjednodušiť vývoj rôznych algoritmov sieťovej steganografie a výskum ich vlastností napr. pre ambiciózných študentov druhého stupňa štúdia.

Pri tvorbe programových prostriedkov je v pláne využívať programovací jazyk Python, knižnicu Scapy a rôzne programy, resp. utility na odchyťovanie prenášaných dát ako Wireshark alebo *tcpdump*. Malá ukážka schopností týchto nástrojov je demonštrovaná na Obr. 4.



Obr. 4: Odchytenie paketu vytvoreného knižnicou Scapy pomocou programu Wireshark

POĎAKOVANIE

Tento príspevok bol podporený výskumnými projektami APVV-17-0208 „Odolné mobilné siete na doručovanie obsahu“ a VEGA 1/0260/23 „Umelá inteligencia, robustné MANET a viacpre-skokové D2D siete integrované do siete 6. generácie“.

LITERATÚRA

- [1] J. Fridrich, *Steganography in Digital Media: Principles, Algorithms and Applications*. Cambridge: University Press, 2010, 466 s. ISBN: 978-05-2119-019-0. DOI: 10.1017/CBO9781139192903.
- [2] D. Levický, *Multimédiá a ochrana ich obsahu*. Košice: Elfa, 2012, 249 s. ISBN: 978-80-8086-199-5.
- [3] B. Li, J. He, J. Huang a kol., A Survey on Image Steganography and Steganalysis, *J. of Information Hiding and Multimedia Signal Processing*, 2011, roč. 2, č. 2, s. 142–172. ISSN: 2073-4212. DOI: 10.1201/B12697-11.
- [4] K. Szczypiorski, HICCUPS: Hidden Communication System for Corrupted Networks, v *Proc. of Intl. Multi-Conference on Advanced Computer Systems*. Międzyzdroje (Poľsko), 2003, s. 31–40.
- [5] W. Mazurczyk, M. Smolarczyk, K. Szczypiorski, Retransmission Steganography and its Detection, *Soft Computing*, 2011, roč. 15, č. 3, s. 505–515. ISSN: 1433-7479. DOI: 10.1007/s00500-009-0530-1.
- [6] J. Lubacz, W. Mazurczyk, K. Szczypiorski, Principles and Overview of Network Steganography, *IEEE Communications Magazine*, 2014, roč. 52, č. 5, s. 225–229. ISSN: 1558-1896. DOI: 10.1109/MCOM.2014.6815916.
- [7] P. Bedi, A. Dua, Network Steganography using the Overflow Field of Timestamp Option in an IPv4 Packet, *Procedia Computer Science*, 2020, roč. 10, č. 171, s. 1810–1818. ISSN: 1877-0509. DOI: 10.1016/j.procs.2020.04.194.
- [8] Raspberry Pi Documentation: The Official Documentation for Raspberry Pi Computers and Microcontrollers [online]. Citované 19.5.2023. Dostupné na internete: <https://www.raspberrypi.com/documentation/>.
- [9] RealVNC Product Documentation [online]. Citované 19.5.2023. Dostupné na internete: <https://help.realvnc.com/hc/en-us/categories/360000301637-Documentation>.
- [10] zram-swap: A Simple Zram Swap Script for Modern Systemd Linux [online]. Citované 19.5.2023. Dostupné na internete: <https://github.com/foundObjects/zram-swap>.
- [11] Apache HTTP Server Version 2.4 Documentation [online]. Citované 19.5.2023. Dostupné na internete: <https://httpd.apache.org/docs/2.4/>.
- [12] SambaWiki [online]. Citované 19.5.2023. Dostupné na internete: https://wiki.samba.org/index.php/Main_Page.
- [13] Mean Well LRS-75-5 Datasheet [online]. Citované 19.5.2023. Dostupné na internete: <https://www.meanwell-web.com/content/files/pdfs/productPdfs/MW/LRS-75/LRS-75-spec.pdf>.

Detection of Events in Data from DIMS Experiment

¹Stanislav LEVENDOVSKÝ, ²Michal VRÁBEL

¹Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

²Fundamental Research Department, National Centre for Nuclear Research, Poland

¹stanislav.levendovsky@student.tuke.sk, ²michal.vrabel@ncbj.gov.pl

Abstract – This work deals with the development and evaluation of methods for the detection and classification of meteors and other objects in the sky. The dataset is constructed from the data of the DIMS Experiment. We test the effect of various dataset transformations on classification accuracy and recall, using the Hough transform, convolutional neural networks, and transfer learning. The methods are also compared with regard to time complexity, due to the limited computational resources in the production. From the results, we evaluate the Hough Transform method to be fast but less accurate. The selected neural networks show better results but are slower using the CPU.

Keywords – convolutional neural networks, DIMS Experiment, Hough transform, image classification

I. INTRODUCTION

Every day, a large number of meteors pass through the Earth's atmosphere and are observed and captured by various detectors and cameras. By analyzing detected meteors, it is then possible to determine their velocities and trajectories. By recording and investigating the time, frequency, and magnitude of flying meteors, we can also learn more about their history and origin which helps in the subsequent studying of the dynamics of the Solar System.

With the rapidly increasing amount of data being collected, the classification of samples is becoming more and more exhausting, inaccurate, and time inefficient if it has to be done manually. With the improvement of techniques in computer vision and deep learning in recent years, it is possible to eliminate the human factor when classifying moving objects from the sky.

The complex process of information discovery and knowledge mining from meteors flying in the sky consists of several chained parts. From video capture and object detection, subsequent processing of the captured objects into a suitable form, classifying the detected objects using different algorithms, to reconstructing the meteors - determining their direction of arrival, velocity, and magnitude, and finally estimating their orbits. In this work, we only deal with the processing of the data into a suitable form for classification algorithms, and implementation and evaluation of the algorithms for meteors classification.

II. DIMS EXPERIMENT

The *Dark Matter and Interstellar Meteoroid Study* Experiment (DIMS)[1], is an international collaborative project of several countries, through their research centers or universities. The main goal of the project is to search for and detect various forms of macroscopic dark matter, or hypothetical cosmic particles that may be components of macroscopic dark matter. One of these hypothetical particles are called *nuclearites* which are believed to reach velocities greater than 250km s^{-1} . In addition, the experiment also aims to search for and detect interstellar meteors that reach velocity greater than 72km s^{-1} [2], as opposed to lower velocities of the meteors. In general, DIMS should be able to track and search for fast-moving particles and objects (events), that appear in the field of view of the detector, which uses high-sensitivity CMOS cameras[3].

There are currently a number of DIMS sky imaging camera modules deployed around the world. Two modules have been deployed from 2021 to 2022 at Japan's Akeno and Kiso observatories. Currently, DIMS modules are deployed in Utah, US. They have been installed near the detectors of the *Telescope Array*[4] experiment, at two locations approximately 17km apart.

A multi-component camera system is currently used to monitor and image the starry sky. In addition to a CMOS camera, capturing the sky at 29.97 or 59.94 frames per second at the 1080p resolution, the system also contains a computer connected to various sensors, GPS receiver, a cooling system, the possibility of connecting to a computer remotely, etc. Events recorded using software *UFOCapture* are stored into the harddrive in a form of multiple outputs:

- Video of the recorded event in raw *avi* format without compression.
- Metadata of the recorded event in *XML* format containing the time and length of the event, latitude, longitude, and altitude of the camera module, country and time zone, camera resolution, camera lens settings, frames per second, etc.
- A composite of a recorded event in *jpg* format, created using the brightest-pixel of a given sequence of video frames (at each pixel of the video).
- A bitmap created from a composite of the video in *bmp* format, that distinguishes the pixels of the moving events and the ambient static noise using different color channels.

III. RELATED WORK

The analyzed articles were selected based on the similarity of the task type - classifying and filtering objects from images or videos. In addition, we selected papers that specifically investigated meteor objects.

In 2015, Siladi, E. et al. [5] used a relatively small sample of 400 data samples (178 meteors and 222 non-meteors) for meteor classification. The hidden layer with 60 neurons achieved an accuracy of 91%. Support vector machine achieved slightly worse accuracy, 89%. The authors rate the result of 10% false positives as unsuitable for real-world implementation.

In 2017, De Cicco, M. et al. [6] seek to extract information from meteors, chipped from comets, to predict the orbits of potentially dangerous comets flying relatively close to Earth. The dataset size of 200 000 samples contained only about 3% meteor samples; the rest was noise, aircraft, clouds, birds, etc. The data consisted primarily of temporal information, describing the trajectory and light curves of given observed objects. Random forest achieved a precision of 90% and a recall of 81%. The paper also describes the automation of the process of classifying samples from CAMS data. A convolutional neural network (CNN) with a depth of 7 layers (5 convolutional and 2 fully connected) achieved a precision of 88.3% and a recall of 90.3%.

In 2018, Galindo, Y. and Lorena, Ana C. [7] tested the use of transfer learning. The dataset of 1000 meteors and 660 other objects in the sky was augmented, but more data did not improve the testing accuracy. They tested *ResNet* architecture with 18, 34, 50, and 101 layers. Leaving the first 6 convolutional layers still frozen, a various number of later convolutional layers were unfrozen during fine-tuning. Pre-trained models increased the testing accuracy by 11 to 18%. The best accuracy was around 96%.

In 2019, Marsola TC. et al. [8] used a dataset containing only 400 images of stellar sky objects, that has been augmented using existing data. They used an input size of 224x224 pixels, to feed the CNN architecture *VGG16* pre-trained on *ImageNet* dataset. Three additional fully connected layers (32, 16, and 8 neurons) were connected to the network and all VGG16 layers remained frozen during the training and fine-tuning. The VGG16 achieved an average accuracy of 84.35% as opposed to 79.89% obtained from a control CNN consisting of 6 layers.

Gural P. S. [9] also used CAMS data to classify meteors. A well-balanced dataset contained 100 000 meteors and 100 000 non-meteor samples. The data were transformed to input sizes of 94x94, 64x64, or 46x46 pixels respectively for the CNN, which they called *MeteorNet*. A manual inspection of the dataset by experts revealed almost 2000 mislabeled data samples. The recall of the classification of the samples by the experts is around 99%, compared to a recall of more than 99.9% produced by the *MeteorNet*. From this fact, authors infer that a properly tuned neural network may be more reliable than the human element in the classification of data samples.

In 2020, Cecil, D. and Campbell-Brown, M. [10] used software *MeteorScan* to capture events from the sky. The dataset contained 56 230 manually labeled samples, 40% of which were meteors. They augmented the dataset by rotating the existing samples to obtain 4-times more samples. The Hough transform was used to preprocess the event images by tracing the bright objects in the image. They used 10% of the data for final testing with an input size of 128x128 pixels. They tested 18 models with wide range of hyperparameters and different numbers of layers. The best CNN model had 6 layers (4 convolutional and 2 fully connected) and achieved an accuracy of 99.8%. It was also comprised of only 41 946 training parameters and took just over an hour to train. The other tested CNNs also achieved a good accuracy from 99.2% to 99.4%.

The most used approach by the authors of analyzed papers is the use of convolutional neural networks. Based on the differences in the data sources, their sizes, and overall shape in comparison to our data, we cannot directly replicate the aforementioned approaches. However, we will use the insights as a basis for forming a suitable solution for our problem.

IV. DATA PREPROCESSING

We were provided with an unlabeled dataset consisting of 3300 data samples with an input size of 192x192 pixels. Each data sample was in the form of a 4D numpy array (sequence of 3D numpy arrays, 3rd dimension representing colors of the pixels). Firstly we developed a script to convert the 4D numpy arrays to the brightest-pixel composites (fig. 1) so we could manually label all of the samples. Then we had to remove invalid samples from the dataset because they were caused by human error and are too atypical to take into consideration.

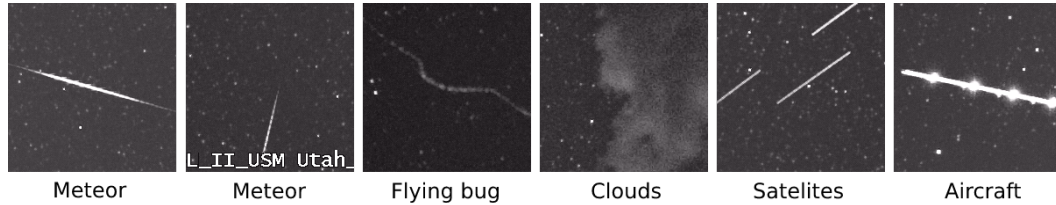


Fig. 1 Example of data samples composites using the brightest-pixel method.

After labeling the whole dataset and examining composites, we have discovered that only about 15% of data samples represent the non-meteors category. A script was developed to address the imbalanced dataset problem by creating non-meteors data samples from raw video recordings. Same input size (192x192 px) random cutouts were used to select areas from source raw video recordings. The random cutouts were validated using a bitmap, mentioned in II so that there are no collisions between them and between captured events from the video. We balanced the dataset in terms of the ratio of category sizes, but we have to keep in mind that more than 80% of the non-meteors are represented by static noise in the form of stars or clouds.

We also needed to find a way to remove text located in the lower 25px of some of the data samples (fig. 1 - second sample from left). The text can be turned off in the UFOCapture software but it was decided that it is important for the DIMS Experiment to identify video recordings. The first developed method for text removal can be applied to an already created brightest-pixel composite. The second method can be used to remove text from all the frames in a 4D numpy array sequence. Using the second method, it is possible to create any type of composite, i.e., a composite that displays one of the axis pixels in time. Both methods work by counting white and black pixels in the lower part of the composite (the second method creates a brightest-pixel composite at runtime) and deciding whether to apply text removal or not according to a threshold value i.e., the count of the detected black and white pixels. Then the dominant composite color is calculated. The first method uses colors from the interval $\langle d - \frac{n}{2}, d + \frac{n}{2} \rangle$, where d is the dominant color and n is the number of adjacent colors chosen, to replace the pixels representing white with a black outline. The second method uses colors from interval $\langle 25, d \rangle$, because the individual frames of the event are darker than the brightest-pixel composite created. Multiple replacement colors are used to blend the colors better as opposed to a single color.

There is currently a noise in the background of the composites, which is caused by the CMOS sensor settings and by light scattered in the atmosphere, etc. Static stars are a second factor that complicates meteor detection. Leaving the noise in could cause a high complexity in the image, e.g., unwanted line detection in places that are made up of adjacent stars arranged in a line. Likewise, when generating composites of axis pixels in time, unwanted horizontal lines emerge from the stars, creating noise. The *OpenCV* library offers function *fastNIMeansDenoising*, but it is not sufficient for us as it reduces the intensity of bright objects and smoothes objects, removing important details from the image. The function cannot completely remove larger stars with higher magnitudes either, it only shrinks them. A noise reduction method was developed that can only be applied to the 4D numpy array sequence. It first divides the sequence of 3D frames into n parts, according to the specified size (number of frames) of one part. Then, in each part, the average color value of each pixel within the sequence is computed and put into a 3D array of average color values. Next, the 3D array of average color values is subtracted from each frame of that part. This step removes the noise and static objects within the parts, leaving only the moving objects. Finally, all the parts are merged into a single part, recreating the 4D numpy array sequence.

V. HOUGH TRANSFORM

According to Leavers V. F. in paper [11] the Hough transform (HT) is used to detect edges or lines, circles, or even in general to find any shapes in an image. Using the HT, we can tell how many lines are in a given trajectory of an object in the sky, and what the detected lines are in terms of the direction of motion, e.g., whether it is a meteor or not.

The *OpenCV* library offers the probabilistic function *HoughLinesP*[12]. It can only be applied to the input in the form of an image, henceforth we used the brightest-pixel composites. We have developed a classifier *HoughTransformClassifier* (HTC) that works with composite with an input size of 192x192px. The classifier supports grid search, which is a form of exhaustive search for the best combinations of hyperparameters used in training¹. By hyperparameters, we mean parameters of the *Canny* edge detection (two thresholds), and the subsequent *HoughLinesP* for line detection (threshold, minimum line length, and maximum allowed gap between points on the same line), applied to data samples. In addition, hyperparameters are also the thresholds that decide the classification of the sample into one of the categories according to conditions (1) and (2). The HTC has been implemented with support for cross-validation and the use of multiple threads, for faster training. Classifier supports various evaluation metrics such as accuracy, recall, and F1 score. It is also possible to export the best hyperparameters produced by grid search.

Based on the character of several transformations (e.g., text removal and noise filtering applied to a sequence of images prior to composite creation) that must be applied to a specific data form, we chose to apply each transformation during HTC training and testing. This approach of dynamically applying the transformations avoids the need to store the different forms of datasets on disk for different transformations applied. It is possible to use the composite image directly as input data, as well as a 4D numpy array that is transformed into a composite at a runtime. We also need to find a suitable combination of transformations from the following options:

- 1) Removal of the text at the bottom of the image (from composite or 4D numpy array).
- 2) Application of *GaussianBlur*.
- 3) Noise removal (from a composite using the *fastNlMeansDenoising* function or from 4D numpy array).
- 4) Edge detection using *Canny* function.
- 5) Morphological operations *dilate* and *erode* to highlight possible object tracks.
- 6) Apply HT using the *HoughLinesP* function to detect lines.

The first meteor detection method will be performed based on the number of detected lines using the condition (1). If the number of lines n_lines is zero or it exceeds a certain threshold t , the sample will be evaluated as a non-meteor.

$$0 < n_lines \leq t \tag{1}$$

When the brightest-pixel composite is created, the temporal information of the event is lost, and we cannot determine the velocity of the detected object². The importance of the line size is lost as a fast meteor flying across the entire image dimension, may have as long a line as a slow satellite. We solve this problem using the second HT method using 2 composites created by that displaying pixels of both of the axis in time. For both composites, the average angle with the x-axis is calculated using the detected lines. Then these two average angles from the two composites are added up (parameter *meteor_angles*), and if they satisfy the condition (2), we classify the object as a meteor.

$$t < meteor_angles < 175^\circ \wedge meteor_angles \neq 90^\circ \tag{2}$$

VI. DEEP NEURAL NETWORKS

We augmented the original dataset and divided the data into 3 parts. First, we divide the data of 5500 samples into 2 parts with a ratio of 60% training data and 40% test data (or 70% to 30%). Approximately 2200 (or 1650) samples are used for the final testing. The remaining 60% (or 70%) of the data are augmented n -times using various transformations. Augmented data are further divided into training and validation data in the ratio of 80% and 20%.

Then we evaluate the impact of augmentation on the testing accuracy using 5-fold cross-validation and a simple CNN with 3 convolutional and 1 fully connected layer. We also test the text removal and noise filtering methods applied to the 4D numpy array. We will also test the effect of image resizing on classification using the *area* interpolation method. The original input size of 192x192 was reduced to 128x128, 96x96, and 64x64 pixels. Using *KerasTuner*³, we try to find more optimal hyperparameters for the CNN architecture, as shown in the table 1.

Another approach we use for meteor classification is transfer learning. The pre-trained models have initialized weights that are obtained from training on *ImageNet* datasets. We have a relatively

¹What we mean by training in the context of the HTC is just fitting the data to the best combination of hyperparameters. The HTC does not contain parameters that are implicitly modified during program execution like in neural networks.

²Another problem with the brightest-pixel composite is that low-magnitude meteors may not be detected due to low brightness, or that the meteors flying in the direction of the CMOS detector appear as a small dot, instead of a line.

³A hyperparameter optimization framework for hyperparameter searching.

Description		Values range	Best parameters
Architecture	CNN	3 - 7 layers	6 (5conv+1dense)
Data augmentation		No, 3n - 6n	5n
Input type		Brightest-pixel composite	Brightest-pixel composite
Transformations		Text removal, Noise reduction	Text removal, Noise reduction
Pooling layers		Max(2x2), Average(2x2)	Max(2x2) - First, Average(2x2) - Hidden
Cost function		Binary cross-entropy	Binary cross-entropy
Optimization alg.		Adam	Adam
Batch size		32 - 64	32
Batch normalization		No, Yes	Yes
Learning rate		0,0001 - 0,1	0,001
Training		25, 50 epochs	25
Early stopping		5, 10 epochs	5
First layer	Input size (px)	64x64 - 192x192	192x192
Conv2D	Filters (size, n)	3x3 - 11x11, 16-64	3x3, 48
	Activation function	ReLU, Tanh	ReLU
Hidden layers	Filters (size, n)	3x3 - 11x11, 16-128	5x5/9x9, 32-128
Conv2D	Activation function	ReLU, Tanh	ReLU
Fully connected	Neurons	32-384	256
Dense	Activation function	ReLU, Tanh	ReLU
Output layer	2 classes	Sigmoid,	
	Activation function	Softmax	Sigmoid

Table 1
 Hyperparameter range to find the best CNN.

small dataset of meteors and non-meteors, but it is very different from the datasets on which the models were pre-trained. For completely different data, we will need to unfreeze several layers to adapt to the character of our data during training. Testing time on the CPU is crucial for us since the samples will be evaluated directly in DIMS modules that do not contain a dedicated GPU. Based on the low test time, we chose MobileNet. We also chose VGG16 (and VGG19 for comparison) and ResNet50V2, based on analyzed existing solutions.

After convolutional layers of pre-trained models, we added one fully connected layer with 256 neurons and a ReLU activation function. We first adapted the models to this layer by training them on 10 epochs using frozen weights. We then further fine-tuned the models by reducing the training rate of the Adam optimization algorithm by a factor of 10, unfroze several convolutional layers, and trained the model on another 10 epochs. In all cases, we used the final model for testing from weights of epoch with the lowest validation error.

VII. RESULTS

We have tested 16 different permutations of transformations using the lines count HTC method. Testing accuracy without applying any transformations was 90%. By applying text removal we have achieved about 95% testing accuracy. By applying both text removal and noise reduction, we were able to boost the accuracy to 97%. The more transformation we added, the bigger chance there was to lose the important information. When we changed the training metric to recall, testing results were worse, but the testing recall achieved about 99.7% in the best scenario. We could definitely improve the recall by increasing the threshold value and using a wider range of parameters in a grid search. The first method's weakness is that it is not capable of distinguishing meteors from satellites and aircrafts. The second method using angles that represent the speed of the moving object partially solves this problem. We tested 6 different permutations. The testing accuracy results were slightly better, using text removal and noise reduction we were able to achieve 97,3% as opposed to 89% with no transformations used.

The HTC testing speeds on the CPU depend on the complexity of the data. The less noise in the images, the faster classifier gets. Text removal methods applied on composite and 4D numpy array can process about 100 and 15 samples per second respectively using a single thread. The noise removal method can handle about 15 samples per second on a single thread. The line count method, without applying text removal and noise reduction methods, can process about 170, and the angular method 110 samples per second on a single thread. After applying transformations the speed boosts up to 1179 and 748 processed samples per second.

Using CNNs and transfer learning, we were able to achieve both accuracy and recall greater than 99% in almost all cases (using 5-fold cross-validation, 70:30 train test ratio, 192x192 input size, and 5n augmentation) as shown in the table 2. The application of text removal and noise reduction improved overall results. The best accuracy of 99.5% and recall of 99.57% on average was achieved by the VGG16. It also achieved more than 99% accuracy and recall on all lower input

sizes tested. It is difficult to directly compare the results with analyzed papers as the datasets differ and some of them do not provide information about the character of the non-meteors category objects. Ours non-meteors category objects are imbalanced, most of them being noise.

CNN model	Trainable params	Depth	Test accuracy	Test recall	Test time (6 CPU threads)
KerasTuner	0,94 m.	6	99,15 ± 0,19%	98,82 ± 0,24%	110 samples/sec
VGG16	14,84 m.	16	99,48 ± 0,28%	99,52 ± 0,36%	22 samples/sec
VGG19	20,15 m.	19	99,30 ± 0,27%	99,28 ± 0,36%	18 samples/sec
MobileNet	3,88 m.	28	99,09 ± 0,48%	99,23 ± 0,41%	197 samples/sec
ResNet50V2	24,09 m.	50	98,54 ± 0,24%	98,47 ± 0,36%	63 samples/sec

Table 2
 Results of the chosen CNN models.

CNN depth and the number of trainable parameters are not necessarily indicators of the neural network speed. CNNs are much slower than HTC as the test times are shown for 6 CPU threads as opposed to a single thread in HTC earlier.

It is important to note that we used an imbalanced dataset, that contained too few non-meteor objects (that were not pure static noise consisting of stars or clouds) for us to be able to evaluate the results as significant and the algorithms as optimal. Therefore it is necessary to extend the dataset and enrich it with different categories of objects. This could be done by applying the HTC on the cutouts of UFOCapture data to get the moving objects and classify them manually.

VIII. CONCLUSION

In this paper we have developed the *HoughTransformClassifier* that may be generally suitable for classifying images (not only meteors in this work) using the number of lines detected or angles between the detected lines in the image. The HTC is relatively fast, depending on the character of the input, being able to achieve a recall of more than 99,7%. Therefore, it is important to emphasize image pre-processing and noise reduction. The developed methods for text removal and noise reduction are slow but generally reusable to filter out text and static objects from moving objects respectively. The pre-trained VGG16 model produced the best results on the accuracy metric, averaging 99,5%, 99,76% at best case. Compared to the HTC, it is slower but more accurate.

HTC, trained on recall and a broader range of parameters using grid search, can be used as an initial filter to filter out the static noise. Then one of the CNN models would be applied to classify samples classified as moving objects. The ideal candidate seems to be the VGG16, adapted to our task, as it achieved the best accuracy result among the tested models. An alternative option is to deploy a different detection trigger on the online camera recording, which would capture moving objects. These could then be directly evaluated using the CNN model without the need of using HTC as an initial filter.

REFERENCES

- [1] S. Abe, D. Barghini, S. Valenti, M. Arahori, M. Bertaina, M. Casolino, A. Cellino, C. Covault, T. Ebisuzaki, Y. Fujiwara *et al.*, “Dims experiment for dark matter and interstellar meteoroid study,” 2021.
- [2] M. Hajdukova, V. Sterken, P. Wiegert, and L. Kornoš, “The challenge of identifying interstellar meteors,” *Planetary and Space Science*, vol. 192, p. 105060, 2020.
- [3] F. Kajino, I. Ide, R. Ide, Y. Tameda, K. Shinozaki, M. Bertaina, A. Cellino, M. Casolino, T. Ebisuzaki, Y. Takizawa *et al.*, “Study for moving nuclearites and interstellar meteoroids using high sensitivity cmos camera,” in *36th International Cosmic Ray Conference (ICRC2019)*, vol. 36, 2019, p. 525.
- [4] T. Abu-Zayyad, R. Aida, M. Allen, R. Anderson, R. Azuma, E. Barcikowski, J. Belz, D. Bergman, S. Blake, R. Cady *et al.*, “The surface detector array of the telescope array experiment,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 689, pp. 87–97, 2012.
- [5] E. Siladi, D. Vida, and E. K. Nyarko, “Video meteor detection filtering using soft computing methods,” in *Proceedings of the International Meteor Conference, Mistelbach, Austria*, 2015, pp. 27–30.
- [6] M. De Cicco, S. Zoghbi, A. P. Stapper, A. J. Ordonez, J. Collison, P. S. Gural, S. Ganju, J.-L. Galache, and P. Jenniskens, “Artificial intelligence techniques applied to automating meteor validation and trajectory quality control to direct the search for long-period comets,” 2017.
- [7] Y. Galindo and A. C. Lorena, “Deep transfer learning for meteor detection,” in *Anais do XV Encontro Nacional de Inteligência Artificial e Computacional*. SBC, 2018, pp. 528–537.
- [8] T. Marsola, A. C. Lorena, S. Itáu Unibanco, and S. Aeronautics, “Meteor detection using deep convolutional neural networks,” *Anais do Simpósio Brasileiro de Automação Inteligente*, vol. 1, p. 104260, 2019.
- [9] P. S. Gural, “Deep learning algorithms applied to the classification of video meteor detections,” *Monthly Notices of the Royal Astronomical Society*, vol. 489, no. 4, pp. 5109–5118, 2019.
- [10] D. Cecil and M. Campbell-Brown, “The application of convolutional neural networks to the automation of a meteor detection pipeline,” *Planetary and Space Science*, vol. 186, p. 104920, 2020.
- [11] V. F. Leavers, *Shape detection in computer vision using the Hough transform*. Springer, 1992, vol. 1.
- [12] I. Macdonald, “Probabilistic hough transform,” *Notes*, pp. 1–4, 2016.

Cold atmospheric plasma treatment of substrate for heating resistive film

¹Tomáš KMEC

¹ Department of Technologies in Electronics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

¹tomas.kmec@student.tuke.sk

Abstract — To improve stainless steel plates surface properties, cold atmospheric plasma treatment of these plates was performed. The plates are used as a substrate for the insulation layer of a heating resistive film. The cold plasma treatment results in the modification of the plates wettability, as indicated by the contact angle between a droplet of water and the steel surface. This modified wettability is then compared with the wettability of a non-modified steel surface with an extra adhesion layer. To perform this comparison, accurate wettability tests using an inert gas atmosphere were conducted.

Keywords — cold atmospheric plasma, steel, contact angle, resistive film, insulation layer

I. INTRODUCTION

Stainless steel is widely used in various industries due to its suitable mechanical, thermal, chemical, and magnetic properties and cost effectiveness. In electronics, especially in appliances using water heating elements, stainless steel is used as a base substrate for heating resistive films enabling heating running cold water e.g. in dishwashers or washing machines. A ceramic coating (Al_2O_3) is generally applied on the steel surface to insulate it from a heating resistive film (NiCr). Defects in the adhesion of this insulation ceramic layer result in electrical breakdowns and damage the heating resistive element. Typical factors that negatively impact the adhesion include organic and inorganic contamination, such as chemically unstable thin layers (10 – 15 Å) of metal oxides and hydroxides that cover stainless steel exposed to air, as well as grease and oils that remain on the steel surface after production steps [2]. Therefore, it is important to have optimal surface cleanness and wettability prior to deposition of the insulation ceramic layer.

While surface modification techniques, including plasma treatment, have been extensively studied, plasma treatment specifically for stainless steel remain until today poorly documented [3].

Plasma is an ionized gas consisting of positively charged ions, negatively charged electrons, and neutral particles. When plasma comes into contact with a particular surface, it alters the surface physicochemical properties, such as its morphology and chemistry. These changes lead to an increase in the free surface energy of the surface that has been treated with plasma and impact the surface wettability [4].

In the tested application of heating resistive elements, insufficient adhesion of the insulation ceramic layer on the steel substrate is solved by the addition of an extra adhesion layer (NiCr). By improving the adhesion of the insulation layer on the steel substrate through the plasma treatment (with a maximum treatment temperature of 50 °C), it may be possible to reduce the need for an extra adhesion layer and thus simplify the whole production process.

II. EXPERIMENTAL PROCEDURE

A. Materials and samples preparation

The wettability experiment was conducted using commercially available steel plates that were laser cut into 8 small samples measuring 20 mm x 30 mm with a thickness of 1.2 mm. 4 of these samples underwent abrasive grit blasting to remove the surface oxidation layer mechanically (Fig. 1). The both blasted (4) and non-blasted (4) steel samples were then degreased using isopropyl alcohol. The non-blasted samples, not intended for the cold atmospheric plasma treatment, were next coated by a 10 µm NiCr adhesion layer to be used as reference samples. The

NiCr adhesion layer is today used in the production process of water heating elements to set the proper adhesion of the Al_2O_3 ceramic layer on the steel surface.

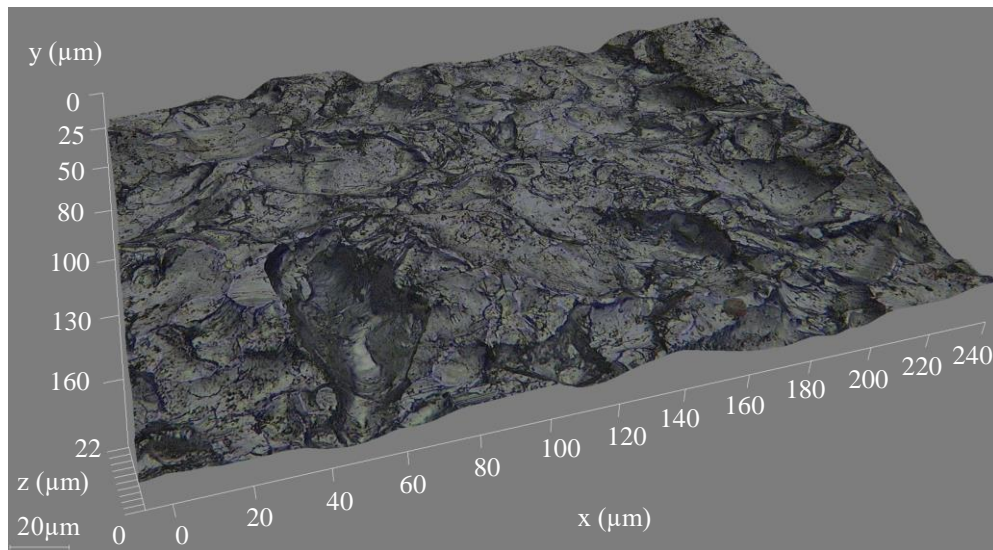


Fig. 1 3D optical analysis of the steel surface roughness after the grit blasting. The optical analysis shows peaks and valleys thus confirming roughness of the surface.

B. Cold atmospheric plasma treatment

An atmospheric plasma torch used in this study had the following setup: voltage of 20 kV, frequency of 1 kHz and an output power of 500 W. A corona discharge with an applied time of 30 s and 120 s was used. A corona discharge, in general, is generated by applying high voltage between two or more sharp electrodes, typically in the form of a needle or a thin wire. The ionization process creates a crown around the active electrode. The used torch included a needle net with a 1 cm mesh [5]. The principle of the applied setup is depicted in Fig. 2. To evaluate the influence of oxygen in the atmosphere (related to level of surface oxidation), a corona discharge was tested in a nitrogen atmosphere with a flow rate of 300 liters per hour, samples are listed in Tab. 1

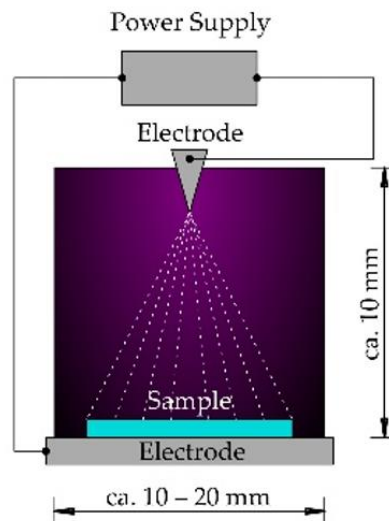


Fig. 2 The principle of corona discharge used in atmospheric plasma torch [5].

C. Contact angle test

Contact angle measurement is a non-destructive method used to determine the wetting behavior of liquids on a solid surface. The contact angle data provide information about the surface tensions and allow for the characterization and evaluation of interactions between two solids, the adhesion between coating and substrate materials, and the wetting behavior of lubricants. By using a heating chamber, measurements can be conducted at elevated temperatures, enabling the analysis of the wetting behavior of melted plastics on coated tools [6].

For the contact angle measurement, a camera system G10 from Krüss was used [7] as shown in Fig. 3. The measuring system consists of a camera, an illuminated background, a precise liquid dropper, and an adjustable stand. The contact angle measurement was used to assess wettability of the steel surface. The liquid applied on the surface was, for every single test, a droplet of clean water. The contact angle between a droplet surface and the steel surface was measured. In Fig. 4 are depicted results from the contact angle measurement. The test was performed for a non-treated and for a cold atmospheric plasma treated steel surface (5 minutes after the treatment process). The results are shown in Tab. 1.

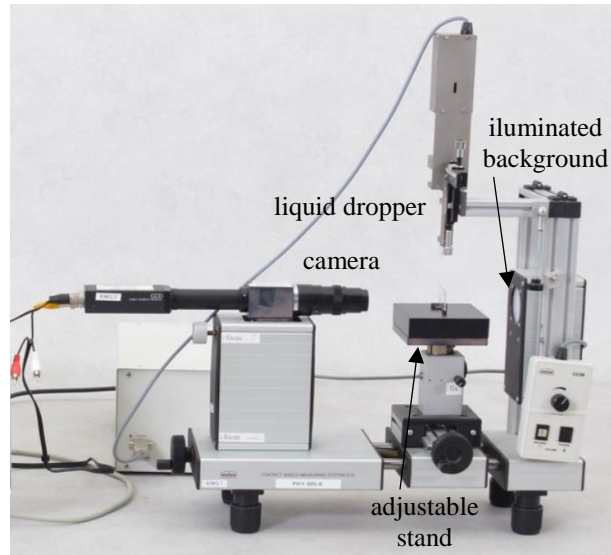


Fig. 3 The Krüss G10 camera system for the contact angle measurement consist from a camera, an illuminated background, a precise liquid dropper and an adjustable stand [7].

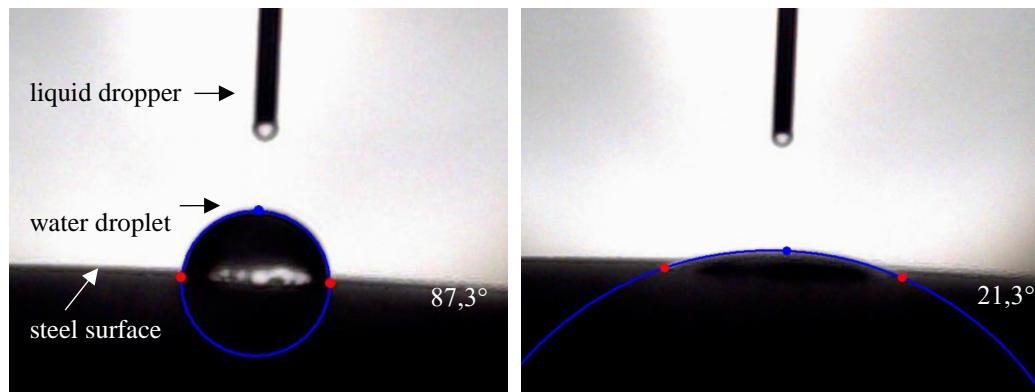


Fig. 4 The contact angle measurement with a Krüss G10 camera system. The left image shows wettability of the steel surface before the plasma treatment, the right image shows wettability after the plasma treatment. These differently treated steel surfaces show different contact angles.

Table 1
Contact angle measurement

No.	Time	Atmosphere	Sample	Initial Contact Angle (°)	Contact Angle after Treatment (°)	Variation (°)
1			Adhesion layer	27.2		
2	30 s	Air	Clear	61.5	40.2	-21.3
3	30 s	Air	Blasted	77.1	50.3	-26.8
4	30 s	Nitrogen	Clear	91.4	58.8	-32.6
5	30 s	Nitrogen	Blasted	71.7	41.9	-29.8
6	120 s	Air	Clear	76.1	22.8	-53.3
7	120 s	Air	Blasted	68.2	12.7	-55.5
8	120 s	Nitrogen	Clear	87.3	21.3	-66
9	120 s	Nitrogen	Blasted	68.5	7.6	-60.9

In Fig. 5 is shown the final arrangement of the tested application – the water heating element on the left side, with all the layers stack up in it on right side (cross sectional view).

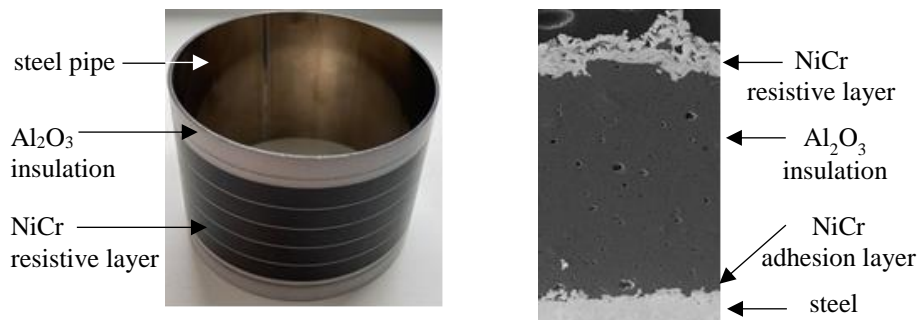


Fig. 5 The water heating element on the left side, a cross sectional view of the layers stack up on the right side

The contact angle measurements of the steel surfaces were performed before the plasma treatment and 5 minutes after the plasma treatment. The steel samples were bare or with the NiCr adhesion layer. The surrounding atmosphere during the measurements was 100 % air or 100 % nitrogen. Thus, the contact angle was measured for every sample, in air or nitrogen atmosphere, before and after the plasma treatment.

The contact angle measurements obtained are summarized in Tab. 1. The table presents the time intervals, atmospheric conditions, sample types, and corresponding contact angles. The contact angles before the plasma treatment are given as the initial values, while the change in the contact angle is in the last column.

The results in Tab. 1 reveal significant changes in the contact angles after the plasma treatment. The contact angles for both bare and blasted steel samples decreased in all cases, indicating an improvement in their surface wettability. The largest decrease in the contact angle was observed for the bare samples treated with nitrogen for 120 seconds, where the angle decreased from 87.3° to 21.3°. These findings suggest that the plasma treatment can effectively enhance the wetting properties of the samples.

These variations of the contact angle measurements were influenced by the applied time, atmosphere, and sample type.

Time effect: comparing the contact angles at different time points, it was observed that the contact angles generally decreased over time after the plasma treatment. This reduction in the contact angle indicates an improvement in the surface wettability. For instance, in the case of the bare samples treated in air, the contact angle decreased from 61.5° to 40.2° after 30 seconds of the plasma treatment. Similarly, for the bare samples treated in air, the contact angle decreased from 76.1° to 22.8° after 120 seconds of the plasma treatment. These results suggest that the longer plasma treatment has a more pronounced effect on reducing the contact angles and thus enhancing the surface wettability.

Atmosphere effect: The atmospheric conditions during the plasma treatment also played a significant role in altering the contact angles. Comparing the contact angles for the same sample type and time point in air and nitrogen atmosphere, it was observed that the nitrogen plasma treatment generally resulted in larger reductions in the contact angles compared to the air plasma treatment. For instance, the bare samples treated with nitrogen for 30 seconds exhibited the contact angle decrease from 91.4° to 58.8°, while the same samples treated with in air showed a decrease from 61.5° to 40.2°. This suggests that the chemical composition of the plasma atmosphere influences the surface properties and, consequently, the contact angles.

Sample type effect: the sample type, particularly the presence or absence of the adhesion layer, also affected the contact angles. Comparing the contact angles for the same time point and atmosphere but for different sample types, it was observed that the adhesion layer influenced the initial contact angles. The samples with the adhesion layer generally exhibited lower initial contact angles compared to those samples without the layer. However, after the plasma treatment, both sample types showed improvements in their surface wettability, resulting in the reduced contact angles. For instance, the bare samples without the adhesion layer treated with nitrogen for 120 seconds had a contact angle decrease from 87.3° to 21.3°, thus after the plasma treatment

is the contact angle lower than if used the extra NiCr adhesion layer. This suggests that the plasma treatment can effectively modify the surface properties of different sample types and give opportunity to remove the adhesion layer from the productions steps.

III. DISCUSSION

The observed variations in the contact angles highlight the complex interplay between time, atmosphere, and sample type during the plasma treatment. The decrease in the contact angles over time can be attributed to several factors. Firstly, in general longer treatments allow for more plasma energy deposition, leading to enhanced surface activation and modification. Additionally, longer treatments provide more time for plasma species to react with the surface, resulting in changes in the surface chemistry and wetting behavior.

The influence of the atmosphere on the contact angles can be attributed to the different reactive species present in nitrogen and air plasmas. Nitrogen plasma, being a chemically more active environment, is likely to induce more significant surface modifications, leading to greater reductions in the contact angles. Conversely, air plasma may have a relatively milder effect on surface properties, resulting in comparatively smaller reductions in the contact angles.

The impact of the sample type on the contact angles can be attributed to the differences in the surface morphology, roughness, and chemical composition. The presence of an adhesion layer may alter the initial contact angle due to differences in surface energy and the interactions between the liquid and solid phases.

To determine which parameter – time, atmosphere, or sample – had the most significant influence on decreasing the contact angle, a comprehensive statistical analysis would be necessary. Without specific statistical tests or data, it is challenging to provide a definitive answer. However, based on the provided information and observations, it is possible to make some general observations:

Time: from the results presented, it can be observed that increasing the time of the plasma treatment generally led to larger reductions in the contact angles. This suggests that time plays a crucial role in modifying the surface wettability. Therefore, it is reasonable to assume that the parameter of time had a considerable influence in decreasing the contact angle.

Atmosphere: the atmospheric conditions during the plasma treatment, specifically air and nitrogen presence, were found to impact the contact angles. Comparatively, the nitrogen plasma treatment resulted in larger reductions of the contact angles compared to the air plasma treatment. This indicates that the atmosphere parameter has a discernible effect on the contact angle but might be less influential compared to time.

Sample type: the presence or absence of the adhesion layer influenced the initial contact angles.

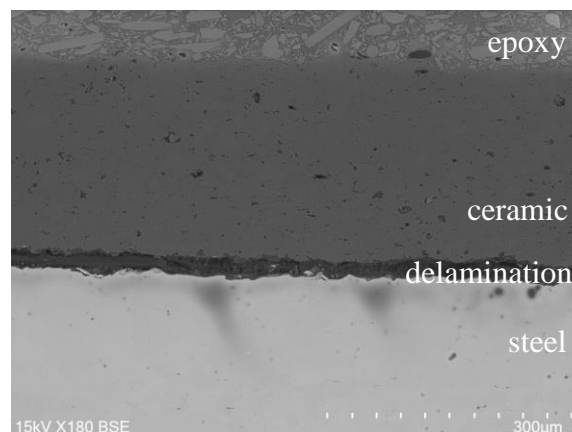


Fig. 6 Delamination of ceramic layer from steel substrate caused by insufficient adhesion.

Fig. 6 depicts an example of a delaminated ceramic coating caused by insufficient adhesion of the coating on the steel surface. The ceramic coating was sprayed directly on the steel surface without any preparation to prepare worst case sample to compare adhesion with plasma treated samples.

To prevent defects in the insulation ceramic layer coated on the steel substrate, the substrate preparation has to be appropriate. Delamination and cracks, formed in the insulation layer result in conductive traces between the heating layer and steel substrate. The resistance of water heating element insulation layer with cracks can drop from $M\Omega$ range to $k\Omega$ range even less and not passed high voltage insulation test. The plasma treatment improves the adhesion of the ceramic coating on the steel substrate. By improving wettability, it is expected to achieve higher quality insulation between the steel substrate and the resistive heating film, thus reducing delamination and cracks in the insulation layer and minimizing the occurrence of shortcuts between the heating layer and steel substrate. [8].

IV. CONCLUSION

The contact angle measurements of water droplets were performed on small stainless steel specimens before and after the cold plasma treatment in different atmospheres: air and nitrogen. To minimize the effect of steel surface oxidation, 4 steel samples underwent the grit blasting treatment. The performed measurements of the contact angles showed that the grit blasting resulted in the lower contact angles (before and after the plasma treatment) compared to the non-blasted specimens. Nitrogen atmosphere reduced the contact angles with higher variation compared to air atmosphere, however the economic feasibility of using nitrogen compared to air is debatable. The time interval of activity is the key to improve wettability with the plasma treatment. However, for industry applications it is necessary to find a proper balance between the time interval of plasma activation and the wettability improvement via production time. Based on the listed reviewed references below, the presented results show an improved wettability trend after the plasma treatment.

Improper wettability of the steel substrate is one of the key factors leading to problems with coating adhesion. By subjecting the steel surface to a 120 s plasma treatment in both (air and nitrogen) tested atmospheres, and with an appropriate surface preparation, the contact angle was reduced compared to the average contact angle of the reference surface with an extra adhesion NiCr layer. Based on these results, in general, it is expected that ceramic coating can be directly applied to steel surface after the cold atmospheric plasma treatment.

REFERENCES

- [1] J. Lin, H. Chang, "Surface modification of SUS304 stainless steel by atmospheric pressure Ar/N₂/O₂ plasma". Nuclear Instruments and Methods in Physics Research B 269. 2011. pp. 1801–1808.
- [2] A. Latifi, M. Imani, M. Khorasani, M. Joupari, "Electrochemical and chemical methods for improving surface characteristics of 316L stainless steel for biomedical applications". Surface & Coatings Technology 221. 2013. pp. 1–12.
- [3] S. Tang, O. Kwon, N. Lu, H. Choi, "Surface characteristics of AISI 304L stainless steel after an atmospheric pressure plasma treatment". Surface & Coatings Technology 195. 2005. pp. 298–306.
- [4] N. Kalel, A. Darpe, J. Bijwe, "Low pressure plasma induced surface changes of some stainless steels". Surface & Coatings Technology 425. 2021. pp. 127700.
- [5] M. Domonkos, P. Ticha, J. Trejbal, P. Demo, "Applications of Cold Atmospheric Pressure Plasma Technology in Medicine, Agriculture and Food Industry". Applied science. 2021. pp. 11, 4809.
- [6] Contact angle measurement device, [Online]. Available: <https://www.iot.rwth-aachen.de/cms/IOT/Das-Institut/Apparative-Ausstattung/Pruefung-Analytik/Topographie/~gqvac/Kontaktwinkelmessgeraet/?lid=1>
- [7] Krüss G10 contact angle measurement system, [Online]. Available: <https://www.geminibv.com/labware/kruss-g10-contact-angle-measurement-system/>
- [8] K. Schneider, V. Belashchenko, M. Dratwinski, S. Siegmann, A. Zagorsky. Thermal spraying for power generation components. ISBN: 978-3-527-31337-2. 2006. Pp. 133-138

Inovácia materiálov pre predmety priemyselnej informatiky so zameraním na manažérske informačné systémy - výukový portál

¹Samuel OKLAMČÁK, ²Tomáš TKÁČIK, ³Ján JADLOVSKÝ

^{1,2,3} Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹samuel.oklamcak@student.tuke.sk, ²tomas.tkacik@tuke.sk, ³jan.jadlovsky@tuke.sk

Abstrakt — Tento článok sa zaoberá tvorbou interaktívneho výukového portálu s využitím generátora statických webov. Cieľom portálu je poskytnúť jednoduchý spôsob tvorby a aktualizácie obsahu výukových materiálov pomocou Markdown. Výukový portál bol pilotne nasadený pre predmet Manažérske Informačné Systémy, vyučovaný na Katedre kybernetiky a umelej inteligencie. Vybraná technológia pre výukový portál nevyžaduje Content Management System ani databázu. Umožňuje tak nasadenie na ľubovoľnom webovom serveri. Taktiež opisuje aktuálne výukové materiály predmetu, inovované materiály, vytvorené prípadové štúdie a overenie poznatkov na vzorovej aplikácii.

Kľúčové slová — databázové systémy, dátový model, informačné systémy, výukový portál

I. ÚVOD

Predmet Manažérske Informačné Systémy (MIS) z portfólia predmetov priemyselnej informatiky má za cieľ oboznámiť študentov s procesom vývoja analytického informačného systému. Cieľom predmetu je, aby študenti na základe vlastnej dátovej množiny vytvorili analytický informačný systém. Dátová množina sa môže skladať z verejne dostupných dát alebo vlastných vygenerovaných dát. Zvyčajne sa jedná o dáta obsahujúce predaje tovarov alebo služieb. Transakčné dáta sa nachádzajú v každom podniku [1]. Globalizácia a potreba rýchlo sa adaptovať na zmeny vyžaduje analytické prostredie pre manažment podniku. Takýto analytický informačný systém pomáha pri dôležitých rozhodnutiach v podniku.

Prvým krokom inovácie výukových materiálov bola analýza výučby predmetu MIS. Pozorovaním a zúčastnením sa samotnej výučby sme identifikovali oblasti, ktoré možno inovovať. Analyzovali sme aktuálnu metodiku tvorby MIS a používané nástroje. Navrhli a vytvorili sme inovované výukové materiály, ktoré možno využiť na cvičeniach alebo pri vypracovaní semestrálneho zadania. Potom sme podľa požiadaviek vyučujúcich predmetu MIS navrhli a vytvorili výukový portál.

V druhej časti článku sa venujeme metodickému postupu tvorby MIS a dostupným výukovým materiálom. V tretej časti opisujeme, ako sme postupovali pri návrhu a realizácii inovácií výukových materiálov. Výukové materiály sa okrem metodického postupu zaoberajú aj teoretickými poznatkami a častými chybami študentov pri tvorbe MIS. Vo štvrtej časti sme prakticky overili inovované výukové materiály pri tvorbe vzorovej aplikácie na fiktívnej databáze predajov. V piatej časti sme zhrnuli požiadavky na výukový portál, navrhnuté riešenie, postup vývoja a samotnú ukážku výukového portálu.

II. AKTUÁLNY STAV VÝUKOVÝCH MATERIÁLOV

Na predmete MIS študenti vytvárajú dátový sklad. Najprv je potrebné vykonať analýzu dát. Je potrebné identifikovať subjekty v podniku a ich vzájomné vzťahy. Dáta prechádzajú ETL (extract, transform, load) procesom [2]. V prostredí podniku sú tieto dáta v rôznych informačných systémoch a je potrebné ich vložiť do jednej relačnej databázy. Pred samotným vkladaním dát sa vytvorí dátový model a vygenerujú sa entity z tohto modelu. Následne sa importujú dáta do databázy. Potom sa prístupí k tvorbe analytického prostredia, dimenzií, hierarchií, tabuliek

faktov, merateľných faktov a kalkulovaných faktov. Nad týmto prostredím vytvoria študenti súbor analýz, aby overili funkčnosť svojej implementácie analytického informačného systému.

Všetky výukové materiály sa nachádzajú na webstránke predmetu MIS, v sekcii Podklady k predmetu [3]. V tejto sekcii nájdú študenti vzorové vypracované zadania, podľa ktorých môžu postupovať. Podľa týchto vzorov následne študenti vytvoria dokumentáciu svojho semestrálneho zadania. Okrem vzorových zadaní sa na webe nachádzajú aj potrebné nástroje od spoločnosti Oracle. Semestrálne zadanie sa vytvára pomocou nástrojov Oracle SQL Developer, Data Modeler, Analytic Workspace Manager a OracleBI doplnku do MS Excelu. K podkladom patrí aj zoznam kľúčových pojmov potrebných pre úspešné pochopenie a vypracovanie zadania. Patria sem poznatky o relačných a multidimenzionálnych databázach. K základným pojmom multidimenzionálnych databáz patria dimenzie, fakty, úrovne, hierarchie a dátové kocky. Podklady dopĺňa aj vhodne vyšpecifikovaná literatúra. Je dôležité, aby študenti správne porozumeli pojmom ako dátový sklad, dátový trh, dátové jazero [4], hviezdicová schéma, schéma snehovej vločky a schéma súhvezdí. Všetky výukové materiály sa nachádzajú v PDF súboroch. Vzorové zadania sú do značnej miery aplikačne orientované a existujúci metodický postup možno zovšeobecniť. Pozorovaniami sme taktiež zistili stratu pozornosti študentov pri hľadaní informácií v týchto materiáloch. Aktuálne výukové materiály zároveň neobsahujú informácie, resp. návody na exportovanie, transformáciu a importovanie dát z iných produkčných systémov.

III. INOVOVANÉ VÝUKOVÉ MATERIÁLY

Po zistení problematických oblastí sme navrhli a vytvorili inovované výukové materiály. Najprv sme vytvorili interný návod v PDF verzii s metodickým a aplikačným postupom na vypracovanie semestrálneho zadania. Tento návod obsahuje všetky potrebné informácie od stiahnutia potrebných nástrojov, oboznámenia sa s prostredím, nastavením prostredia, až po vytvorenie analytického prostredia a analýzu dát viacerými spôsobmi. Tento návod nám poslužil ako základ, na ktorom sme následne budovali inovované verzie výukových materiálov. Vytvorili sme dve prípadové štúdie, ktoré prezentovali rozšírený metodický postup riešenia dielčích úloh pri tvorbe manažérskych informačných systémov. Ďalšie výukové materiály sa zaoberali novými nástrojmi od iných spoločností, ako aj od spoločnosti Oracle. K príkladom z praxe sme pridali verejne dostupnú databázu DATAcube [5] od Štatistického úradu SR.

A. Prípadová štúdia 1

V prípadovej štúdii 1 sme navrhli a vytvorili súbor univerzálnych procesov a vzorových príkladov, ktoré majú za cieľ demonštrovať proces exportu, transformácie a importu dát. Cieľom nebolo zamerať sa na konkrétne podnikové informačné systémy. Keďže každý podnikový informačný systém ukladá štruktúrované a neštruktúrované dáta do relačných, resp. dokumentových databázových systémov [1], rozhodli sme sa preto zamerať na databázové systémy. V štúdii opisujeme hlavne prácu s tabuľkovými dátami. Tabuľky sú najčastejšie používané v databázových systémoch na uloženie a prepojenie dát. Štruktúra prípadovej štúdie je znázornená v Tab. 1.

Tab. 1 Štruktúra prípadovej štúdie 1.

Názov	Import dát z produkčných systémov do RDBS Oracle
Aktuálny stav	Pri výučbe predmetu je používaný ako ukázkový import dát zo súborov typu CSV (Comma Separated Values)
Cieľ	Vytvoriť študijné materiály zaoberajúce sa importom dát z rôznych databázových systémov, bez ohľadu na produkčný informačný systém
Riešenie	Presne definovať potrebné predpoklady a znalosti na vykonanie tohto kroku počas tvorby MIS. Vytvoriť flow diagram, ktorý všeobecne opisuje proces importu dát, a vytvoriť sériu príkladov ako ukážku tohto procesu
Body realizácie	1. návrh a definícia potrebných predpokladov pre vykonanie procesov importu dát 2. návrh a tvorba flow diagramu postupu 3. výber databázových systémov 4. návrh a tvorba príkladov

Začali sme návrhom a definíciou vhodných predpokladov, a potrebných informácií pre úspešný import dát. V skratke bolo potrebné navrhnuť univerzálny súbor procesov, ako implementovať ETL. Vytvorili sme postupy a diagramy postupných krokov, ktoré bolo potrebné

vykonať. Vybrali sme vhodné DBMS (Database Management System) a vhodné dáta pre simuláciu podnikového prostredia. Pre vytvorenie a naplnenie databázových systémov sme použili prostriedky cloudovej platformy Microsoft Azure. Vytvorené prostredie sme mohli použiť na tvorbu vzorových príkladov, ktoré budú potom používané študentmi.

Stanovili sme súbor najdôležitejších otázok, na ktoré potrebuje riešiteľ nájsť odpoveď. Základom je súbor informácií, ktoré potrebujeme od cieľového používateľa (resp. zákazníka), aby sme doručili požadované výstupy. Potrebujeme informácie ohľadom informačných systémov, bezpečnostných pravidiel, pracovných časov, prihlásení, verzií a pod. Z toho teda vznikla ďalšia otázka ohľadom informačných systémov. Potrebujeme sa oboznámiť s informačnými systémami a s používanými databázami. Následne zisťujeme zloženie dát, merné jednotky, lokality, frekvenciu obnovy dát, metadáta [6] a pod. Tieto dôležité fakty je potrebné konzultovať aj s manažmentom podniku.

Simulované produkčné systémy sme založili na MySQL, SQL Server (Microsoft SQL), MariaDB databázach. Použili sme rôzne verzie a rôzne aplikácie na manažment databáz. Pri MariaDB sme použili phpMyAdmin. Následne pri MySQL sme použili MySQL Workbench. Pri databáze od Microsoftu sme použili SSMS (SQL Server Management Studio) a mssql-cli nástroj. Prípadová štúdia pozostávala zo série príkladov, ktoré sa zaoberali exportom dát pomocou CSV, formátom súborov programu MS Excel typu XLSX a SQL (Structure Query Language). Taktiež sme použili aj migračné prostredie od Oracle a vlastný automatizovaný skript v programovacom jazyku Python. Vlastné skripty obsahovali aj transformáciu dát pre Oracle.

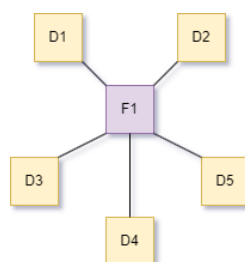
B. Prípadová štúdia 2

V prípadovej štúdií 2 sme sa zamerali na dátové modely, schémy dátových modelov a postup tvorby dátového modelu. Dátový model je dôležitý pri návrhu a realizácii dátovej kocky. Dátová kocka obsahuje tabuľky dimenzií a tabuľky faktov. Tieto tabuľky je potrebné prepájať cez primárne a cudzie kľúče. K tomu používajú študenti nástroj Oracle Data Modeler. Štruktúra prípadovej štúdie je znázornená v Tab. 2.

Tab. 2 Štruktúra prípadovej štúdie 2.

Názov	Dátové modely v nástroji Oracle SQL Developer Data Modeler
Aktuálny stav	Pri výučbe predmetu je používaný najčastejšie príklad s hviezdicovou schémou
Cieľ	Vytvoriť študijné materiály zaoberajúce sa procesom tvorby dátových modelov s ukázkami hviezdicovej schémy, schémy snehovej vločky a schémy súhvezdí
Riešenie	Presne definovať potrebné predpoklady a znalosti pre návrh dátového modelu a vytvoriť ukážkové príklady dátových modelov pomocou Oracle Data Modeler
Body realizácie	1. návrh a definícia potrebných predpokladov 2. návrh a tvorba flow diagramu postupu 3. výber vhodných ukážkových dát 4. návrh a tvorba schém pomocou ukážkových dát

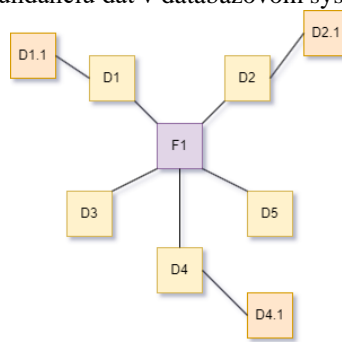
Študenti môžu vytvoriť viacero modelov. K tým základným patrí hviezdicová schéma. Pri tejto schéme je len jedna tabuľka faktov v strede modelu a je obkľúčená tabuľkami dimenzií [2], [6]. V tabuľke faktov sa potom nachádzajú merné jednotky, resp. fakty. K faktom patria napr. predajná cena, preprava a pod. Neskôr je možné generovať kalkulované fakty napr. predajná cena mínus profit dílera. Na Obr. 1 je hviezdicová schéma, ktorá bola najčastejšie používaná na predmete MIS.



Obr. 1 Hviezdicová schéma [6].

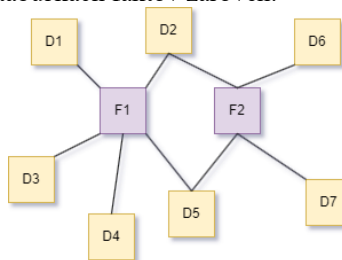
Našou snahou bolo, aby študenti viac používali schému snehovej vločky a schému súhvezdí. Pre spomínané schémy neboli vytvorené ukážkové dátové modely ani definované kritéria, ako by

mali vyzerat'. Schéma snehovej vločky má rovnako jednu tabuľku faktov a viacero dimenzií. Tieto dimenzie môžu (ale nemusia) mať ďalšie poddimenzie. Je to vhodné v komplikovanejších situáciách. Cieľom je znížiť redundanciu dát v databázovom systéme [4], [6].



Obr. 2 Schéma snehovej vločky [6].

Najzložitejšia je schéma súhvezdí. Táto schéma obsahuje viacero tabuliek faktov. Dimenzie môžu byť použité vo viacerých tabuľkách faktov zároveň.



Obr. 3 Schéma súhvezdia [6].

Pre vytvorenie jednotlivých relačných databáz sme použili Oracle databázu spravovanú v rámci Centra moderných metód riadenia a priemyselnej informatiky (CMMRaPI) na KKUI FEI TU. Na tejto databáze sme po tvorbe dátových modelov nasadili ukázkové datasety. Štúdiu sme uzavreli zhrnutím dôležitých poznatkov. V Tab. 3 sme porovnali jednotlivé rozdiely schém dátových modelov. Táto porovnávacia tabuľka uľahčí študentom pochopenie tejto problematiky.

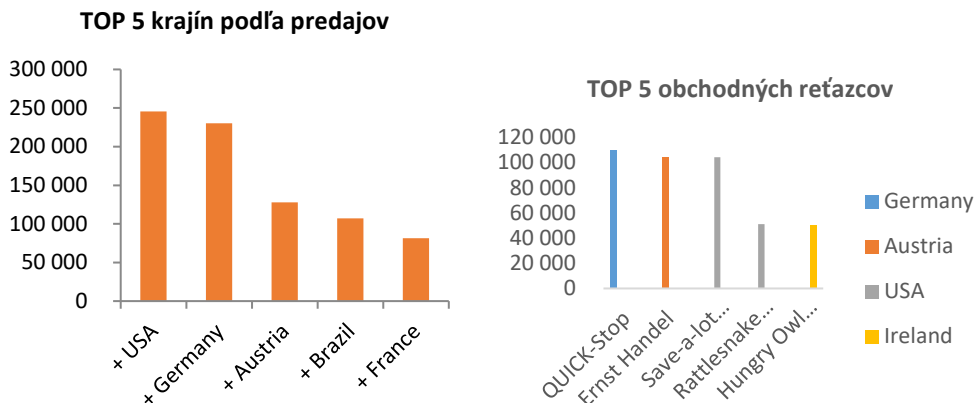
Tab. 3 Porovnanie schém dátových modelov.

	Hviezdicová	Snehová vločka	Súhvezdie
Prvky	jedna tab. faktov m. tab. dimenzií 0 tab. pod dimenzií	jedna tab. faktov m. tab. dimenzií n. tab. pod dimenzií	f. tab. faktov m. tab. dimenzií 0 tab. pod dimenzií
Normalizácia	Ne-normalizované	Normalizované	Normalizované
Výkon	Menej cudzích kľúčov – zvýšený výkon	Viac cudzích kľúčov – znížený výkon	Viac cudzích kľúčov – znížený výkon
Dizajn	Jednoduchý na tvorbu a pochopenie	Komplikovanejší ako hviezdicová schéma	Komplikovaný na pochopenie pri veľkých štruktúrach
Úložisko	Vyžaduje viac diskového miesta – vyššia redundancia dát	Potrebuje menej diskového miesta – snaha potlačiť redundanciu dát	Minimalizácia diskového miesta – nízka redundancia dát

IV. REALIZÁCIA VZOROVEJ APLIKÁCIE

Pre overenie inovovaných výukových materiálov sme vytvorili vzorovú aplikáciu analytického informačného systému. V tejto vzorovej aplikácii sme použili Northwind B2B dataset [7]. Simulovali sme rozloženie dát na viaceré RDBMS (Relational DBMS). Relačné tabuľky objednávok a zákazníkov boli nasadené na MariaDB, kde Python skript postupne pridával jednotlivé objednávky. Dáta o dodávateľoch a produktoch sme vložili na SQL Server. MySQL obsahoval údaje o zamestnancoch, lokalitách a prepravcoch. Pomocou príkladov prípadových štúdií sme získali zdrojové dáta, upravili ich a vložili do Oracle databázy. Použili sme schému súhvezdia. Nad naplnenou relačnou databázou sme vytvorili analytické prostredie. Vytvorili sme súbor analýz dát cez viacero nástrojov. Použili sme Oracle Analytic Workspace Manager

a OracleBI doplnok pre MS Excel. Paralelne sme vyexportovali dátovú množinu do CSV súborov a použili ich v nástroji Oracle Analytics Desktop. V tomto nástroji sme vytvorili analytické prostredie. Grafy na Obr. 4 sú ukážkou jednej z viacerých vytvorených analýz.



Obr. 4 Ukážka analýzy dát.

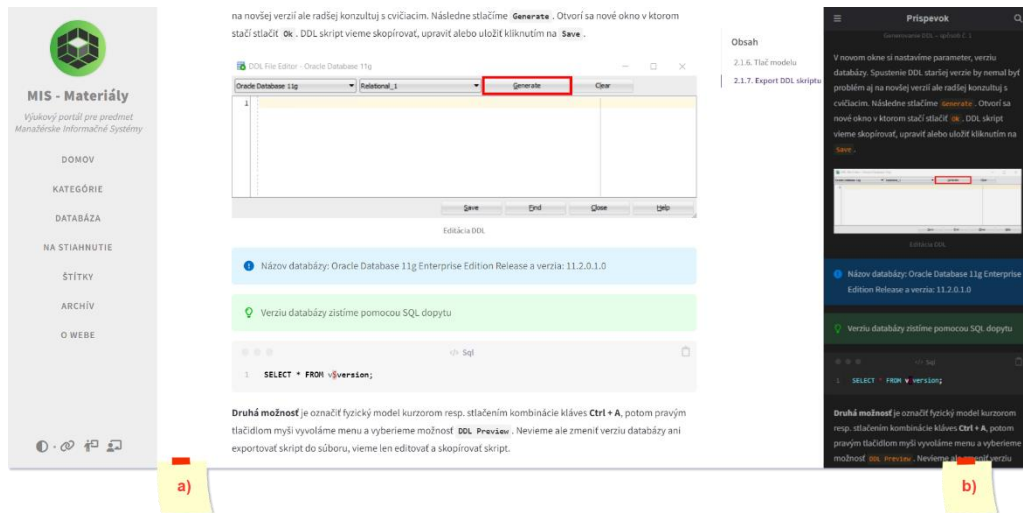
V. VÝUKOVÝ PORTÁL

Inovované výukové materiály bolo potrebné umiestniť na web. Najprv sme zosumarizovali a vyhodnocovali požiadavky vyučujúcich predmetu MIS. Na základe požiadaviek sme navrhli vhodné riešenie. Po odsúhlasení návrhu sme pokračovali výberom správnej technológie. Pri výbere technológie sme brali do úvahy tieto požiadavky:

- tvorba moderného statického webu,
- nevyžaduje CMS (Content Management System), resp. je možné použiť akýkoľvek webový server,
- jednoduchá tvorba a úprava obsahu,
- je open source (bez obmedzení),
- vývoj pokračuje, veľká komunita vývojárov a používateľov, široký výber doplnkov a tém.

Spomedzi dostupných technológií, ktoré spĺňajú uvedené požiadavky, sme zvolili generátor *Jekyll*. Tento generátor je Open Source, umožňuje vytvárať a upravovať obsah pomocou Markdown súborov, podporuje šablónovací jazyk Liquid a nevyžaduje databázu alebo CMS [8]. Generátor *Jekyll* vygeneruje jednoduché HTML, CSS a JS súbory, ktoré je možné umiestniť na akýkoľvek webový server. Markdown umožňuje automatizáciu tvorby a úpravy obsahu. Obsah vytvorený v dokumente Markdown je možné s drobnými úpravami exportovať ako PDF súbor.

Rozhodli sme sa pre textovú podobu výukových materiálov. Rôzne štúdie [9], [10] ukazujú, že pri použití audiovizuálnych materiálov sa nepreukázali významné rozdiely výsledkov v porovnaní s textovými výukovými materiálmi. Jednotlivé kroky metodiky tvorby MIS sú reprezentované ako články na webe [11]. Každý článok má svojho autora, dátum, obsah a pod. Z praktického hľadiska sme pridali aj tzv. čas na prečítanie. Na základe našich pozorovaní sme zistili, že študenti vyžadujú jednoduché, čitateľné a krátke texty. Nastavili sme automatickú detekciu farebnej schémy klienta. Web tak dokáže prispôsobiť tému preferenciám používateľa. Použili sme základnú svetlú a tmavú tému. Dizajnový jazyk sme sa snažili čo najviac prispôsobiť webovým stránkam samotného predmetu MIS. Okrem výukových materiálov je možné na webe nájsť aj informácie ohľadom pripojenia k databáze, návody na inštaláciu a nastavenie nástrojov, a súbory na stiahnutie. Sekcia na stiahnutie obsahuje PDF návody pre tlač, vzorové dáta, inštaláčne súbory nástrojov a pod. Ukážku svetlej webovej a tmavej mobilnej verzie webu vidíme na Obr. 5.



Obr. 5 Ukážka výukového portálu – a) desktopová verzia, b) mobilná verzia.

VI. ZÁVER

Cieľom tohto článku bolo sumarizovať prínosy výsledkov analýzy aktuálnych materiálov na stránke predmetu MIS, predstaviť inovované výukové materiály, a rozšírenie výukových materiálov o prípadové štúdie tvorby manažérskych informačných systémov. Inovované výukové materiály sú uverejnené na novovytvorenom výukovom portáli pre predmet MIS. Výsledné príklady z prípadových štúdií budú používané pri výučbe cvičení predmetu MIS. Prípadová štúdia 1 pojednávajúca o importe dát do RDBS Oracle je vhodná aj pre iné predmety priemyselnej informatiky, v ktorých je potrebné pracovať s dátami a databázami. Pri výbere vhodných technológií pre výukový portál sme uvažovali aj o rozšírenom použití týchto technológií pre webové stránky iných predmetov alebo personálne stránky vyučujúcich. Výukový portál vytvára jednotné prostredie pre všetky výukové materiály, čo má potenciál viesť k rozmanitejším a sofistikovanejším semestrálnym zadaniam.

LITERATÚRA

- [1] BASL, Josef a Blažiček Roman. Podnikové informační systémy: podnik v informační společnosti. 3., aktualiz. a dopl. vyd. Praha: Grada, 2012. 323 s. ISBN 978-80-247-4307-3 (viaz.)
- [2] LACKO, Luboslav. Databáze: datové sklady, OLAP a dolování dat s příklady v Microsoft SQL Serveru a Oracle. Brno: Computer Press, 2003. ISBN 80-7226-969-0
- [3] CMMRaPI. Webstránka predmetu: Manažérske informačné systémy [online]. CMMR a PI [cit. 2023-05-11]. Dostupné na internete: <http://matlab.fe.i.tuke.sk/mis/index.html>
- [4] NexTech. Dátové sklady, tržiště a jezera: Jak správně plavat v záplavě dat a informací. Computerworld: Ucelený informační zdroj pro IT profesionály. Praha: IDG Czech, a.s., 2023, roč. 34, č. 3, s. 10-11. ISSN 1210-9924
- [5] Štatistický úrad Slovenskej republiky. DATAcube [online]. Štatistický úrad Slovenskej republiky [cit. 2023-05-12]. Dostupné na internete: https://slovak.statistics.sk/wps/portal/ext/Databases/DATAcube_sk/
- [6] OKLAMČÁK, Samuel. Inovácia materiálov pre predmety priemyselnej informatiky so zameraním na manažérské informačné systémy - výukový portál : Diplomová práca. Košice : Fakulta elektrotechniky a informatiky Technická univerzita v Košiciach, 2023. 87 s. Školiteľ: doc. Ing. Ján Jadlovský, CSc.
- [7] Microsoft. Northwind and pubs sample databases for Microsoft SQL Server. [cit. 2022-12-15]. Dostupné na internete: <https://github.com/Microsoft/sql-serversamples/tree/master/samples/databases/northwind-pubs>
- [8] Jekyll. Transform your plain text into static websites and blogs. [cit. 2023-05-15]. Dostupné na internete: <https://jekyllrb.com/>
- [9] BREIMER, Eric, Michelle CONWAY, Jami COTLER a YODER, Robert. A study of video-based versus text-based labs for a management information systems course. In: Proceedings of the 16th annual joint conference on Innovation and technology in computer science education [online]. New York, NY, USA: ACM, 2011, 2011-06-27, s. 128-132 [cit. 2023-05-16]. ISBN 9781450306973. Dostupné z: doi:10.1145/1999747.1999785
- [10] BRIMER, Eric, COTLER, Jami a YODER, Robert. 2012. Video vs. text for lab instruction and concept learning. Journal of Computing Sciences in Colleges. 27, 6 (June 2012), 42–48.
- [11] CMMRaPI. MIS - Materiály [online]. CMMRaPI [cit. 2023-05-16]. Dostupné na internete: <http://matlab.fe.i.tuke.sk/mis/materialy/>

Reinforcement learning a radiálne siete pre potreby riadenia pohybu

Ján VAŠČÁK

Katedra Kybernetiky a Umelej Inteligencie, Fakulta Elektrotechniky a Informatiky,
Technická Univerzita v Košiciach, Slovenská Republika

jan.vascak@tuke.sk

Abstrakt – Táto práca sa zaoberá problematikou vytvorenia riadiacich reinforcement learning agentov, ktorí využívajú fuzzy systémy a ich porovnanie s implementáciami agentov založených na neurónových sieťach. Hlavným cieľom bolo implementovať modely založené na neurónových sieťach DQN a A2C, modifikovať tieto modely pre funkciu s radiálnou sieťou (RBF) a napokon porovnať ich v simulačnom prostredí OpenAI GYM na diskretných ako aj spojitých úlohách. Získané výsledky z experimentov preukazujú možnosť dostatočne kvalitného tréningu RBF sietí s využitím RL.

Príslušné slová – fuzzy logika, hybridná výpočtová inteligencia, reinforcement learning, riadenie pohybu

I. ÚVOD

Učenie s posilňovaním, známe v origináli ako *reinforcement learning* (RL) patrí v súčasnosti k najdôležitejším metódam strojového učenia. Využíva najmä prístup pokus–omyl a využíva odmenu a trest spätnej väzby z prostredia na úpravu alebo zlepšenie svojho správania a nakoniec na dosiahnutie optimálnej stratégie. Keďže RL má nízke nároky na predchádzajúce znalosti prostredia a môže vykonávať online učenie v prostredí reálneho času, túto metódu je možné využiť všade tam, kde je problém získať dostatok reálnych dát a vyžaduje sa rýchle prispôbenie zmeneným podmienkam, či parametrom, čo je práve prípad robotiky. Základnou úlohou RL je použiť parametrizovanú funkciu na aproximáciu mapovania zo stavu na akciu, ktorá môže byť určená tzv. *hodnotovou funkciou*, kde sa hodnotí buď vhodnosť daného stavu (*hodnotová funkcia stavu*), alebo vhodnosť konkrétnej kombinácie stavu a akcie, čo je tzv. Q-learning (QL). Klasický RL je založený na prehľadovej tabuľke, ktorá opisuje hodnotovú funkciu v malom diskretnom priestore stavu a akcie. Stavový priestor je však v skutočnom systéme zvyčajne spojitý, čiže predstavujúci v podstate nekonečný počet stavov a vyjadrenie vysokodimenzionálneho stavového priestoru pomocou prehľadovej tabuľky zaberá veľa pamäte. Jedným z možných východísk pre riešenie problému je použitie metódy *mapovania funkcií*.

Keďže neurónová sieť dokáže aproximovať funkciu s ľubovoľnou presnosťou a má silnú schopnosť adaptácie a zovšeobecnenia, tak v [1] navrhli metódu, ktorá kombináciou neurónovej siete s RL upravuje Q-funkciu (vhodnosť kombinácie stavu a akcie), čím možno dosiahnuť lepšie výsledky. Metóda využíva dynamický algoritmus zhlukovania a algoritmus najmenších štvorcov na nastavenie parametrov skrytej vrstvy a váh výstupnej vrstvy, ktoré môžu takto rýchlo a stabilne viesť ku konvergencii riešenia s dobrou schopnosťou zovšeobecňovať, avšak v diskretnom stavovom priestore. V [2] je predstavený algoritmus metódy učenia *actor–critic* v spojitom stavovom priestore. Aby bolo možné pracovať s veľkými diskretnými alebo spojitými stavovými priestormi, tak algoritmus využíva upravenú radiálnu sieť (Radial Basis Function – RBF). Podobnej problematike sa venuje aj práca [3], ktorá taktiež rieši implementáciu prístupu algoritmu *actor–critic* s použitím rôznych typov RBF funkcií a navrhnuté riešenie je aplikované na ovládanie inverzného kyvadla. V práci [4] je ukázané riešenie vysporiadania sa so spojitým stavovým priestorom učením hodnotových funkcií pomocou hlbokých radiálnych sietí, konkrétne algoritmus Q-learning (QL) je tu realizovaný pomocou radiálnych funkcií. Napokon, v práci [5] je navrhnutá fuzzy aproximačná architektúra, ktorá je kombinovaná s pôvodným iteračným algoritmom QL. Poznatky z vyššie spomenutých prác v ďalšom využijeme na vlastný návrh fuzzy RL pre potreby riadenia rôznych druhov pohybu.

Na základe vyššie uvedeného je štruktúra tejto práce nasledovná. V kapitole II je popísaný výskumný cieľ práce a použité metódy. V kapitole III sú popísané konkrétne pohybové úlohy v rámci OpenAI GYM, ktoré boli vybrané pre vývoj a porovnávanie implementovaných modelov.

V kapitole IV sú uvedené porovnávacie experimenty a výsledky, ktoré boli realizované v tejto práci. Napokon v kapitole V sú zhrnuté hlavné výsledky.

II. CIEĽ A POPIS POUŽITÝCH METÓD

Cieľom práce je porovnanie implementácie agentov založených na konvenčných neurónových a RBF (radiálnych) sieťach. Porovnanie bude prebiehať v jednoduchom simulačnom prostredí formou RL, kde je využívaný aktuálny stav prostredia a doposiaľ dosiahnutá odmena, ktoré slúžia ako vstupy do agenta, v tomto prípade mobilného robota. Po spracovaní tohto stavu agent vráti akciu, ktorá by mala byť vykonaná.

Cieľom teda bude porovnať tieto jednotlivé prístupy z hľadiska:

- ako náročné je ich implementovať,
- ako rýchlo sa dokážu naučiť vyberať akcie v jednotlivých prostrediach,
- rozdielov vo výkone medzi diskretnými a spojitými prostrediami.

A. Metódy RL

RL je jednou z troch základných paradigiem strojového učenia spolu s učením s učiteľom a učením bez učiteľa. RL sa líši od učenia s učiteľom v tom, že nie je potrebné prezentovať označené dvojice vstup/výstup a nie je potrebné, aby boli explicitne opravené suboptimálne činnosti. Namiesto toho sa zameriavame na nájdenie rovnováhy medzi prieskumom neprebádaného územia a využívaním súčasných poznatkov.

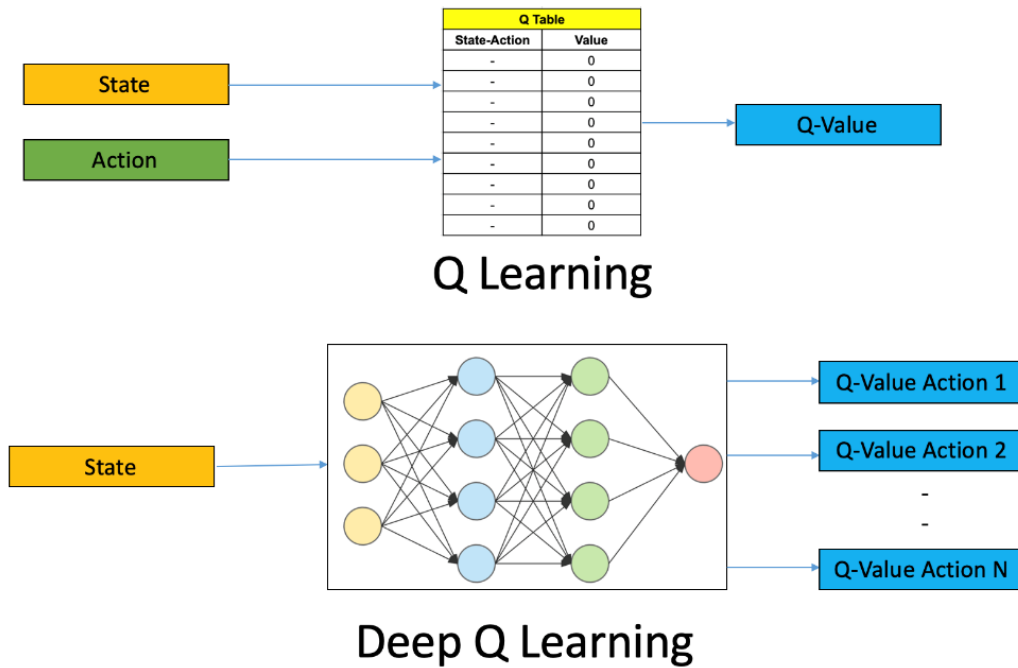
Pri kontrolovanom učení majú údaje o tréningu klúč k požadovanej odpovedi, takže model je tréningovaný so správnou odpoveďou sám, zatiaľ čo v RL neexistuje žiadna odpoveď, ale samotný algoritmus rozhodne, čo robiť. V prípade absencie tréningových údajov sa musí učiť zo svojich skúseností. Ide o prijatie vhodných opatrení na maximalizáciu odmeny v konkrétnej situácii. RL využíva aktuálny stav prostredia a doposiaľ dosiahnutú odmenu ako vstup a po spracovaní tohto stavu algoritmus vráti akciu, ktorú by model, resp. agent využívajúci takýto model, mal vykonať. V prípade RL neexistuje fáza učenia a fáza života ako u neurónových sietí, ale RL algoritmus neustále nastavuje svoj model.

QL je tzv. *model-free* algoritmus učenia na zistenie hodnoty akcie A v konkrétnom stave S , tzv. Q hodnoty. Nevyžaduje model prostredia a dokáže zvládnuť problémy so stochastickými prechodmi a odmenami bez potreby ďalších úprav. V závislosti od toho, kde sa agent v prostredí nachádza, rozhodne o ďalšom kroku. Prechod medzi jednotlivými stavmi je zabezpečovaný akciami, ktoré vyberá QL algoritmus. Za každú akciu dostane agent vykonávajúci QL odmenu alebo trest. Ak je dosiahnutý tzv. stav ukončenia, kde už nie je možné vykonať ďalšiu akciu, tak sa uzatvára jedna epizóda algoritmu. V našom prípade takýmto stavom ukončenia je cieľ cesty. Q hodnota teda určuje, aká kvalitná je akcia, ktorá bola vykonaná v konkrétnom stave S , čiže ju môžeme zapísať ako $Q(A, S)$.

Pôvodne navrhnutý (konvenčný) QL obsahuje tzv. Q tabuľku, ktorá kombinácii stav–akcia priradí konkrétnu hodnotu. Na začiatku sa zvyčajne inicializuje na 0, vid' Obr. 1. Agent začína v počiatočnom stave, kde vykoná nejakú akciu, za ktorú dostane odmenu. Túto skúsenosť si vo forme odmeny (kladná alebo záporná) zapíše do tabuľky. Týmto jednoduchým princípom si agent postupne vytvára tabuľku Q hodnôt, podľa ktorých môže robiť rozhodnutia. Tento tabuľkový prístup je obmedzený na prostredia s menším počtom stavov a akcií a je nevhodný pre prostredia so spojitými akciami alebo stavmi.

Jedným z modernejších prístupov vylepšujúcich konvenčný QL je neurónová sieť založená na hlbokom QL (DQN), kde namiesto tabuľky použijeme na predikovanie Q hodnoty neurónovú sieť. Tento prístup je robustnejší, ale tiež nevhodný pre spojitá prostredia. Pri hlbokom QL používame neurónovú sieť na aproximáciu funkcie Q hodnoty. Stav je daný ako vstup a ako výstup je generovaná Q hodnota všetkých možných akcií. Porovnanie medzi QL a hlbokým QL je znázornené na Obr. 1.

Zo štruktúrneho hľadiska existujú aj pokročilejšie techniky pre RL agentov nazývané *actor-critic* ako napr. A2C, kde QL je možné vnímať iba ako časť *critic*. Tieto techniky obsahujú neurónovú sieť so spoločnými vrstvami, ktoré sa ďalej delia na vrstvu *actor* a vrstvu *critic*. A2C je možné na rozdiel od QL spoľahlivo využiť aj v spojitých prostrediach (stavy a aj akcie). *Actor* vykonáva akcie v prostredí a *critic* má za úlohu ohodnotiť tieto akcie. Používajú na to tzv. *funkciu výhodnosti*, ktorá vypočítava, o koľko je konkrétny pár stav–akcia lepší ako ostatné možné (vykonateľné) akcie. Bližšie informácie o technike *actor-critic* je možné nájsť napr. v [7] a [8]. V [4] bola využitá RBF sieť pre realizáciu *actor-critic*, kde RBF sieť nahrádza práve túto dvojvrstvu v sieti *actor-critic*.



Obr. 1 Schéma porovnávajúca konvenčný (tabuľkový) QL a DQN [6].

B. Radiálne siete

Radiálne siete sú vždy trojvrstvové neurónové siete, ktoré vo svojej skrytej vrstve ako aktivačné funkcie využívajú radiálne (bázové) funkcie (RBF) a skrátene sa nazývajú RBF siete, vid' Obr. 2. Princíp radiálnej siete bol odvodený od funkcie tzv. vizuálneho kortexu mozgu živých organizmov, ktorý je zodpovedný za transformáciu zrkovných vnemov do štruktúry mozgu. Z tohto dôvodu sa skrytá vrstva aj nazýva *vrstvou receptívnych polí*, kde vzhľadom na vlastnosti RB funkcií ϕ , sa tieto postupne postupne aktivujú z daného intervalu aktivačných hodnôt a aktivácie susedných funkcií sa môžu navzájom prelínať, čoho analógiu môžeme nachádzať vo fuzzy logike. Z toho dôvodu sa aj radiálne siete často využívajú spolu s fuzzy logikou. Podmienku RB funkcie spĺňa celá rada známych analytických funkcií, z ktorých najznámejšou je Gaußovská funkcia a túto budeme aj v ďalšom používať.

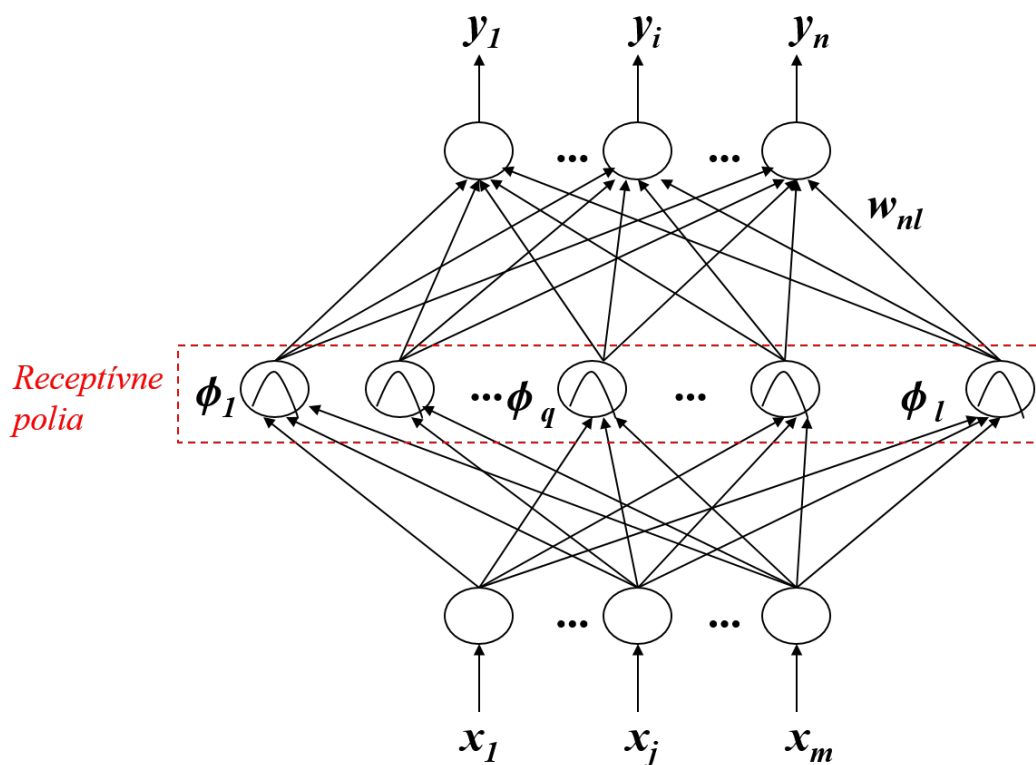
III. VYBRANÉ POHYBOVÉ ÚLOHY

Aby sa otestovali metódy uvedené v predošlej kapitole, bolo zvolené prostredie OpenAI GYM [9], čo je open source prostredie s nástrojmi na vývoj a porovnanie algoritmov RL a úloh, ktoré riešia. Konkrétne sa bude jednať o prostredia *Mountain car*, *Cart pole* a *Lunar lander*, ktorých popis bol pre naše potreby mierne upravený.

V prostredí *Mountain car* je auto na jednorozmernej dráhe umiestnenej medzi dvomi vyvýšeninami. Cieľom je vyjsť s autom na vyvýšeninu vpravo. Motor auta však nie je dostatočne silný na to, aby na jeden pokus zdolal takéto prevýšenie. Jediným spôsobom ako uspieť, je preto jazdiť tam a späť, až kým sa nezíska potrebná energia, vid' Obr. 3 (a).

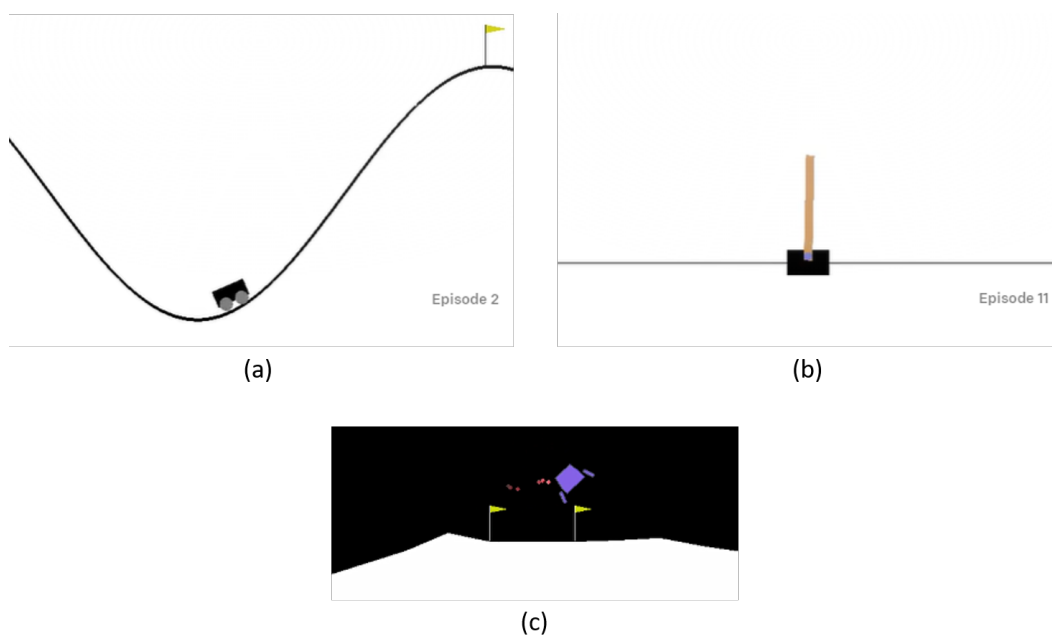
Prostredie *Cart pole* zobrazuje problém známy aj ako inverzné kyvadlo, ktoré je zospodu pripevnené kĺbom k vozíku (na rozdiel od zaveseného kyvadla) a ktorý sa pohybuje po vodorovnej dráhe na obidve strany s cieľom udržať kyvadlo vo vzpriamenej podobe, vid' Obr. 3 (b). Systém sa ovláda pôsobením kladnej alebo zápornej sily na vozík. Znamienko závisí od smeru pôsobenia. Odmena +1 sa poskytuje za každý časový krok, pri ktorom kyvadlo zostane vzpriamené. Epizóda sa ukončí, keď je kyvadlo viac ako 15 stupňov od vertikálneho smeru, alebo sa vozík pohne o viac ako 2,4 dĺžkovej jednotky od stredu.

Úloha *Lunar lander* (Obr. 3 (c)) je z uvedených úloh najkomplexnejšia vzhľadom na počet možných akcií (spolu štyri) ako aj spôsob priradovanie odmien. V tomto prípade má lunárny pristávací modul zostúpiť z hornej časti obrazovky na pristávaciu plochu vymedzenú zástavkami v spodnej časti obrazovky, pričom pri jej dosiahnutí má mať nulovú rýchlosť. Pri úspešnom zvládnutí tejto operácie sa odmena pohybuje v rozmedzí 100–140 bodov. Ak sa modul vzdialí od pristávacej plochy, odmenu stratí. Epizóda sa skončí, ak pristávací modul havaruje alebo sa zastaví, alebo ak získa ďalších -100 alebo +100 bodov. Každý kontakt so zemou je +10. Spustenie



Obr. 2 Schéma radiálnej siete.

hlavného motora je -0,3 bodu na každú časovú periódu. Prostredie sa považuje za vyriešené, ak agent dosiahne aspoň 200 bodov. Pristátie mimo pristávacej plochy je tiež možné, ale s menším počtom bodov. Palivo je nekonečné, takže agent sa môže naučiť lietať a potom pristáť na prvý pokus. K dispozícii sú štyri samostatné akcie: nerobiť nič, spustiť pomocný motor s orientáciou vľavo, spustiť hlavný motor, spustiť pomocný motor s orientáciou vpravo.



Obr. 3 Prostredia Mountain Car (a), Cart pole (b) a Lunar lander (c).

IV. POROVNANIE MODELOV A ICH VYHODNOTENIE

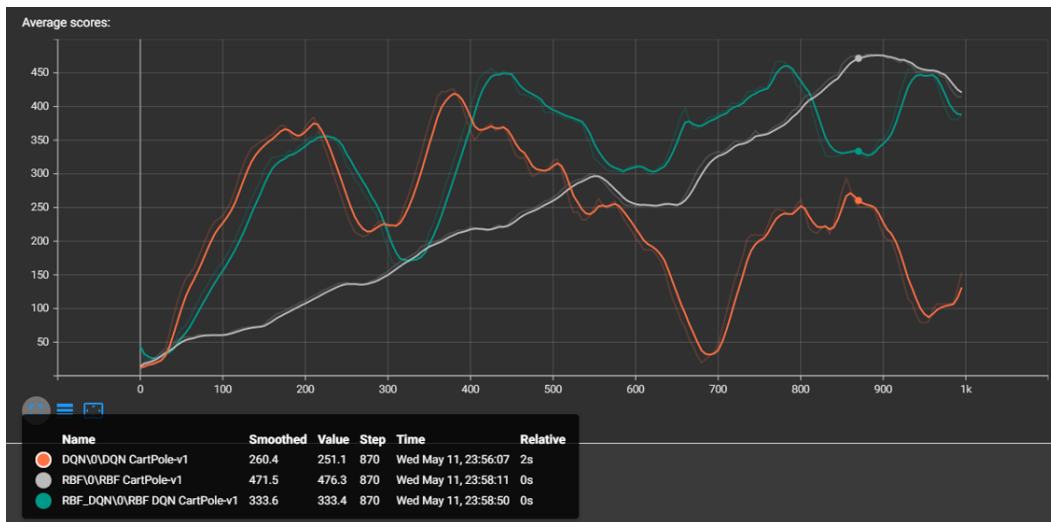
Experimenty pri riadení pohybov v prostrediach popísaných v predošlej kapitole je možné rozdeliť do dvoch skupín, a to experimenty s DQN sieťami a experimenty s A2C sieťami. Pre lepšie porovnanie boli vyhodnotenia vykonávané v diskretných ako aj v spojitých stavových prostrediach, pričom na základe vykonaných experimentov vieme, že agenti majú väčšinou lepšie výsledky v diskretných stavových prostrediach a spojité sú pre nich problematické, avšak v spojitých stavových prostrediach by agentom mala pomôcť práve RBF vrstva [1]. Vo všetkých prípadoch boli modely hodnotené dvakrát, ale do úvahy sa bral ten lepší výsledok.

A. Experimenty s DQN sieťami

Na vyhodnotenie úspešnosti DQN a RBF-DQN modelov sa implementovali modely vo všetkých troch diskretných stavových prostrediach. Pri hodnotení boli nastavené parametre takto:

- Počet tréningových epizód: 1000
- Rýchlosť učenia siete: 0.0001

Na Obr. 4 je zobrazený výsledok testovania DQN modelu v diskretnom prostredí *Cart pole*. Prostredie sa považuje za vyriešené dosiahnutím skóre 195. Na obrázku vidno, že DQN agenti sa zastavili na hodnote 425,4, RBF-DQN na hodnote 466 a samotná RBF sieť dosiahla skóre 476. Najvyššie možné skóre dosiahnuteľné agentmi je 500.



Obr. 4 Test DQN siete v diskretnom prostredí *Cart pole*.

V prostredí *Mountain car*, ktoré sa považuje za vyriešené dosiahnutím skóre 0, všetky tri prístupy zlyhali, keď získali nasledovné skóre: DQN: -118, RBF-DQN: -143 a samotná RBF: -155.

Posledné prostredie, na ktorom sa testovali DQN modely, je *Lunar lander*. Prostredie sa považuje za vyriešené dosiahnutím skóre 200 (najvyššie dosiahnuteľné skóre je 314). DQN a RBF agenti vyriešili prostredie – priemerné skóre je 231 a 226, RBF-DQN agentovi sa prostredie nepodarilo vyriešiť, nakoľko jeho skóre je iba 13. Tento fakt môže byť spôsobený komplexnou povahou prostredia *Lunar lander* oproti ostatným dvom prostrediam, čo si vyžaduje podstatne väčší počet neurónov v skrytej vrstve RBF siete pre aproximáciu aktivačnej funkcie.

B. Experimenty s A2C sieťami

Na vyhodnotenie úspešnosti A2C a RBF-A2C modelov boli vzájomne porovnané konvenčná A2C sieť a sieť A2C s RBF aktivačnou funkciou, a to v diskretných ako aj v spojitých prostrediach. Pri hodnotení boli nastavené parametre takto:

- Počet tréningových epizód: 5000
- Rýchlosť učenia siete: 0,0001

Konvenčná A2C sieť:

Jedná sa o základné modely A2C sietí, oproti ktorým sa bude porovnávať. Najprv všetky modely boli otestované v diskretnom prostredí a následne v spojitom. V prostredí *Mountain car* všetci agenti sa zastavili na hodnote -200, čiže zlyhali (požadovaná hodnota skóre je 0). Agenti podobne

zlyhali aj v prostredí *Lunar lander*, nakoľko najvyššie dosiahnuté skóre bolo 60 (požadované je aspoň 200). V prostredí *Cart pole* boli všetci agenti po 700 epizódach úspešní, nakoľko najvyššie dosiahnuté skóre bolo 500 (minimálne skóre je 195).

V ďalšom sa pokračovalo v testovaní v spojitých prostrediach. V prostredí *Mountain car* bola úloha vyriešená počas prvých 200 epizód. V prostredí *Lunar lander* sa úlohu nepodarilo vyriešiť, nakoľko najvyššie skóre bolo 94 a požaduje sa hodnota 200.

A2C s RBF aktivačnou funkciou:

Aj v tomto prípade boli modely otestované v diskretnom prostredí a následne v spojitom. V prípade diskretného prostredia *Cart pole* je vidieť, že RBF aktivácia znížila výkon siete. Skóre od začiatku až do konca sa držalo v okolí 20 bodov. Podobne to bolo aj v diskretnom prostredí *Lunar lander*, kde RBF aktivácia znížila výkon siete a aj drasticky zvýšila čas tréovania.

Odlíšna situácia nastala v spojitých prostrediach, zatiaľ čo v prípade prostredia *Lunar lander* boli výsledky podobné, avšak v prípade *Mountain car* sa dosiahli porovnateľné výsledky s konvenčnou A2C sieťou a testovanie bolo iba nepatrne dlhšie, avšak výhodou oproti tomu je stabilnejšie tréovanie. V Tabuľke 1 sa nachádza sumarizačné porovnanie použitia jednotlivých metód v daných modeloch z hľadiska ich úspešnosti.

Tabuľka 1

Porovnanie úspešnosti použitých metód pri riadení simulačných modelov: a – úspešné riadenie; n – neúspešné riadenie.

Metóda	Prostredie	Mountain car	Cart pole	Lunar lander
DQN	diskrétné	n	a	a
	spojité	–	–	–
RBF-DQN	diskrétné	n	a	n
	spojité	–	–	–
RBF	diskrétné	n	a	a
	spojité	–	–	–
A2C bez RBF	diskrétné	n	n	n
	spojité	a	a	n
A2C s RBF	diskrétné	n	n	n
	spojité	a	a	n

V. ZÁVER

Úspešne sa podarilo vytvoriť modely založené na neurónových sieťach DQN a A2C, modifikovať ich pre funkciu s RBF sieťou, porovnať ich v prostredí OpenAI GYM na diskretných a spojitých úlohách a zodpovedať otázku, či je možné tréovať fuzzy systémy pomocou princípov RL.

Vyhodnotenie jednotlivých modelov ukázalo, že použitie RBF vrstiev namiesto aktivačných funkcií sa neosvedčilo pri RL úlohách, dopredné neurónové siete s tradičnými aktivačnými funkciami budú skoro vo všetkých prípadoch lepšie. Taktiež bolo overené, že je možné tréovať RBF siete v RL úlohe.

POĎAKOVANIE

Táto práca je finančne podporená Vedeckou grantovou agentúrou MŠVVaŠ SR a SAV prostredníctvom projektu VEGA 1/0480/22 — EDEN: EDge-Enabled inteligeNtné systémy.

LITERATÚRA

- [1] Y.-M. Du, B.-B. Yan, Y.-C. Jiang, and Q.-S. Liu, *Reinforcement learning method based on radial basis function neural network*. World Scientific, 2017, pp. 21–27.
- [2] C.-G. Li, M. Wang, Z.-J. Huang, and Z.-F. Zhang, “An actor-critic reinforcement learning algorithm based on adaptive RBF network,” in *IEEE International Conference on Machine Learning and Cybernetics*, vol. 2, 2009, pp. 984–988.
- [3] Y.-H. Cheng, J.-Q. Yi, and D.-B. Zhao, “Application of actor-critic learning to adaptive state space construction,” in *IEEE Proceedings of 2004 International Conference on Machine Learning and Cybernetics*, vol. 5, 2004, pp. 2985–2990.
- [4] K. Asadi, N. Parikh, R. E. Parr, G. D. Konidaris, and M. L. Littman, “Deep radial-basis value functions for continuous control,” 2021.
- [5] L. Buşoniu, D. Ernst, B. De Schutter, and R. Babuška, “Continuous-state reinforcement learning with fuzzy approximation,” in *Adaptive Agents and Multi-Agent Systems III. Adaptation and Multi-Agent Learning*, K. Tuyls, A. Nowe, Z. Guessoum, and D. Kudenko, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 27–43.
- [6] A. H. Zarif, “Aoi minimization in energy harvesting and spectrum sharing enabled 6g networks,” 2021. [Online]. Available: <https://dx.doi.org/10.21227/aynt-my59>
- [7] V. R. Konda and V. S. Borkar, “Actor-critic-type learning algorithms for markov decision processes,” *SIAM Journal on Control and Optimization*, vol. 38, no. 1, pp. 94–123, 1999.
- [8] I. Grondman, L. Buşoniu, G. A. D. Lopes, and R. Babuska, “A survey of actor-critic reinforcement learning: Standard and natural policy gradients,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 1291–1307, 2012.
- [9] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, “OpenAI Gym,” 2016.

Vplyv prosumerov na distribučnú sieť

Kristián HABORÁK, Dušan MEDVEĎ

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

kristian.haborak@student.tuke.sk, dusan.medved@tuke.sk

Abstrakt — Tento príspevok sa zaoberá vplyvom prosumerov na distribučnú sieť. Prosumeri sú súčasťou elektrickej siete. Ich prítomnosť prináša kladné ale aj záporné stránky. Obnoviteľné zdroje energie, ktoré využíva prosumer na získanie energie, sú budúcnosťou po vyčerpaní nerastných surovín. Preto je potrebné neustále navyšovanie produkcie energie z obnoviteľných zdrojov lokálne aj vo svete. Cieľom príspevku je citlivostná analýza vstupných parametrov rôznych zostáv solárnych systémov tvoriaci off-grid model. Tento model tvoria prosumeri a užívatelia, ktorí sú od nich plne závislí. Parametre zostáv sú vložené do fotovoltaického geografického informačného systému, ktorý slúži na výpočet produkcie energie vo vybranej oblasti. Zmenou jednotlivých vstupných parametrov sa dosahujú rozdiely, ktoré sa analyzujú z hľadiska sebestačnosti modelu, produkcie energie, finančného hľadiska a uvádzajú sa návrhy na vylepšenie modelu.

Kľúčové slová — prosumer, obnoviteľné zdroje energie, off-grid, produkcia energie

I. ÚVOD

Pre účely tohto príspevku sme sa zamerali na porovnávanie viacerých zostáv solárnych systémov s postupne sa meniacimi vstupnými parametrami, čím sa určí dôležitosť zmeny parametra na výstup. Vstupnými parametrami sú denná spotreba, akumulácia schopnosť, ročné obdobie, inštalovaný výkon, sklon fotovoltaických panelov. Výstupom je výroba elektrickej energie, sebestačnosť solárneho systému, využitý potenciál slnečnej energie, počítacia investícia. Jednotlivé zostavy sú umiestnené do off-grid modelu, pričom sú vystavované rôznym podmienkam. Získané dáta boli analyzované, hľadalo sa optimálne riešenie pre danú situáciu. Cieľom optimalizácie je celoročná sebestačnosť modelu a jeho užívateľov. Skonstruovanie off-grid modelu poukazuje na realnosť takejto prevádzky v dnešných podmienkach, pri aktuálnej efektívnosti jednotlivých komponentov.

II. POSTUP PRI NÁVRHU HYBRIDNÉHO MODELU MERANIA

Na realizáciu modelu je využitý geografický informačný systém (PVGIS) [1]. Prvým krokom k výpočtu, je voľba umiestnenia solárneho systému. Tento krok je veľmi dôležitý, lebo podmienky daného prostredia určujú, koľko je daný systém schopný vyrobiť energie. Druhým vstupným parametrom je inštalovaný špičkový výkon solárneho systému. Túto informáciu je možné zistiť z technického listu solárneho panela a vynásobiť počtom kusov, ktorý solárny systém obsahuje. Tretím vstupným parametrom je kapacita batérie. Tu je potrebné dávať pozor a uviesť použiteľnú kapacitu danej batérie. Štvrtým vstupným parametrom je dolný hraničný limit vybíjania batérie, ktorý je možné určiť z hĺbky vybitia batérie. Táto veličina hovorí o úrovni nabitia, ktoré batéria počas vybíjania neprekročí. Výrobca batérie udáva v technickom liste až 95 hĺbku vybitia. Je potrebné dodať, že čím nižšia dolná hranica vybitia batérií, tým menej cyklov batéria vydrží. Batérie pri ostrovej prevádzke tvoria podstatnú časť investície, preto je potrebné zvážiť kúpu kvalitnejších batérií za vyššiu cenu. Piatym vstupným parametrom je spotreba objektu za deň. Tento údaj charakterizuje spotrebu energie daného objektu počas 24 hodín. Tento údaj je možné zistiť na základe faktúry za elektrinu, smart aplikácie, približným výpočtom. Denná spotreba je premenlivá v závislosti od časti dňa a je rozdelená na hodiny. Predpoklad je, že najviac energie sa minie počas večera [1].

Cursor:
 Selected: 49.173, 20.862
 Elevation (m): 491
 PVGIS ver.: 5.2

Use terrain shadows:
 Calculated horizon
 Upload horizon file

Nie je vybratý žiadny súbor

PERFORMANCE OF OFF-GRID PV SYSTEMS

Solar radiation database* PVGIS-SARAH2

Installed peak PV power [Wp]* 50

Battery capacity [Wh]* 600

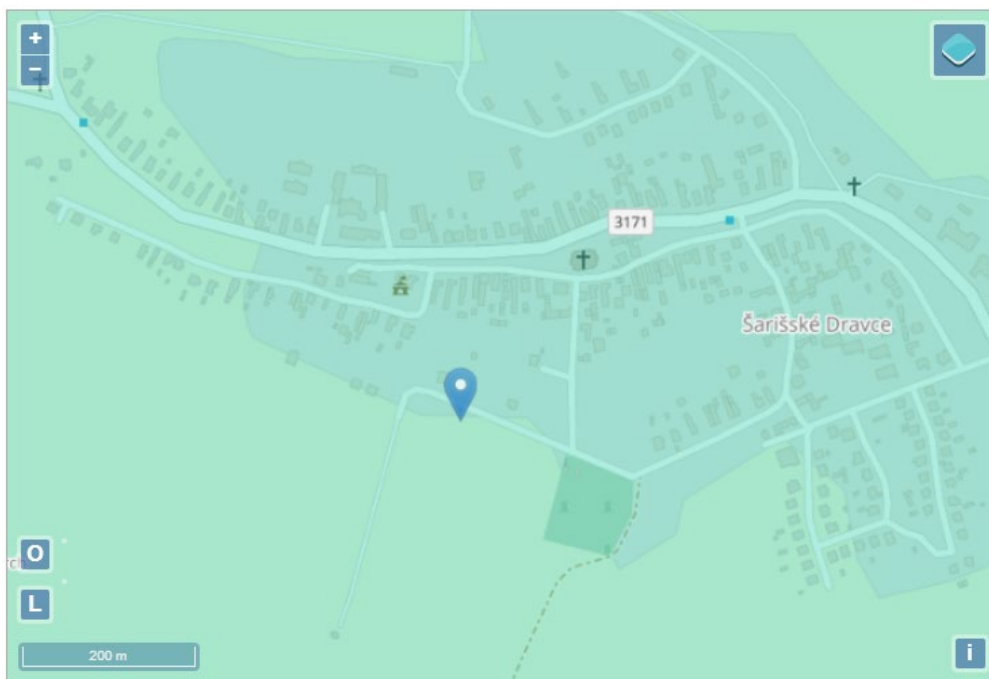
Discharge cutoff limit [%]* 40

Consumption per day [Wh]* 300

Upload consumption data Nie je vybratý žiadny súbor

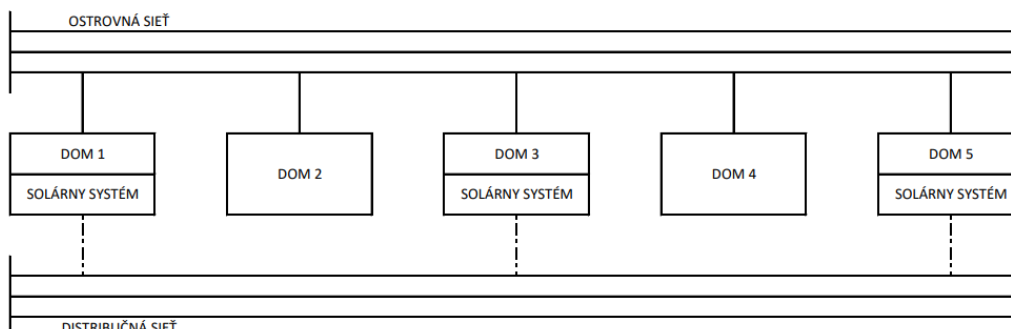
Slope [°]* 35

Azimuth [°]* 0



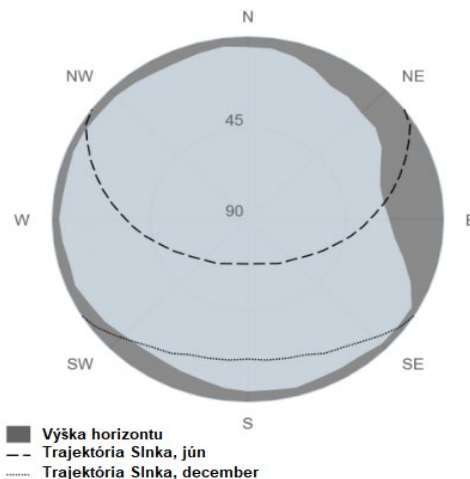
Obr. 1 Prostredie programu PVGIS [1]

Ostrovná prevádzka pozostáva z piatich nízkoenergetických domov, pričom majitelia troch domov získavajú elektrickú energiu za pomoci solárnych panelov a dvaja užívatelia nevlastnia solárne panely. Títo užívatelia, sú energeticky plne závislí od majiteľov vlastniacich FV systém. Jednotlivé domy sú pripojené do ostrovnej siete, ktorú spoločne zdieľajú všetci užívatelia.



Obr. 2 Hybridný model merania [4]

Sieť je ostrovného typu, preto je potrebné s energiou narábať šetrne. Varenie v domácnostiach je realizované plynom, aby sa znížila energetická náročnosť domov. Je potrebné použiť spotrebiče, ktoré majú menšiu spotrebu. Jednotlivé meniče energie musia medzi sebou komunikovať, aby sa energia získaná z FV panelov využila na spotrebu domácností alebo uložila do batérií. Ďalším predpokladom je, že všetky domy majú rovnaké spotrebiče. Jednotlivé domy majú wi-fi sieť tvorenú bezdrôtovým routerom na ovládanie komunikačného centra. V prípade nedostatku energie je možné chýbajúcu energiu doplniť z distribučnej siete. Na realizáciu modelu je využitý fotovoltický geografický informačný systém (PVGIS). Je potrebné priradiť modelu určitú lokalitu. Model je umiestnený na súradniciach 49.173, 20.862 v nadmorskej výške 491 m na okraji dediny Šarišské Dravce, Slovenská republika. Vplyv horizontu a trajektórie Slnka na produkciu elektrickej energie je zobrazený na nasledujúcom obrázku [3].



Obr. 3 Výška horizontu a trajektória Slnka na mieste modelu [2]

III. PRINCÍP VÝPOČTU

Pri výpočte výstupnej energie ostrovného modelu sa berie do úvahy slnečné žiarenie pre každú hodinu, pričom sa zohľadňujú hodnoty namerané v rokoch 2005 až 2020. PVGIS má zadefinovanú 67 % účinnosťou ostrovného solárneho systému. Solárne panely sú orientované smerom na juh s predpokladom získania najviac energie. Elektrocentrála s prevádzkovým výkonom 2000 W je schopná prevádzky 13 hodín, doplní chýbajúcu energiu v daných mesiacoch do výšky 26 kWh. Nádrž elektrocentrály sa doplní na začiatku každého mesiaca. Dolná hranica vybitia batérie je určená na 20%. Výpočet prebieha v nasledujúcich krokoch:

- Pre každú hodinu sa vypočíta slnečné žiarenie dopadajúce na solárne panely a výstupný výkon solárneho systému
- Ak je výkon solárneho systému väčší ako spotreba energie za danú hodinu, uloží sa zvyšok energie do batérií
- Ak je výkon solárneho systému menší ako spotreba energie, získa sa chýbajúca energia z batérií
- Ak sa batérie naplnia, energiu ktorú nie je možné spotrebovať ani uložiť, uloží do premennej nezachytená energia
- Ak batérie dosiahnu zvolený dolný hraničný limit vybíjania, vypočíta sa chýbajúca energia a pridá sa deň, počas ktorých sa systému vybila energia

IV. VÝSLEDKY MERANIA

Bolo vykonaných 5 meraní, pričom v tejto časti sa analyzuje vplyv sklonu panelov na výrobu elektrickej energie.

Tabuľka 1

Vplyv sklonu panelov na výrobu elektrickej energie vstupné parametre zostava 3 [4]

Zostava 3			
Meranie	1	2	3
Inštalovaný špičkový výkon (Wp)	2730	2730	2730
Kapacita batérie (Wh)	6748	6748	6748
Dolný hraničný limit vybíjania batérie (%)	20	20	20
Spotreba za deň (Wh)	11670	11670	11670
Sklon solárneho panelu (°)	30	45	60
Orientácia solárneho panelu (°)	0 - Juh	0 - Juh	0 - Juh

Tabuľka 2
Vplyv sklonu panelov na výrobu elektrickej energie namerané hodnoty model A [4]

Model A						
Meranie	1	2	3	1	2	3
Sklon panelov	30	45	60	30	45	60
Mesiac	Priemerná produkcia energie (Wh/d)			Priemerná nezachytená energia (Wh/d)		
Január	3158,9	3442,4	3527	8,7	102,3	204,6
Február	4409,9	4539,9	4521,6	197,9	400,3	496
Marec	6055,6	6032,1	5892,5	961,1	1175,9	1137,6
Apríl	7352	7189,4	6898,4	1571,8	1559,3	1245
Máj	7660,2	7390,8	6950,3	1485,4	1249,5	791,6
Jún	8278,3	7974,8	7453,9	1592	1237	678,9
Júl	8369,4	8107,8	7658	1625,2	1327,8	770,9
August	7903,1	7746,8	7463,3	1633,4	1581,8	1190,3
September	6623,6	6599,5	6464,8	1110,8	1350,1	1281,9
Október	5090,4	5214,1	5200,5	381,1	714,8	862,3
November	3125,7	3357,5	3425,4	21,5	140,7	237,3
December	2256,3	2515,1	2617,7	0	20,3	63,9
Mesiac	Plná batéria dni (%)			Batéria na 20% dni (%)		
Január	2,6	11,1	14,9	100	100	100
Február	15,8	20,1	22,1	100	100	100
Marec	36,3	37,5	37,5	100	100	100
Apríl	48,1	46,9	44	100	100	100
Máj	48,2	44,6	37,3	100	100	100
Jún	53,5	49,2	39,6	100	100	100
Júl	61,5	56,9	45	100	100	100
August	59,3	58,5	51,8	100	100	100
September	46,5	48,8	47,5	100	100	100
Október	29	32,9	34,9	100	100	100
November	4,6	11,9	13,8	100	100	100
December	0	3,5	6,1	100	100	100

Tabuľka 3
Vplyv sklonu panelov na výrobu elektrickej energie porovnanie meraní 1-2-3 [4]

Meranie	1	2	3
Priemerná mesačná výroba model A (kWh)	535	533,7	518,1
Ročná výroba model A (kWh)	6420,5	6404	6217,4
Chýbajúca energia model A za rok (kWh)	6358,1	6374,6	6561,2
Nezachytená energia model A za rok (kWh)	970,1	993,5	818,5

Tabuľka 4
Vplyv sklonu panelov na výrobu elektrickej energie vstupné parametre [4]

Zostava 6			
Meranie	4	5	6
Inštalovaný špičkový výkon (Wp)	2730	2730	2730
Kapacita batérie (Wh)	13496	13496	13496
Dolný hraničný limit vybíjania batérie (%)	20	20	20
Spotreba za deň (Wh)	11670	11670	11670
Sklon solárneho panelu (°)	30	45	60
Orientácia solárneho panelu (°)	0 - Juh	0 - Juh	0 - Juh

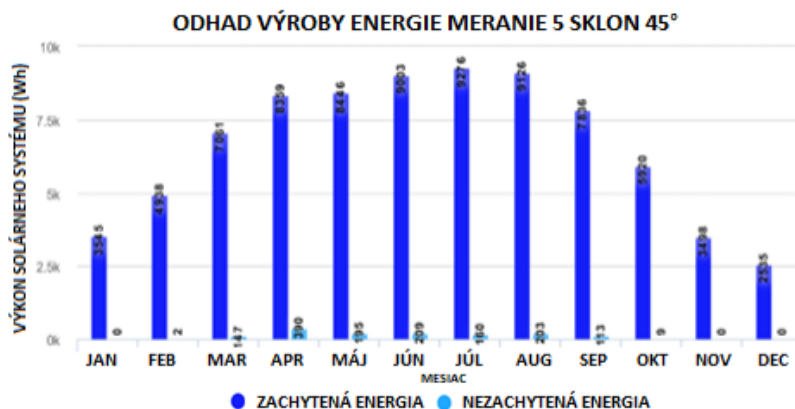
Tabuľka 5
Vplyv sklonu panelov na výrobu elektrickej energie namerané hodnoty model B [4]

Model B						
Meranie	4	5	6	4	5	6
Sklon panelov	30	45	60	30	45	60
Mesiac	Priemerná produkcia energie (Wh/d)			Priemerná nezachytená energia (Wh/d)		
Január	3167,6	3544,7	3731,6	0	0	0
Február	4607,7	4938,3	5013,2	0	1,9	4,5
Marec	6949,6	7060,8	6904,3	67,1	147,2	125,7
Apríl	8521,4	8359	7940,1	402,4	389,6	203,2
Máj	8800,4	8445,7	7732,1	345,3	194,6	9,8
Jún	9439,3	9003,1	8131,4	431	208,7	1,4
Júl	9666,5	9275,9	8426,7	328	159,7	2,2
August	9295,3	9125,7	8612,5	241,2	202,9	41,1
September	7680,6	7836,1	7669,4	53,8	113,5	77,3
Október	5471,6	5919,6	6038,8	0	9,3	23,9
November	3147,2	3498,1	3662,7	0	0	0
December	2256,3	2535,4	2681,6	0	0	0
Mesiac	Plná batéria dni (%)			Batéria na 20% dni (%)		
Január	0	0	0	100	100	100
Február	0	0,2	0,5	99,8	98,4	97,3
Marec	5,7	9,9	9,1	85,7	83,3	84,3
Apríl	18,3	17,9	13,8	70,6	72,5	77,1
Máj	16,5	10,9	1	70,4	77	87,1
Jún	21	12,5	0,4	63,5	73,5	86,5
Júl	18,1	11,1	0,2	61,7	70,2	85,7
August	14,9	13,7	4,4	65,3	67,5	76,2
September	4,6	9,4	7,7	80	73,5	75,4
Október	0	1,8	3,6	98,8	93,2	90,1
November	0	0	0	100	100	99,8
December	0	0	0	100	100	100

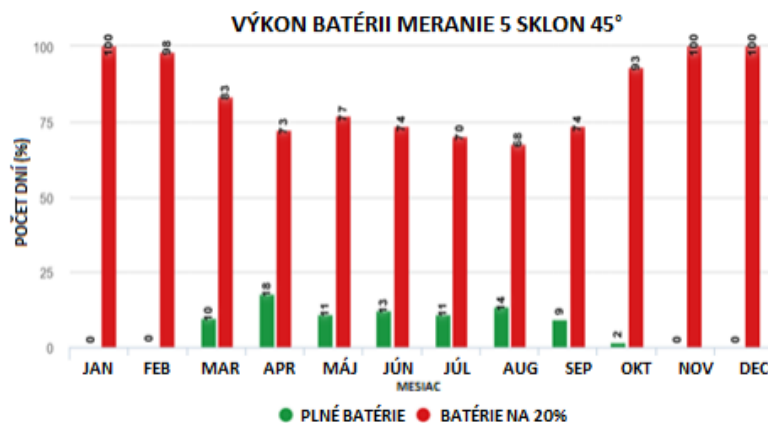
Meraním sa zistilo, že priemerná produkcia energie za deň nepostačuje na pokrytie dennej spotreby. Nízky inštalovaný špičkový výkon, vysoká denná spotreba a nízka akumulácia schopnosť spôsobuje, že model A nie je schopný ostrovnej prevádzky. V mesiacoch marec až september, dosahuje najvyššiu priemernú dennú produkciu 30° sklon na hodnote 8622 Wh. V daných mesiacoch 45° sklon vyrobí o 178 Wh denne energie menej ako 30° sklon na hodnote 8444 Wh. Sklon 60° vyrobí o 705 Wh denne energie menej ako 30° sklon na hodnote 7917 Wh. Na základe týchto zistení je možné povedať, že 30° sklon optimalizuje výrobu v jarnej a letnej sezóne. V mesiacoch október až február dosahuje najvyššiu priemernú dennú produkciu 60° sklon na hodnote 4226 Wh. V daných mesiacoch 30° sklon vyrobí o 496 Wh denne energie menej ako 60° sklon na hodnote 3730 Wh. Sklon 45° vyrobí o 138 Wh denne energie menej ako 60° sklon na hodnote 4087 Wh. Na základe týchto zistení je možné povedať, že 60° sklon optimalizuje výrobu v jeseni a v zime. Najviac vyrobenej energie za mesiac 605,6 kWh a najvyššiu ročnú výrobu 7266,9 kWh dosahuje model B pri použití 45° sklonu panelov. Z hľadiska potenciálu nezachytenej energie dominuje 30° sklon. Potenciál nezachytenej energie už nie je taký výrazný ako pri modeli A, preto je potrebné doplniť jednotlivé zostavy o ďalšie solárne panely. Model B potrebuje v najlepšom prípade ďalších 5511,7 kWh energie na úplnú sebestačnosť. Sebestačnosť sa nedosiahne ani v prípade montáže solárneho systému na každý dom, v takom prípade chýba modelu 667,1 kWh energie

Tabuľka 6
Vplyv sklonu panelov na výrobu elektrickej energie porovnanie meraní 4-5-6 [4]

Meranie	4	5	6
Priemerná mesačná výroba model (kWh)	601,6	605,6	582,6
Ročná výroba model (kWh)	7219,5	7266,9	6991,3
Chýbajúca energia model za rok (kWh)	5559,2	5511,7	5787,3
Nezachytená energia model za rok (kWh)	171,1	130,6	44,6



Obr. 4 Meranie 5 odhad výroby energie sklon panelov 45° [4]



Obr. 5 Meranie 5 výkon batérií sklon panelov 45° [4]

V. ZÁVER

Hlavným cieľom článku bolo navrhnúť hybridný model merania, vhodne zvoliť vstupné veličiny merania, opísať postup pri meraní modelu a porovnať jednotlivé zostavy z hľadiska vplyvu sklonu panelov. Výsledky ukázali že zmena jednotlivých parametrov na vstupe ovplyvní získanú energiu na výstup. Takýmto približným spôsobom je možné optimalizovať vstupné veličiny, čím sa dosiahne maximálna efektívnosť modelu merania. V súčasnosti je výstavba takéhoto modelu nevýhodná, čo spôsobuje hlavne efektívnosť FV panelov, ktorá je pre komerčné účely na nízkej úrovni. Zvýšenie efektivity by prinieslo viac záujmu zo strany investorov.

POĎAKOVANIE

Tento príspevok vznikol vďaka Agentúre na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0576 a APVV-21-0312, ako aj Vedeckou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu SR a SAV, podporená grantom VEGA 1/0757/21.

ZOZNAM POUŽITEJ LITERATÚRY

- [1] EU Science Hub, „PVGIS Online Tool“, [online]. [cit. 2023-02-27]. Dostupné na internete: https://joint-research-centre.ec.europa.eu/pvgis-online-tool/getting-started-pvgis_en
- [2] EU Science Hub, „PVGIS data sources and calculation methods“, [online]. [cit. 2023-02-28]. Dostupné na internete: https://joint-research-centre.ec.europa.eu/pvgis-online-tool/getting-started-pvgis/pvgis-data-sources-calculation-methods_en
- [3] R. Müller, T. Behrendt, A. Hammer, A.A. Kemper, „A new algorithm for the satellitebased retrieval of solar surface irradiance in spectral bands“, [online]. Remote Sens. 2012, 4, 622–647 [cit. 2023-03-02]. Dostupné na internete: <https://www.mdpi.com/2072-4292/4/3/622>
- [4] K. Haborák, „Vplyv prosumerov na distribučnú sieť“. Diplomová práca. Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2023.

Budúcnosť vodíkových technológií

Martin KRIŠŠÁK, Dušan MEDVEĎ

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

martin.krissak@student.tuke.sk , dusan.medved@tuke.sk

Abstrakt — Tento vedecký článok popisuje analýzu využitia vodíkových technológií v súčasnosti, ako aj v budúcnosti. Prvá časť článku je venovaná informáciám, ktoré o vodíku poznáme. Sú tu popísané základné poznatky ako sú: akým spôsobom je možné vyrobiť, uskladniť a transportovať vodík, načo slúžia palivové články a rozdiely medzi nimi. V druhej časti je vytvorená analýza využitia vodíka. Ak si niekto predstaví, kde by mohol byť vodík ako prvý využiteľný, tak si pravdepodobne predstaví automobily. V tejto analýze bolo uvedené porovnanie súčasného stavu, to znamená porovnaní spaľovacích a elektrických áut. Záver analýzy tvorí grafické porovnanie, prečo sú elektrické autá krokom vpred, čo ešte nie je dokonalé a čo by ich mohlo v budúcnosti nahradiť.

Kľúčové slová — vodík, palivové články, akumulácia elektrickej energie

I. ÚVOD

V dnešnej dobe, plnej moderných technológií a výdobytkov 21. storočia, si čoraz viac uvedomujeme aj náš dopad na životné prostredie. Krajiny celého sveta vytvárajú ročne obrovské množstvo emisií CO₂, čo negatívne pôsobí na životné prostredie. Podľa zozbieraných údajov z roku 2020 sa celosvetovo vytvorí najviac emisií práve v odvetviach priemyselnej výroby, energetiky a dopravy. [1]

Množstvo emisií v doprave zohráva kľúčovú úlohu v celosvetovom rozsahu. EÚ každým rokom sprísňuje právne predpisy v rôznych odvetviach, vrátane sektora dopravy, z dôvodu zníženia množstva vyprodukovaných emisií, ktoré negatívne pôsobia na životné prostredie. [1]

Z týchto dôvodov sa uvažuje o rôznych alternatívnych náhradách, ktoré by mohli pomôcť znížiť množstvo emisií a splniť tak požadované ciele. V prípade možných alternatív sa najčastejšie uvažuje o využití vodíka ako nosiča energie.

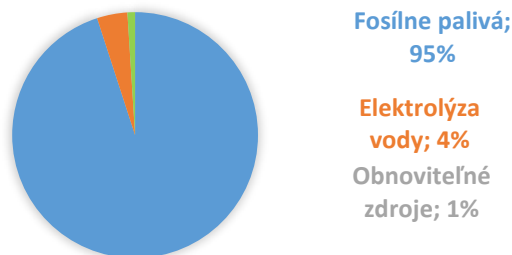
II. VÝROBA, USKLADNENIE A TRANSPORT VODÍKA

A. Výroba vodíka

Výrobu vodíka vieme uskutočniť viacerými možnými spôsobmi, či už z fosílnych zdrojov energie ako sú uhlie, ropa a zemný plyn, ako aj pomocou využitia obnoviteľných zdrojov energie, akými sú energia vody, slnka, vetra a biomasy či v neposlednom rade pomocou využitia jadrovej energie.

V súčasnej dobe sa výroba vodíka realizuje predovšetkým z fosílnych zdrojov energie a to v podobe reformingu zemného plynu, ľahkých uhľovodíkov ako aj splyňovaním uhlia. V oveľa menšom množstve sa vyrába vodík pomocou elektrolýzy vody ako aj pomocou využitia obnoviteľných zdrojov energie.

SÚČASNÁ SVETOVÁ PRODUKČIA VÝROBY VODÍKA

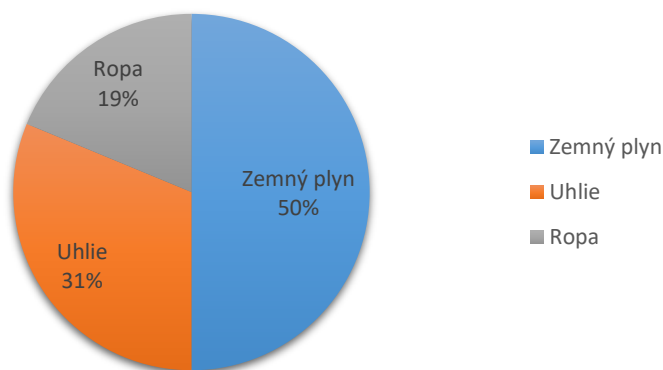


Obr. 1 Celosvetová produkcia výroby vodíka [2]

Ako aj môžeme vidieť podľa Obr. 1, ktorý popisuje súčasný svetový stav výroby vodíka vo svete, tak až v 95% sa vodík vyrába práve pomocou fosílnych palív, v 4% elektrolýzou vody a iba 1% pomocou využitia obnoviteľnej energie. [2]

Vodík je najviac zastúpený chemický prvok v celom vesmíre. [3] Vodík nie je zdroj, ale je energetickým nosičom. Keďže vodík sa v elementárnej forme skoro nenachádza, tak z toho dôvodu ho nevieme ani lacno ťažiť. Vodík sa nachádza v zásade vždy ako súčasť iného prvku a preto je potrebné ho z neho odobrať, vďaka čomu musíme rátať aj s nejakými stratami. Vodík nie je možné použiť ako hlavný zdroj energie, ale je možné konštatovať, že vodíkové technológie sú emisne tak čisté, aký typ hlavného zdroja energie využijeme na jeho výrobu.

Využitie fosílnych palív na výrobu vodíka



Obr. 2 Využitie fosílnych palív na výrobu vodíka [2]

Aby sa mohol stať vodík plne hodnotnou alternatívou do budúcnosti k fosílnym palivám, tak musíme nájsť iný vhodne ekologicky čistý zdroj energie, ktorý je nevyhnutný na jeho výrobu. V súčasnosti, ako môžeme vidieť aj podľa Obr. 2, je na výrobu vodíka fosílnymi palivami využívaný v 50% zemný plyn, v 31% uhlie a z 19% ropa. [4]

B. Uskladnenie vodíka

Skladovania vodíka je veľmi dôležité z hľadiska vodíkoveho hospodárstva. Vodík je potrebné skladovať v mieste výroby, ako aj v mieste spotreby. V dnešnej dobe sa len veľmi malé percento vodíka skladuje. Prevažná väčšina sa rovno spotrebovávajú v rámci výrobného procesu bez toho, aby sa skladoval. Príkladom takéhoto procesu môže byť napríklad výroba vodíka zo zemného plynu. Takto získaný vodík sa ihneď využíva na syntézu amoniaku. V súvislosti s vodíkovou energetikou je otázka skladovania vodíka čoraz aktuálnejšia a dôležitejšia. [6]

Uskladnenie vodíka je kľúč k zavedeniu vodíka k praktickým využitiam. Aby sa vodík vedel stať konkurencieschopným energetickým médiom, tak by bolo potrebné zefektívniť náklady na celý systém, prevádzku, zvýšiť životnosť, opotrebenie ako aj montáž tak, aby to bolo prijateľné pre danú aplikáciu. [5]

Vodík možno skladovať prevažne tromi hlavnými spôsobmi: v plynnom stave, kvapalnom stave a skladovanie vodíka viazaného vo forme hydridov alebo iných chemických zlúčenín. [6]

C. Transport

Využitelný vodík sa nachádza v plynnom alebo kvapalnom stave. Vodík skvapalňuje pri teplote nižšej ako je bod varu 20,39K (-252,76°C). V kvapalnom stave je možné uskladniť vodík s hustotou 70,79 kg/m³ pri normálnom tlaku.

Problémom skladovania vodíka v kvapalnom stave je nízka teplota. Prekročením teploty bodu varu môže dôjsť k vyparovaniu vodíka a preto sa využívajú špeciálne izolované kryogénne nádoby. [7]

Naopak, plynný vodík je možné distribuovať v zásade dvoma spôsobmi: vysokotlakovými kontajnermi alebo potrubím. Preprava pomocou využitia vysokotlakových nádrží môže byť zjednodušená pomocou cestnej, železničnej alebo námornej dopravy. To môže výrazne pomôcť k zavedeniu vodíka ako nosiča energie z dôvodu, že nebude potrebné budovať novú infraštruktúru. Naopak, prenos vodíka potrubím môže byť dobrým spôsobom ako distribuovať vodík vo väčších množstvách, akurát bude potrebné urobiť niekoľko úprav z dôvodu, že vodíkové skrehnutie urýchľuje tvorbu trhlín a tým výrazne skracuje životnosť systému. [8]

III. PALIVOVÉ ČLÁNKY

Palivový článok je elektro-chemické zariadenie, ktoré slúži na premenu vodíka na energiu elektrickú. [9]

Palivové články sa delia na základe viacerých kritérií, ako napríklad podľa typu elektrolytu, prevádzkovej teploty či praktického využitia.

Na základe týchto charakteristík sa ovplyvňuje, pre ktoré využitie budú dané palivové články najvhodnejšie. V súčasnosti existuje niekoľko rôznych typov palivových článkov. Každý z nich je na rôzne využitie, niektoré je výhodnejšie využiť v interiéri, iné len v exteriéri. Každý z nich má svoje výhody i nevýhody.

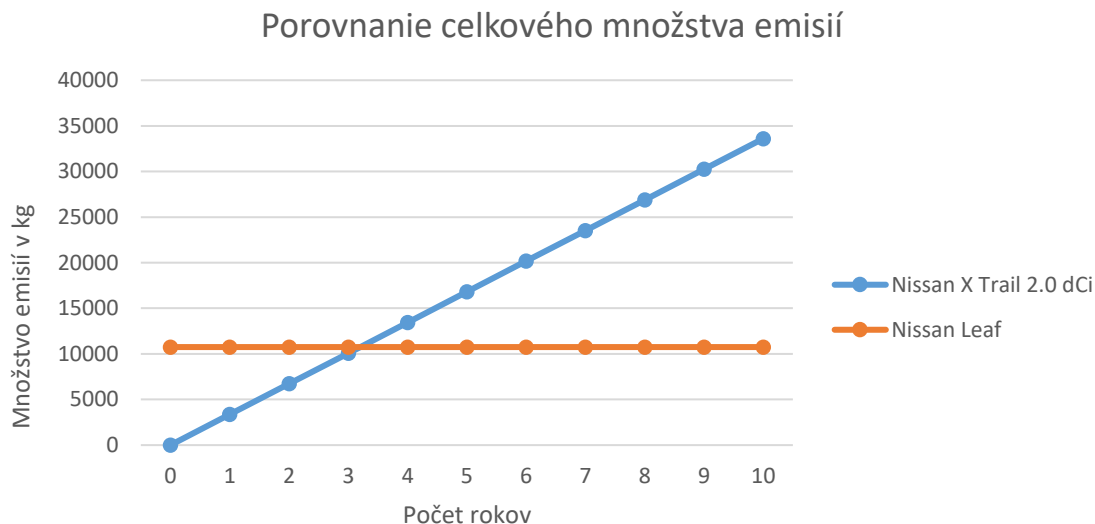
Na základe týchto kritérií rozlišujeme palivové články nasledovne [9]:

Druh palivového článku	Spôsob použitia	Teplota článku; účinnosť	Výhody	Nevýhody
P.č. s polymérnou membránou (PEMFC)	Pre prenosné a stacionárne použitie, pre automobilový priemysel	50 - 100°C; účinnosť 35 - 60 %	Kompaktné prevedenie, dlhá životnosť, rýchly nábeh, pokročilý vývoj	Drahý katalyzátor; potrebuje palivo vysokej chemickej čistoty; zložitá regulácia teploty a množstvo vody.
Alkalický p.č. (AFC)	Vo vesmíre, vojenstve, v ponorkách, v doprave.	90 - 100°C; účinnosť 60 %	Nízke náklady na diely a prevádzku; žiadny kompresor; vysoká reakčná rýchlosť katódy.	Sú veľké; citlivé na znečistenie vodíka a kyslíka.
P.č. s tavenými uhlíčitami (MCFC)	Pre dosiahnutie veľkého výkonu.	600 - 700°C; účinnosť 45 - 50 %	Vysoká účinnosť, nie je viazaný na jedno konkrétne palivo, kogeneračné.	Vysoká teplota spôsobuje koróziu; dlhý nábeh, krátka životnosť.
P.č. s kyselinou fosforečnou (PAFC)	Pre dosiahnutie stredného až veľkého výkonu.	150 - 200°C; účinnosť 40 %	Dobre znáša prítomnosť nečistôt v palivu; kogeneračné.	Nízka účinnosť, obmedzená životnosť; drahý katalyzátor.
P.č. s tuhými oxidmi (SOFC)	Pre dosiahnutie stredného až veľkého výkonu.	700 - 1000°C; účinnosť 60 %	Bez veľkých nárokov na palivo; môže používať zemný plyn, vysoká účinnosť.	Vysoká teplota spôsobuje koróziu; dlhý nábeh, krátka životnosť.
P.č. pre priamu reakciu metanolu (DMFC)	Pre prenosné, mobilné a stacionárne použitie.	40 - 60°C; účinnosť 20 %	Kompaktný; beží na metanol; žiadny kompresor.	Usporiadanie p.č. do súboru je zložitá; pomalá odozva; nízka účinnosť.

IV. ANALÝZA SÚČASNÉHO STAVU

Analýza súčasného stavu porovnáva množstvo vytvorených emisií bežným spaľovacím autom a elektrickým autom. Pre túto modelovú úlohu bol aj vytvorený návrh ostrovnej fotovoltaickej elektrárne pre dobíjanie elektromobilu.

Z analýzy vyplýva, že spaľovacie autá sú väčším zdrojom emisií ako elektrické autá a to aj pri zohľadnení výroby a recyklácie akumulátorov. V analýze bol použitý Nissan X Trail, ktorý pri prevádzke a najazdení 200 000km vytvorí 33 600kg CO₂. Elektrické auto Nissan Leaf vyrobí emisií 10 748,4 kg CO₂. Z toho vyplýva, že pri tomto návrhu je elektrické auto 3-krát šetrnejšie.



Obr. 3 Porovnanie celkového množstva vyrobených emisií spaľovacím autom Nissan X Trail 2.0 dCi a elektrického auta Nissan Leaf

Elektrické autá sa na prvý pohľad vidia ako riešenie problému so znečistením prostredia. Ako je vidieť z tejto analýzy, elektrické autá priamo neprodujú počas prevádzky emisie CO₂. Elektrické autá prichádzajú už s tzv. uhlíkovým dlhom. V prípade tejto analýzy sa uhlíkový dlh vyrovná po 3 rokoch, resp. 60 000km. Od tohto momentu sú elektrické autá šetrnejšie v porovnaní s autami so spaľovacími motormi.

Ako je zrejmé z tejto analýzy, elektrické autá sú krok vpred pri šetrení emisií v doprave. Zároveň, ale ešte stále majú čo zlepšiť a vytvárajú príliš veľké množstvo emisií. Riešením tohto problému by mohol byť vodík, keďže pri svojej výrobe, skladovaní a preprave nie je zdrojom emisií.

V. ZÁVER

Budúcnosť zavedenia vodíka ako nosiča energie je závislá od viacerých faktorov ako sú: cenová dostupnosť, prevádzkové náklady, bezpečnosť, účinnosť celého systému ako aj konkurencieschopnosť.

Vodík sa bežne nevyskytuje ako chemicky čistý. Pri vzniku sa ihneď spája do molekúl. Pre využitie vodíka je preto potrebná energia pre rozdelenie.

Z tohto dôvodu sa uvažuje o využití energie, ktorá je dostupná tzv. zadarmo a nie je potrebné ju vyrábať. Ide predovšetkým o obnoviteľné zdroje, ako sú slnečná a veterná energia. Vodík má viaceré predpoklady pre úspešné zavedenie ako nosiča energie a úspešné vytvorenie tzv. bezuhlíkovej dopravy.

POĎAKOVANIE

Tento príspevok vznikol vďaka Agentúre na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0576 a APVV-21-0312, ako aj Vedeckou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu SR a SAV, podporená grantom VEGA 1/0757/21.

LITERATÚRA

- [1] Z. Kamenská, „*Emisie z dopravy sú veľký problém! Ako ich môžeme znížiť?*“, [online]. [cit. 11.05.2023]. Dostupné z: <https://europskydialog.eu/2022/04/20/emisie-z-dopravy-su-velky-problem-ako-ich-mozeme-znizit/>
- [2] V. Kafková, D. Karas, V. Cyprichová, „*Vodík*“, 1. vyd. Leopoldov: Centrum výskumu a vývoja, 2021. 196 s. ISBN 978-80-973973-0-2.
- [3] Toyota, „*Vodík*“, [online]. [cit. 11.05.2023]. Dostupné z: <https://www.toyota.sk/elektrifikovany/vodik/vodik>
- [4] N. Jasmínská, T. Brestovič, M. Lázár, „*Výroba a uskladnenie vodíka*“, 1. vyd. Košice: TU, 2015. 135 s. (Edícia vedeckej a odbornej literatúry). ISBN 978-80-553-2378-7.
- [5] N. Jasmínská, „*Výroba vodíka z alternatívnych zdrojov a jeho akumulácia*“, tézy habilitačnej prednášky k habilitačnému pokračovaniu za docenta v odbore 5.2.6 Energetické stroje a zariadenia. 1. vyd. Košice: TU, 2017. 37 s. ISBN 978-80-553-3130-0.
- [6] J. Balajka, „*Vodík a iné nové nosiče energie*“, Bratislava: Alfa, 1982.
- [7] T. Brestovič, „*Výskum a vývoj efektívnej kompresie a uskladnenia vodíka pomocou nových metalhydridových zliatin*“, tézy inauguračnej prednášky k vymenúvaciemu konaniu za profesora v odbore 5.2.6 Energetické stroje a zariadenia. 1. vyd. Košice: TU, 2019. 53 s. ISBN 978-80-553-3370-0.
- [8] TÚV SÚD, „*Vodík umožňuje prepojenie čistej energie a jej využitia*“, [online]. [cit. 11.05.2023]. Dostupné z: <https://www.tuvsud.com/sk-sk/temy/vodik/spoznajte-vodikovy-hodnotovy-retazec/skladovanie-a-distribucia-vodika>
- [9] P. Štubařová, „*Jak fungují palivové články?*“, [online]. [cit. 11.05.2023]. Dostupné z: <https://elektro.tzb-info.cz/elektromotory-pohony-a-stroje/16987-jak-funguji-palivove-clanky>
- [10] M. Kriššák, „*Budúcnosť vodíkových technológií*“. Bakalárska práca. Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2023.

Transforming a Gantt Chart Into an Equivalent Hierarchical Data Structure

¹Lívia KOVÁČOVÁ, ²Branislav MADOŠ

^{1,2} Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Kosice, Slovak Republic

¹livia.kovacova@student.tuke.sk, ²branislav.mados@tuke.sk

Abstract — The paper describes a view of the development process of a web application whose main aim is to provide the possibility of transforming a Gantt chart into an equivalent hierarchical data structure. The final implementation of such software supports working with a Gantt chart and creating its transformation into an equivalent hierarchical data structure. Prior to the implementation of this software, a more detailed analysis of the current state of software tools for project management was carried out by comparing the features they support. The possibilities of the Gantt chart for visualizing the sequence of activities in project management were also analyzed. As part of this analysis, the concept of the project and project management was introduced. Based on the analysis, a solution in the form of a web application was designed, implemented, tested, and evaluated. The server part of the implementation was built using the Spring Boot framework, which communicates with a PostgreSQL relational database. On the client side, the React library was used in combination with other, smaller libraries to partially support visualization of charts and diagrams. The design was created mainly using the Material UI library, with necessary customizations made using CSS style sheet language. The output product is a fully functional web-based application that offers the possibility of generating a Gantt chart, editing it, and creating a dependency diagram that represents a hierarchical data structure equivalent to a Gantt chart.

Keywords — Gantt chart, java, project management, react, web application

I. INTRODUCTION

Project management is a crucial part of any project implementation, because it represents a significant area dedicated to managing a range of specialized activities that need to be executed during the project lifecycle. Project management primarily provides flexible planning and management of project processes and resources, which ultimately significantly affects the quality, cost-effectiveness, and resource utilization of the project, thereby influencing its overall success.

Software tools for project management significantly ease otherwise time-consuming activities related to project management. They accompany the project management process from initial planning of project activities, defining their characteristics, through monitoring project progress, to modification and optimization of project activities in line with the project's nature and requirements.

An essential component of these tools is the Gantt chart - one of the fundamental supportive means of project management. The broad range of possibilities that support project management, from its initiation to project finalization is the main reason for its popularity. Its effective utilization leads to the successful completion of projects with a high degree of success. Gantt chart is considered as one of the earliest project management tools and provides various options for visualizing the sequence of project activities. Its success is attested by the fact that it was integrated as a planning tool in production as early as the 1920s [1] and remains a popular project management tool to this day.

Given the high potential of the Gantt chart in project management, a web application has been developed to support its creation and transformation into an equivalent hierarchical data structure - the dependency diagram - providing a closer and more comprehensive view of the sequence and parameters of project activities associated with the available project resources. The resulting software solution offers user-friendly and understandable functionalities for working with the Gantt chart as well as its transformation. This paper describes the key features of the development process and the resulting web application.

II. PROJECT AND PROJECT MANAGEMENT

A. Project

In project management, there are various definitions of the term "project," although most of them concur in using terms such as resources, project objectives, requirements or time.

Standard for project management system defines a project as a unique process consisting of a series of coordinated and controlled activities with start and end dates, carried out to achieve a specific goal that meets specific requirements, including constraints imposed by time, cost, and resources [2].

The concept of a software project author in [3] defines as the execution of activities for specific user requirements, where these activities and their sequence are specified by the software process in the project plan. According to [3], a software project is also characterized as a time-limited effort that is undertaken to create a unique outcome represented by the terms of the project agreement, whereas the fulfillment of the agreement is provided through a set of technical and management activities.

B. Project Management

The subject of project management is precisely the management activities providing coordination of the development activities of the project defined above. According to authors in [4], project management is the planning, organizing, directing, and controlling of company resources for a relatively short-term objective that has been established to meet specific goals and objectives.

It is necessary to clarify the concept of project management in the context of software engineering too. In this regard, it is possible to speak of software development project management (SDPM). Wysocki in [5] defines SDPM as a discipline that involves approaching the characteristics of the software to be developed, selecting the most appropriate software development life cycle, and subsequently selecting a suitable approach to project management while ensuring the most effective provision of customer requirements to provide business value.

III. GANTT CHART

As mentioned before, the Gantt chart represents a significant part of software tools used in project management. The study [6] also lists it as one of the tools for successful and effective project planning, with versatile applications, particularly for visualizing the project's timeline.

Regarding its origins, according to the authors in [1], the breakthrough year for the Gantt chart was 1920. During this time, the Gantt chart was used as a tool for production planning but was not directly associated with project management. Since then, the Gantt chart has undergone several phases of developmental changes, from its application in serial production to its adoption in projects and project management. Currently, it constitutes an indispensable component of planning, control, and monitoring in any project, including software projects.

The Gantt chart is a type of bar chart that represents activities or events plotted against time. Activities refer to the amount of work to be completed from one point in time to another, while events represent the starting and ending points of these activities or their groups [7]. It is also possible to add any number of attributes to each activity, providing a more detailed description of its nature and progress. The main components of a Gantt chart include a list of activities, their durations, scheduled dates, as well as records of the progress made [7].

With a focus on the practical application of the Gantt chart, it is possible to define additional properties for individual development activities as well as the entire chart. Some of these properties include relationships to other activities, current activity status, activity completion percentage, resources required for activity, higher-level units for activity groups or vertical line indicating the current date. These features enhance the Gantt chart's practicality by capturing the relationships between activities, tracking their progress, providing real-time status updates, managing resource allocation, organizing related activities into groups, and visually representing the current date on the chart.

The use of Gantt chart in project management needs to be carefully considered, like any other project management tool, considering the nature of the project and the benefits and limitations of the Gantt chart. According to author in [8], some of the benefits of Gantt chart include providing a visual representation of the project's complexity, interactive visualization of the project plan, and transparency of the work environment by displaying project activities and resources. On the other hand, potential limitations may arise, such as the complexity of creating the chart for complex projects, which can result in challenges when updating the chart in response to project changes.

IV. SOFTWARE TOOLS FOR PROJECT MANAGEMENT

In order to design the web application as effectively as possible, software tools for project management were analyzed. These tools offer diverse functionalities for project management methodologies, supporting the digitization of processes such as diagram creation, task allocation, resource allocation, project status monitoring, project version management, and communication within project teams. During the analysis process, several software tools for project management were compared in terms of provided functionality, technical features, and specific characteristics.

The compared software tools included MS Project, Zoho Projects, Redmine, Wrike, and Planview. Each software tool offers nearly the same functionality, but their support depends on the chosen version of the tool. Some features are only available in paid or enterprise versions of the respective software. Key features include cross-platform support, Gantt chart creation support, critical path visualization, milestone tracking, file sharing, reports generation, project resource management, team management, team communication, and support for different user roles.

V. DESIGN AND IMPLEMENTATION OF WEB APPLICATION

The developed solution aims to be a software tool that supports a range of system requirements, including the creation, editing, and visualization of a Gantt chart as a means of project management. It also includes the creation and display of a transformation from the project's Gantt chart to a dependency diagram, which represents a hierarchical data structure equivalent to the Gantt chart. To fulfill these requirements, the application's functionality had to be implemented, including user registration, user login, project creation, display of user projects, creation of the Gantt chart by creating project phases and their tasks, editing Gantt chart generated based on the definition of project phases and tasks, visualization of the project's Gantt chart, generation of the dependency diagram, and display of the dependency diagram.

As it is a project management tool, it is necessary for it to be available in real-time for all members of the project team. For this reason, a web application was the most suitable approach for the solution. Like most web-based solutions, this particular solution consists of a server-side and a client-side component too. The server part handles the manipulation of data that will be stored within the PostgreSQL relational database. The client-side is responsible for presenting data to the user and handling user inputs required for creating instances and components of the web application.

A. *Web application solution concept*

The implementation of the solution followed a design process that involved describing various concepts of the developed software. Among the key concepts that needed to be described and defined were the concept of the Gantt chart task along with its parameters, the concept of Gantt chart phase, the concept of the dependency diagram representing a hierarchical data structure equivalent to the Gantt chart, and the concept of user roles which will utilize the software functionalities.

As mentioned, the task in the Gantt chart is defined by specific parameters, which can be divided into two groups: basic task parameters and additional task parameters. Among the basic task parameters that are part of the solution and necessary for task definition, the task identifier and duration are included. Parameters with informational character were also added to the task definition, including the task description, human and financial resources, and task status. The additional task parameters allow addressing possible logical issues during chart rendering. Relationship dependencies with other tasks and task priority are among the added task attributes.

The concept of Gantt chart phases enhances the clarity of the chart, particularly in projects with a large number of tasks. By visual representation of the project's phases, they provide an overview of the project and the sequence of project phases from a broader perspective. They also play a role in the visualization of the Gantt chart transformation, specifically in the dependency diagram.

The transformed diagram, called the dependency diagram, serves as a hierarchical structure equivalent to the Gantt chart. Its main purpose is to provide the user with a clearer representation of tasks in terms of their relationships and dependencies, with tasks being clearly depicted as part of project phases. Through the dependency diagram, it is possible to track the tasks that need to be completed to transition to the next project phase and monitor the utilization of human and material resources throughout the different phases of the project. This is achieved by displaying relevant information related to the tasks included in the dependency diagram.

The project is executed by a project team consisting of members with different competencies, who are subject to the project's organizational structure. Therefore, the concept of user roles was

introduced in the final solution, specifically two specific roles: the project manager and the project team member. The project manager is an individual who defines the characteristics of the project and its project tasks that the Gantt chart contains, as well as its transformation. In the solution, the project manager has possibility of actions to create a project, create a Gantt chart, and a dependency diagram by defining the parameters of project tasks. The project team member is a person assigned to both the project and its individual tasks by the project manager. As project members, they have access to the diagrams for viewing as soon as they are created by the project manager.

B. Server part structure

The server-side of the application is built on the Spring Boot framework in collaboration with the PostgreSQL relational database, which together provide the fundamental functionalities of the application from a backend implementation perspective.

The application's entities, services, controllers, and endpoints were implemented to facilitate communication and services to the client-side of the application.

Regarding the database, the individual entities were created using the Spring Data JPA interface. Access to the entities in the database is ensured by *repository* classes, which are also part of the Spring Data JPA interface. They support standardized CRUD (Create, Read, Update, Delete) operations.

The communication between the server-side and client-side is ensured through HTTP REST requests to individual REST API endpoints. When creating individual endpoints, it was necessary to consider the relationships between the entity being requested and other entities of the application. Almost all service components and entities of the application have REST controllers created, which call methods from the created Spring Framework service services. These services were created to separate the layer of handling HTTP requests from the business logic layer of the application. The architecture of the developed web application can be seen in Fig. 1. This type of architecture ensures separation of different functional parts and enhances modularity, reusability, and maintainability.

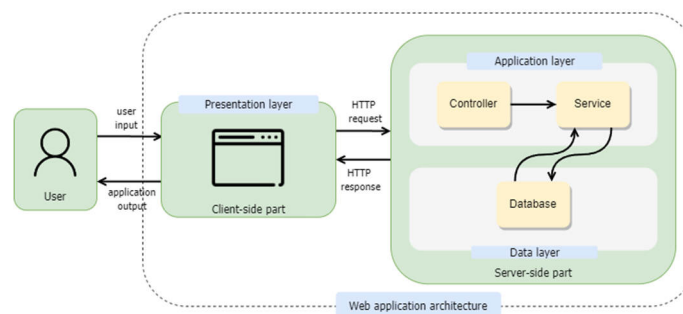


Fig. 1 Architecture of the developed web application. Architecture follows a multi-layered client-server model of architecture.

C. Security

The security of the developed application, which represents a project management tool, is extremely important, as it involves various types of sensitive information about the companies that will use this tool in the future. The implemented security mechanisms support security restrictions, including the fact that any user can register in the application, but during registration, the user's role as either a project manager or a team member is defined. Additionally, only registered users are allowed to log into the application, and only users with the project manager role can create projects and diagrams.

These restrictions were enforced through several specific steps implementing security mechanisms. The first step involved creating a Role entity in which records for the manager role and the team member role are stored.

Next step is storing user passwords in the encrypted form in the database. This was achieved using the *PasswordEncoder* interface provided by the Spring Security framework. Of course, criteria were established for the format of user passwords to minimize their predictability.

The Spring Security framework, in conjunction with JSON Web Token (JWT) authentication technology, provided the capability for authentication in the application. The authentication principle is based on the fact that a logged-in user possesses a JSON Web Token, which verifies the validity of the user's login because this token is generated by the server-side upon logging

into the system. Possessing JSON Web token is important for verifying the origin of client requests to the server and for accessing endpoints that are restricted to authorized requests.

D. Client part structure

The client-side of the application was built using the React library, which provides a user interface development framework for web applications. The principle behind creating this part involved creating React components of the application and client-side services that facilitate communication with the server-side of the application.

The concept of the created web application is single-page. For each component that needs to be displayed on the client-side, *React Routes*, which are part of an important component called *Routes*, have been created. *Routes* component allows the rendering of the desired component based on the defined path in *Route*.

Regarding the architecture, we can talk about two significant groups of entities that form the client-side. These are components and services. Application components provide the presentation of the application's content, which can vary depending on the application's state, which changes based on user actions or data obtained from the server-side of the application. Components representing individual subpages of the web application include specific and shared components that together create the content of the parent component.

The services of the client-side application communicate with the server-side of the application. This communication, which involves sending and processing HTTP requests to the server, is facilitated by using the *axios* library. These services provide functionality for actions such as registration, login, and retrieving information about application entities such as users, projects, project tasks and phases, and diagrams.

In terms of component design, and ultimately the entire application's design, components from the *Material UI* library were used to achieve a unified look of the user interface. These components offer an intuitive and responsive design with modern elements.

E. Gantt chart and dependency diagram visualization

The visualization of the Gantt chart and dependency diagram is provided by components utilizing the *react-gantt-task* and *react-flow* libraries, which have been modified for the needs of the implemented application.

The visualization process consists of several steps. The data representation of the Gantt chart is generated by the server-side of the application based on user inputs when defining project phases and tasks. In the client-side, the generated data representation of the Gantt chart is visualized using a react component that includes modified and logically arranged components from the *react-gantt-task* library. This visualization is obtained by making an HTTP request to the server-side to retrieve the generated data.

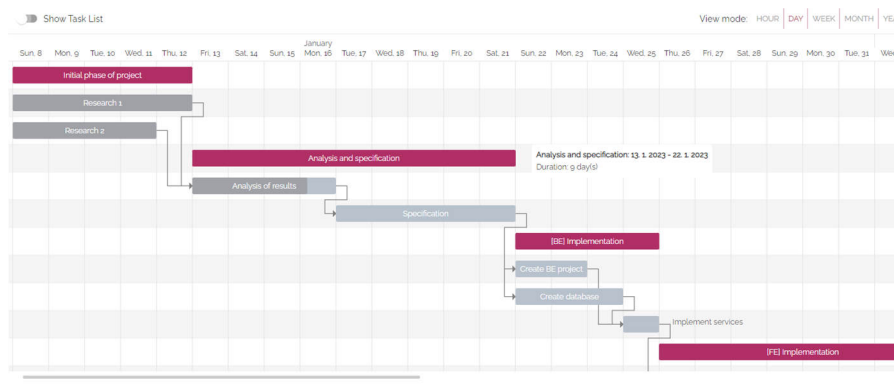


Fig. 2 Visualization of Gantt chart

The data representation of the transformation of Gantt chart into a dependency diagram is performed on the client-side of the application. The process involves executing a recursive function to create and configure individual components of the dependency diagram, which is based on the project phases from the previously generated Gantt chart. These components contain data structures for Gantt chart tasks, including task identifiers and an array of predecessors. These attributes are crucial for generating the transformation into an equivalent hierarchical data structure, which the dependency diagram represents. The visualization of the Gantt chart and its transformation into a dependency diagram can be seen in Fig. 2 and Fig. 3.

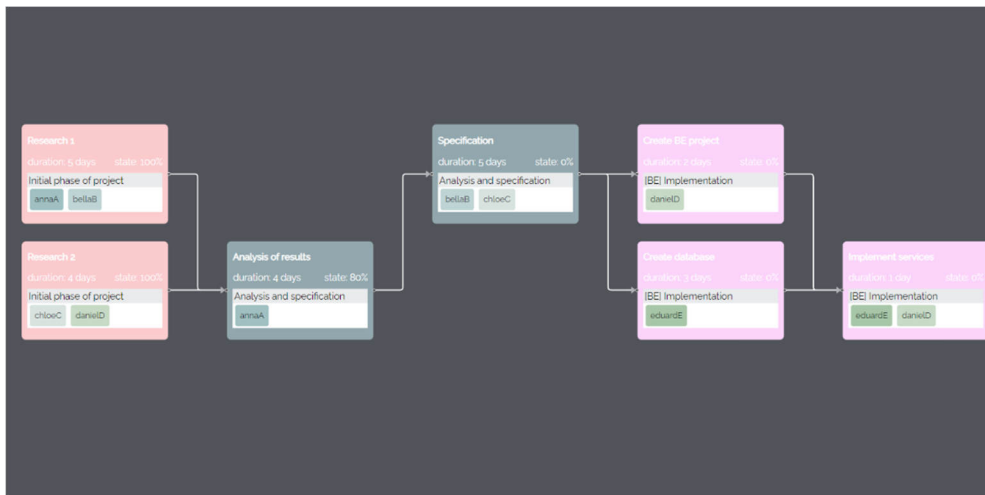


Fig. 3 Visualization of dependency diagram

VI. RESULTS

The final state of the implemented solution was obtained through testing. During the testing of the application, various phases and project tasks were defined as input for creating both the Gantt chart and the dependency diagram. The main changes in terms of defining the parameters of project phases and tasks included predecessors, duration, and team members assigned to work on tasks. It was also important to test different numbers of project tasks and phases that serve as input for creating the Gantt chart and subsequently the dependency diagram. In all tested cases, the application was able to generate the diagrams correctly and produced the expected results.

The solution was also tested in terms of the time complexity of the Gantt chart generation method. In this function, the tasks in the Gantt chart are ordered based on the phases that contain project tasks and the project's start date defined during project creation. The results indicate that the generation time of the chart significantly increases with the number of created tasks, which can be a challenge for complex projects. However, this issue is already resolved by supporting the grouping of tasks into larger units, namely project phases.

VII. CONCLUSION

The developed solution is a web application that can be expanded and improved with additional functionalities. Currently, it supports the creation and management of Gantt charts, as well as the transformation of the Gantt chart into an equivalent hierarchical data structure in the form of a dependency diagram. There are several possibilities for further enhancements, including adding features to facilitate user communication, enhancing interactivity in the visualization of the dependency diagram, supporting user membership in specific organizations, and optimizing the method for generating Gantt chart.

ACKNOWLEDGEMENTS

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: 313011W988), co-financed by the ERDF.

REFERENCES

- [1] WILSON, James M. "Gantt charts: A centenary appreciation", In: *European Journal of Operational Research*, vol. 149, no. 2, pp. 430–437, 2003. ISSN 0377-2217. Available at: doi: [https://doi.org/10.1016/S0377-2217\(02\)00769-5](https://doi.org/10.1016/S0377-2217(02)00769-5).
- [2] STN ISO 10006:2019. Quality management - Guidelines for quality management in projects. In: *Bratislava: Office for Standardization, Metrology, and Testing of the Slovak Republic*, 2019.
- [3] BIELIKOVÁ, Mária. Software Engineering: Principles and Management. In: *Slovak University of Technology*, 2000. ISBN 8022713228.
- [4] KERZNER, H. Project Management. In: *New York: Van Nostrand Reinhold*, 1986.
- [5] WYSOCKI, R.K. Effective Software Project Management. In: Wiley, 2010, chap. *The Changing Landscape of Software Development*, p. 7. Wiley Desktop Editions. ISBN 9780470446539. Available at: <https://books.google.sk/books?id=n3ry2WtdTbMC>.
- [6] JONES, Capers. Software project management practices: Failure versus success. In: *CrossTalk: The Journal of Defense Software Engineering*. 2004, vol. 17, no. 10, pp. 5-9.
- [7] KERZNER, H. Project Management: A Systems Approach to Planning, Scheduling, and Controlling. Wiley, 2009. ISBN 978-0-470-27870-3.
- [8] LUTKEVICH, Ben. Gantt chart. TechTarget, 2021. Available at: <https://www.techtarget.com/searchsoftwarequality/definition/Gantt-chart>.

Návrh simulátora fuzzy kognitívnych máp pre potreby navigácie vozidiel

Ján VAŠČÁK

Katedra Kybernetiky a Umelej Inteligencie, Fakulta Elektrotechniky a Informatiky,
Technická Univerzita v Košiciach, Slovenská Republika

jan.vascak@tuke.sk

Abstrakt – Táto práca sa zameriava na navigáciu vozidla v neznámom prostredí s prekážkami. Navigácia sa vykonáva prostredníctvom fuzzy kognitívnych máp. Článok poskytuje teoretické základy fuzzy kognitívnych máp a ich využitie na navigačné účely. Naším hlavným cieľom je vyvinúť vlastný navigačný simulátor. Po jeho dokončení program umožní užívateľovi vybrať vytvorenú fuzzy kognitívnu mapu a následne vizualizovať správanie vozidla podľa vybranej mapy. Prostredie simulátora bude slúžiť ako nástroj na realizáciu experimentov. Cieľom experimentov je nájsť funkčné súvislosti medzi zmenami parametrov mapy a navigáciou vozidla. Na vyhodnotenie, či sú fuzzy kognitívne mapy vhodné na túto úlohu, sa najprv realizovala manuálna navigácia. Potom sa porovnali výsledky.

KLúčové slová – fuzzy kognitívne mapy, navigácia, simulátor, vizualizačné nástroje

I. ÚVOD

Navigácia vozidiel je problém, ktorý stále do veľkej miery riešia ľudia. Automatické systémy, ktoré sú schopné dosiahnuť tento cieľ, sú požadované z viacerých dôvodov. Je to napríklad šetrenie ľudských zdrojov, efektívnejšia spotreba paliva alebo zlepšenie plánovania dopravy. Preto sú pre ďalší výskum v tejto oblasti potrebné nové metódy a prístupy k riešeniu problému navigácie. Na riešenie navigačných problémov existuje celý rad metód. Cieľom tohto článku je opísať a pomocou vlastného simulátora implementovať prístup využívajúci fuzzy kognitívne mapy (FKM) na navigačné účely.

II. OPIS FUZZY KOGNITÍVNYCH MÁP

A. Fuzzy kognitívne mapy

FKM je orientovaný graf, ktorého uzly predstavujú pojmy a spojenia medzi nimi predstavujú vzťahy. Teória FKM zahŕňa celú triedu rôznych modifikácií FKM, čo je aj dôvodom toho, že definícia FKM je všeobecnejšia a rozsiahlejšia [1]. Nasledujúci stručný výklad je opisom FKM použitej výlučne v tejto práci. Aktivačné hodnoty pojmov môžu byť z intervalu $[0; 1]$, kde 0 znamená, že pojem nie je platný, 1 znamená, že je platný. Hodnoty medzi nimi charakterizujú neurčitost'. Vzťahy možno nastaviť z intervalu $[-1; 1]$ a predstavujú znalosti experta v danej oblasti. Na účely realizácie experimentov bola použitá nasledujúca FKM. K dispozícii bolo niekoľko vstupných uzlov alebo pojmov. Vstupné hodnoty zo snímačov boli najprv fuzzifikované pomocou funkcií príslušnosti uzlov. Potom boli nastavené na príslušné uzly FKM ešte v požadovanom intervale. Potom sa vypočítal nový krok stavu FKM. Nakoniec sa výstupné uzly alebo pojmy FKM opäť defuzzifikovali pomocou funkcií príslušnosti príslušných uzlov. Získané výstupné hodnoty sa použili v navigácii na riadenie vozidla a určenie, či má zastaviť. Výhodou použitia FKM je jednoduchý opis vzťahov v zložitom dynamickom systéme [1]. Príkladom takéhoto systému je navigácia vozidla medzi prekážkami.

B. Príklady použitia fuzzy kognitívnych máp

FKM sa dajú použiť pri riešení mnohých problémov. Ako už bolo spomenuté, sú určené na opis zložitých dynamických systémov. Okrem navigácie sa dajú a úspešne používali aj v iných oblastiach, napríklad v medicíne. V [2] je opísaný príklad použitia FKM pri diagnostike rakoviny močového mechúra. Úspešnosť diagnostického systému bola 80–90%. Použitie FKM je vhodné aj na účely regulácie [3]. V robotike sa FKM môže použiť na stabilizáciu pohybu robota alebo na realizáciu robotického futbalu. Výhodou FKM je, že dokážu pracovať s neurčitými informáciami (napr. zo senzorov).

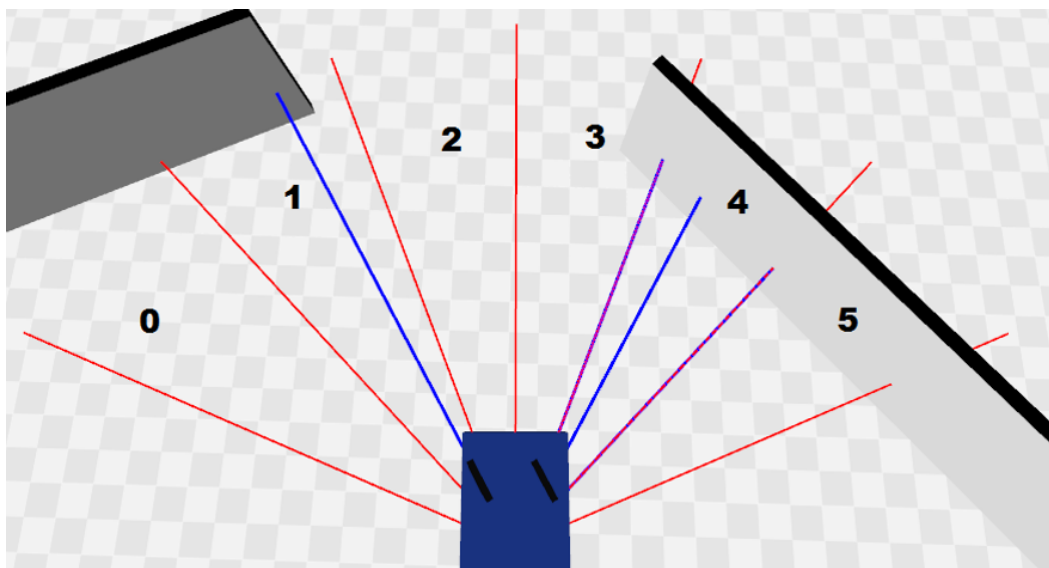
V [4] je opísaná navigácia využívajúca FKM. Cieľom bolo pomocou robota nájsť pohár v neznámom prostredí s prekážkami. Parametre pohára boli vopred špecifikované. Využíval sa aj systém hodnotenia situácie (Situation Assessment Framework - SAF). Jeho úlohou bolo generovať čiastkové ciele, ktoré sa vyhodnocovali pomocou FKM. Potom sa robot rozhodol, čo má urobiť. Úloha podobná v tejto práci je opísaná v [5]. Na simuláciu navigácie v 2D prostredí bola použitá metóda umelého potenciálového poľa (APF). Získané výsledky sa použili na testovanie navigácie založenej na FKM. Najprv sa vytvorila štruktúra FKM definovaním uzlov a spojení. Potom sa iteratívne menili parametre konkrétnej FKM. Na tento účel sa použili genetické algoritmy (GA). Po určitom počte iterácií sa zistilo, že FKM kopíruje správanie metódy APF až na 90%.

III. IMPLEMENTÁCIA SIMULÁTORA

V tejto práci bol použitý dostupný FKM editor *Irrlicht Engine* na vytváranie a úpravu FKM používaných v navigácii [6]. Program, ktorý bol vytvorený, bol 3D vizualizáciou s implementovanou fyzikou, ktorá simuluje reálny svet. Umožňuje používateľovi vytvoriť prostredie s prekážkami a nastaviť polohu cieľa a vozidla, manuálne navigovať vozidlo do cieľa a potom spustiť FKM navigáciu a porovnať trajektórie oboch prístupov.

Všetky prekážky sú reprezentované ako steny s konštantnou výškou 3 metre a šírkou 0,2 metra (v metrickej systéme fyzikálnej simulácie). Žiadne iné prekážky sa nepredpokladajú. Vozidlo je dlhé 3 metre a široké 1,5 metra. Maximálny polomer riadenia je $\pm 30^\circ$. Vozidlo je typu s pohonom zadných kolies.

Dôležitou otázkou je aj implementácia snímačov. Každý snímač predstavuje uhlový sektor 0,4 radiánu a dosah 8 metrov. Je ich 6 a smerujú dopredu, vid' Obr. 1, kde červené čiary predstavujú hranice snímačov a modré čiary sú vizualizáciou najbližšej vzdialenosti k prekážke vzhľadom na obrisy vozidla. Snímače by mali poskytovať najbližšiu vzdialenosť k akejkol'vek prekážke v predpísanom rozsahu a v určenom uhlovom sektore. Táto implementácia sa vykonáva pomocou tzv. lúčov. Každý senzor má 4 lúče. Nevýhodou môže byť, že niekedy sa prekážka môže nachádzať v tzv. slepom uhle snímača a preto si ju nemusí všimnúť.



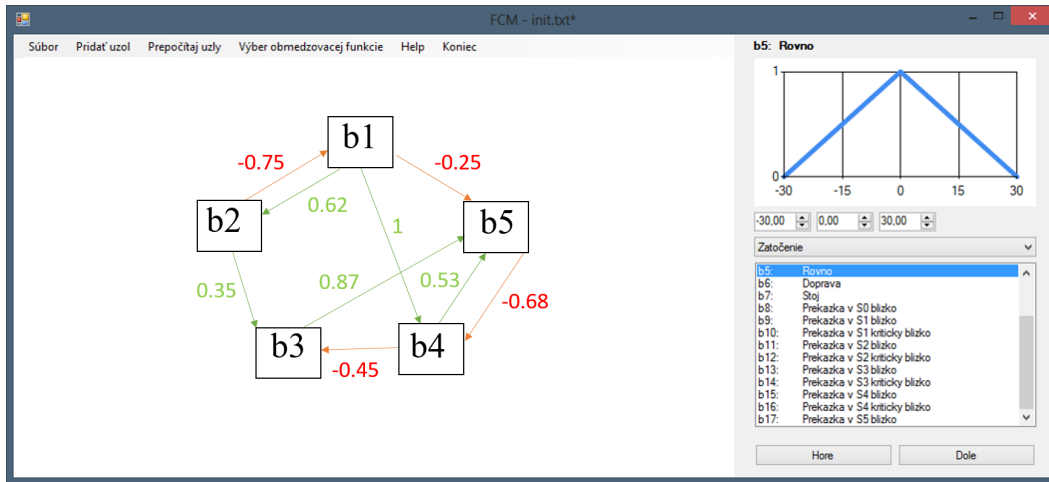
Obr. 1 Ukážka vlastného simulátora – vizualizačné okno.

Súčasťou navrhnutého simulátora je aj návrhové okno, vid' Obr. 2, kde sú zobrazené základné nástroje na konštruovanie jednak štruktúry FKM a jednak definovanie funkcií príslušnosti pre jej uzly.

IV. EXPERIMENTY A ICH VYHODNOTENIE

Na testovanie navigačného správania bola použitá manuálne navrhnutá a nastavená FKM, vid' Obr. 3 zobrazujúci celkovú štruktúru FKM a Obr. 4 zobrazujúci definície funkcií príslušnosti jednotlivých uzlov mapy. FKM bola potom testovaná na rôznych scénach s rôznymi prekážkami a rôznymi východiskovými polohami vozidla, vid' Tabuľka 1, ktorej výsledky potvrdzujú univerzálnosť použitia navrhutej FKM pri navigácii vozidla.

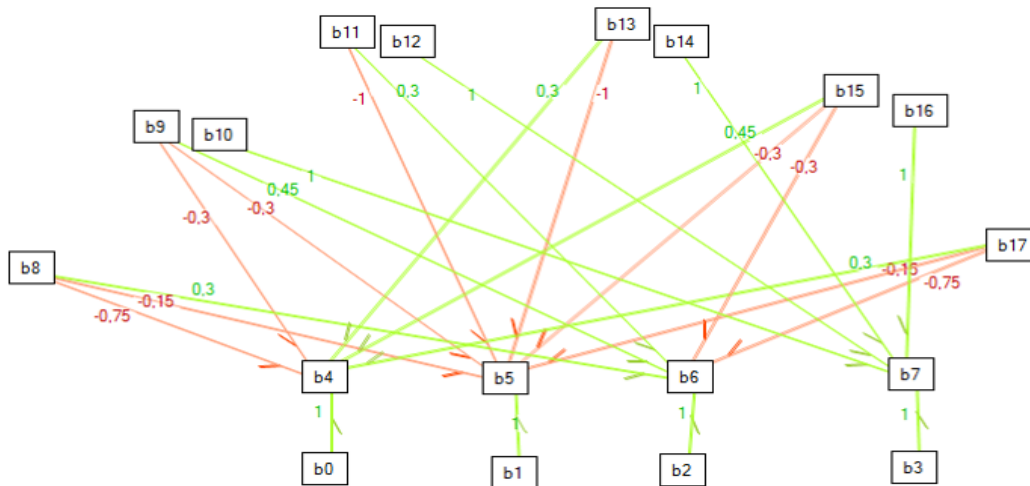
Uzly mapy predstavujú konkrétne pojmy. Vstupné uzly sú v Obr. 4 označené ako $b0 - b3$ a $b8 - b17$. Hodnoty týchto uzlov sa nastavujú pomocou fyzikálnej simulácie. Následne sa vypočíta



Obr. 2 Ukážka vlastného simulátora – návrhové okno.

ďalší iteračný krok FKM, a tak sa nastaví výstupné uzly riadenia $b4 - b6$ a uzol $b7$ testujúci podmienku, či má vozidlo zastaviť. Ich výstupy predstavujú akčné zásahy riadenia vozidla počas simulácie.

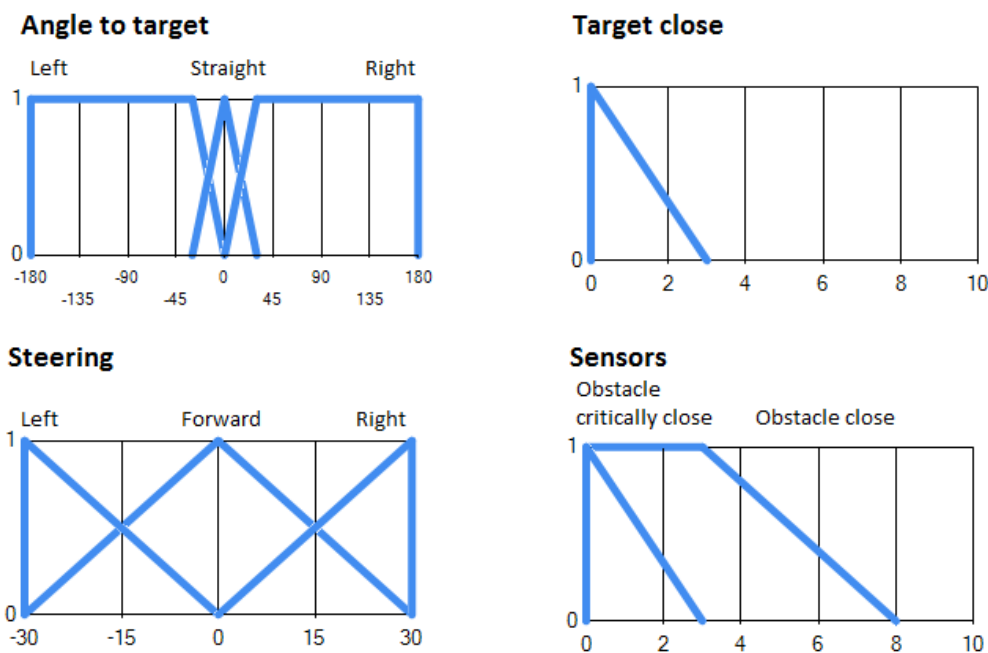
Cieľom návrhu je získať rovnaké alebo podobné výsledky správania FKM ako pri navigácii riadenej človekom. Preto sa pred navigáciou FKM realizuje manuálna navigácia. Celkovo sa vykonalo 9 experimentov, každý s vlastnou scénou s prekážkami. Na každej scéne sa vykonali 3 testy s rôznymi východiskovými pozíciami, viď Obr. 5.



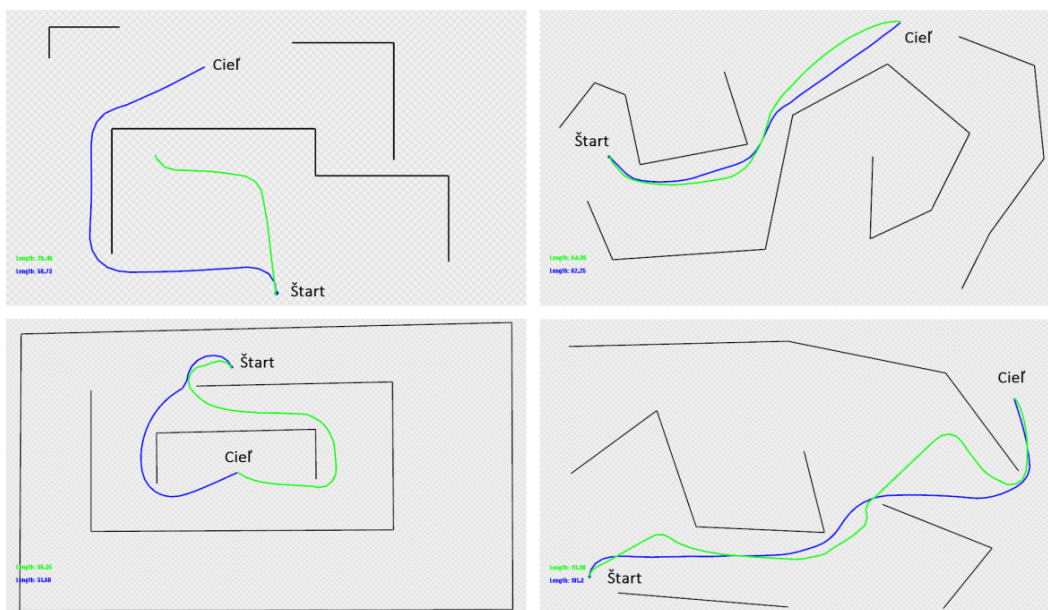
Obr. 3 Finálny návrh FKM určenej pre navigáciu vozidla.

Porovnanie výsledkov jednotlivých testov sú uvedené v Tabuľke 1. Posledné dva stĺpce predstavujú rozdielové porovnanie manuálnej a FKM navigácie v absolútnom a relatívnom (percentuálnom) pomere voči výsledkom manuálneho riadenia považovaného za etalón. Symbol „-“ znamená, že navigácia pomocou FKM nebola schopná nájsť cieľ. Je to výsledok implementácie navigácie, pri ktorej vozidlo nemôže cúvať. Preto sa môže dostať do tzv. robotického pasce, keď všetky snímače hlásia prekážku pred sebou a vozidlo sa nedokáže rozhodnúť, kam má ísť (okrem prípadu cúvania).

Vo väčšine prípadov bola navigácia FKM porovnateľná voči manuálnej navigácii (niekde o niečo horšia a niekde o niečo lepšia). Výnimkou boli prípady, keď sa cesta nenašla, čomu sa dalo zabrániť realizáciou spätného pohybu vozidla, avšak to si vyžaduje komplexnejší návrh FKM s mechanizmom pamätania si už prejdenej cesty, aby sa vylúčila cesta, ktorá viedla do pasce. Taktiež sa vyskytlo aj niekoľko veľmi dlhých ciest pomocou FKM navigácie (napr. experiment č. 5, test č. 1), ktoré sa objavili, keď vozidlo zvolilo nesprávne odbočenie na križovatke, čo bolo spôsobené príliš krátkym dosahom snímačov vozidla, a tým sa im dá tiež predísť.



Obr. 4 Definície funkcií príslušnosti pre jednotlivé uzly FKM z Obr. 3.



Obr. 5 Príklady výstupov experimentov. Zelená farba predstavuje navigáciu FKM, modrá farba manuálnu navigáciu.

V. ZÁVER

Zámerom tejto práce bolo predstaviť interaktívne založený simulátor FKM pre navigačné účely. V rámci toho bol testovaný nielen návrh samotnej FKM, ale aj overovaná možnosť využitia takéhoto prostriedku v danom type úloh. Z tohto dôvodu boli výstupy riadenia pomocou FKM porovnávané voči manuálnemu riadeniu (taktiež v tomto simulátore) ako etalónu, ktorý môžeme považovať za najlepší spôsob riadenia. Výsledky ukázali veľmi podobné správanie sa, aj keď ručné riadenie bolo väčšinou mierne lepšie, ale jednalo sa skoro vždy o zanedbateľné rozdiely. Preto je možné záverom konštatovať, že FKM je vhodným prostriedkom na návrh automatickej navigácie. Dôležitým faktorom je aj spôsob nastavenia parametrov FKM, ktoré sa dajú zlepšiť implementáciou techník automatickej adaptácie FKM a ktoré budú predmetom ďalšieho vývoja simulátora.

Tabuľka 1
Výsledky experimentov porovnávajúce manuálne riadenie vozidla a riadenie pomocou FKM.

Číslo experimentu	Číslo testu	Dĺžka trajektórie		Absolútny rozdiel	Percentuálny rozdiel
		Manuálne	FKM		
1	1	44.6	48.4	3.8	8.52
	2	67.6	72.9	5.3	7.84
	3	28.1	27.3	-0.8	-2.85
2	1	58.7	–	–	–
	2	46.6	–	–	–
	3	31.2	31.0	-0.2	-0.64
3	1	88.5	–	–	–
	2	62.3	65.0	2.7	4.33
	3	57.1	56.6	-0.5	-0.88
4	1	75.5	77.8	2.3	3.05
	2	66.6	71.8	5.2	7.81
	3	96.4	100.5	4.1	4.25
5	1	101.1	165.5	64.4	63.7
	2	68.8	70.4	1.6	2.33
	3	51.7	66.4	14.7	28.43
6	1	101.3	112.0	10.7	10.56
	2	76.8	77.8	1.0	1.3
	3	109.9	110.0	0.1	0.09
7	1	73.4	97.4	24.0	32.7
	2	74.7	74.6	-0.1	-0.13
	3	115.6	114.6	-1.0	-0.86
8	1	46.4	46.2	-0.2	-0.43
	2	34.8	45.6	10.8	31.03
	3	36.1	37.9	1.8	4.99
9	1	39.6	–	–	–
	2	29.3	27.3	-2.0	-6.83
	3	28.7	27.5	-1.2	-4.18

POĎAKOVANIE

Táto práca je finančne podporená Agentúrou na podporu výskumu a vývoja prostredníctvom projektu APVV-20-0247, ENISaC — Edge-enabled inteligentné snímanie a výpočty.

LITERATÚRA

- [1] E. Jiya, O. Georgina, and A. O., "A review of fuzzy cognitive maps extensions and learning," *Journal of Information Systems and Informatics*, vol. 5, no. 1, pp. 300–323, Mar. 2023.
- [2] P. P. Groumpos, "Fuzzy cognitive maps: basic theories and their applications in medical problems," in *19th Mediterranean Conference on Control & Automation (MED)*, 2011, pp. 1490–1497.
- [3] C. Stylios and P. Groumpos, "Mathematical formulation of fuzzy cognitive maps," in *Proceedings of the 7th Mediterranean Conference on Control and Automation (MED99)*, 06 1999, pp. 1–12.
- [4] N. Ganganath, M. Walker, and H. Leung, "Fuzzy cognitive map based situation assessment framework for navigation goal detection," in *2013 IEEE International Conference on Systems, Man, and Cybernetics*, 2013, pp. 1444–1449.
- [5] O. Motlagh, S. H. Tang, A. R. Ramli, N. Ismail, and D. N. Nia, "A novel approach to motion modeling using fuzzy cognitive map and artificial potential fields," in *11th International Conference on Control Automation Robotics & Vision*, 2010, pp. 1963–1967.
- [6] H. C. Park and N. Baek, "Design of selfengine: A lightweight game engine," in *Information Science and Applications*, K. J. Kim and H.-Y. Kim, Eds. Singapore: Springer Singapore, 2020, pp. 223–227.

SAT kódovanie kardinalitných ohraničení pomocou binárneho stromu

¹Lubomír RUSNÁK, ²Marián MACH

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

¹lubomir.rusnak@student.tuke.sk, ²marian.mach@tuke.sk

Abstrakt — Pri riešení úloh prostriedkami výrokovej logiky je hľadanie riešenia ekvivalentné vytvoreniu logického modelu a nájdeniu podmienok splniteľnosti tohto vytvoreného modelu. Článok sa zameriava na problematiku vytvárania logických modelov pri riešení úloh používajúcich kardinalitné ohraničenia, obmedzujúce počet vybraných alternatív z dostupnej množiny alternatív. Dôraz je na modelovaní takýchto ohraničení v tvare CNF vytváranom nad logickými symbolmi pomocou negácie, konjunkcie a disjunkcie, pričom sa hlbšie venuje ich kódovaniu metódou využívajúcou binárny strom. V rámci experimentovania bolo realizovaných niekoľko experimentov s týmto kódovaním a jeho porovnaním voči binomickému kódovaniu ako aj P-H kódovaniu v rámci zvoleného testovacieho problému.

Kľúčové slová — binárny strom, kardinalitné ohraničenia, logický model, totalizer, výroková logika

I. ÚVOD

Pri použití výrokovej logiky pre riešenie úloh je možné zostaviť logický model úlohy, s jeho následným testovaním na splniteľnosť [1]. V prípade splniteľnosti hodnoty logických symbolov, ktoré vedú na splniteľný model, kódujú hľadané riešenie. Častým prípadom pri vytváraní logického modelu je reprezentácia rozličných ohraničení, ktoré musí model spĺňať. Jedným typom takýchto ohraničení sú kardinalitné ohraničenia. Tieto ohraničenia modelujú situáciu, keď existuje viac alternatív, z ktorých možno validne vybrať iba určitý počet – výber iného počtu je považovaný za nevalidný. Najčastejší tvar kardinalitných ohraničení je voľba maximálne (minimálne, presne) b alternatív z celkového počtu n alternatív ($0 < b < n$).

Problémom je, že SAT solver (softvér používaný na dokazovanie splniteľnosti logických modelov) takto definovaným ohraničeniam nerozumie. Preto je potrebné takéto vysokoúrovňové vyjadrenie kardinalitných ohraničení zakódovať do tvaru logického výrazu reprezentovaného vo forme CNF (konjunktívnej normálnej formy). Pre toto kódovanie kardinalitných ohraničení existuje viacero metód [2].

Tento článok sa zameriava na metódu kódovania kardinalitných ohraničení využívajúcu štruktúru binárneho stromu [3]. Štruktúra článku je nasledovná – v druhej kapitole sa venuje danej metóde kódovania a v tretej kapitole je uvedená séria experimentov s danou metódou na úlohe hľadania riešenia nesplniteľného problému.

II. METÓDA KÓDOVANIA S BINÁRNÝM STROMOM

A. Reprezentácia celého čísla

Táto metóda je založená na simulovaní aritmetickej operácie súčtu nad doménou nezáporných čísel (reprezentujúcich počty splnených alternatív) pomocou logických symbolov. Predpokladajme, že je potrebné reprezentovať celé číslo z intervalu $[0, n]$. Pre jeho reprezentáciu je potrebných n logických symbolov. Pomocou týchto symbolov jednotlivé čísla z daného intervalu budú vyjadrené

$$\begin{array}{l} 0 \quad \sim x_1 \sim x_2 \sim x_3 \dots \sim x_n \\ 1 \quad x_1 \sim x_2 \sim x_3 \dots \sim x_n \\ 2 \quad x_1 x_2 \sim x_3 \dots \sim x_n \\ \dots \\ n \quad x_1 x_2 x_3 \dots x_n \end{array} \quad (1)$$

kde symbol \sim reprezentuje negáciu. Teda reprezentované číslo je vyjadrené počtom splnených symbolov, pričom tie splnené symboly sa vždy „nasúvajú“ zľava. Táto podmienka je vyjadrená pomocou

$$x_i \vee \sim x_{i+1} \quad (2)$$

pre $i=1, \dots, n-1$.

Výhodou takejto reprezentácie je, že možno jednoducho vyjadriť nielen konkrétne číslo, ale aj podinterval. Napríklad podmienka, že číslo má minimálne hodnotu 2 je vyjadrená ako

$$x_1 x_2 ?? \dots ? \quad (3)$$

pomocou presného stanovenia hodnôt prvých dvoch symbolov, zatiaľ čo ostatné symboly vpravo ostávajú neurčené. Podobne podmienka, že číslo má hodnotu maximálne 2 bude

$$?? \sim x_3 \dots \sim x_n \quad (4)$$

pomocou stanovenia hodnoty symbolov od tretej až po poslednú pozíciu, zatiaľ čo prvé dva symboly ostávajú neurčené.

B. Súčet dvoch čísel

Na ilustráciu uvažujme súčet dvoch sčítancov – jeden je celým číslom z $[0,3]$ a druhý zase celým číslom z $[0,2]$. Prvý bude reprezentovaný tromi symbolmi, zatiaľ čo pre reprezentáciu druhého je potrebné použiť dva symboly. Výsledný súčet bude definovaný piatimi symbolmi

$$a_1 a_2 a_3 + b_1 b_2 = v_1 v_2 v_3 v_4 v_5 \quad (5)$$

Spôsob určovania pravdivostnej hodnoty symbolov v_1 až v_5 je daný hodnotami symbolov sčítancov a závislosťnými pravidlami. Ilustrujme to na symbole v_4 .

Ak v_4 je pravdivý, znamená to, že hodnota výsledku sčítania je číslo minimálne 4 (teda 4 alebo 5). Aby toto platilo pre výsledok, tak pre sčítance musí platiť

$$\begin{array}{ccc} a_1 a_2 a_3 + b_1 b_2 = v_1 v_2 v_3 v_4 v_5 & & \\ \min 0 & \min 4 & \min 4 \\ \min 1 & \min 3 & \min 4 \\ \min 2 & \min 2 & \min 4 \\ \min 3 & \min 1 & \min 4 \\ \min 4 & \min 0 & \min 4 \end{array} \quad (6)$$

Keďže prvý sčítanec nemôže mať hodnotu 4 a druhý nemôže mať hodnotu 3 ani 4, tak do úvahy prichádzajú iba dva riadky z predchádzajúcej tabuľky. Tieto dva riadky sa vyjadria dvomi podmienkami

$$a_2 \wedge b_2 \rightarrow v_4 \quad a_3 \wedge b_1 \rightarrow v_4 \quad (7)$$

Podobne možno uvažovať pre prípad, že v_4 je nepravdivý, znamená to, že hodnota výsledku sčítania je číslo maximálne 3 (teda 1, 2 alebo 3). Aby toto platilo pre výsledok, tak pre sčítance musí platiť

$$\begin{array}{ccc} a_1 a_2 a_3 + b_1 b_2 = v_1 v_2 v_3 v_4 v_5 & & \\ \max 0 & \max 3 & \max 3 \\ \max 1 & \max 2 & \max 3 \\ \max 2 & \max 1 & \max 3 \\ \max 3 & \max 0 & \max 3 \end{array} \quad (8)$$

Keďže druhý sčítanec nemôže mať hodnotu 3, do úvahy neprichádza prvá alternatíva predchádzajúcej tabuľky. Navyše hodnota prvého je vždy maximálne 3 a hodnota druhého sčítanca vždy maximálne 2, a teda toto nie je potrebné explicitne vyjadriť. Výsledkom sú tri podmienky

$$\sim a_2 \rightarrow \sim v_4 \quad \sim a_3 \wedge \sim b_2 \rightarrow \sim v_4 \quad \sim b_1 \rightarrow \sim v_4 \quad (9)$$

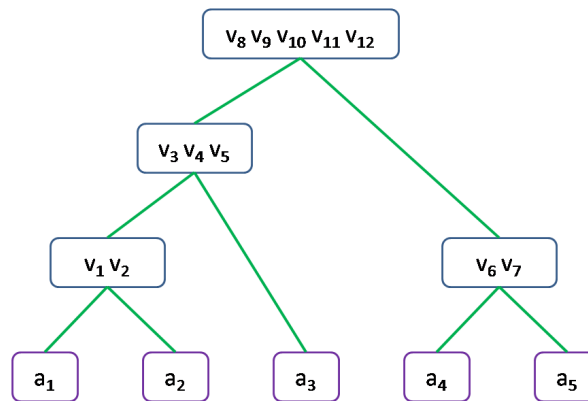
Podobným spôsobom sa vygenerujú podmienky pre každý symbol výsledku sčítania.

C. Súčet viacerých čísel

Pri sčítaní viac než dvoch čísel sa použije opakované sčítavanie po dvojiciach čísel s vytváraním medzisúčtov a skladaním týchto medzisúčtov v ďalších sčítaniach. Vzniká tak binárny strom, ktorého listy reprezentujú pôvodné sčítance, medziľahlé uzly sú medzisúčtami a koreň stromu reprezentuje celkový výsledok súčtu (takáto štruktúra sa niekedy označuje ako totalizer).

Keďže je mnoho spôsobov ako vytvárať takýto binárny strom, preferuje sa taký spôsob, ktorý vyžaduje čo najmenší počet pomocných premenných pre reprezentáciu medzivýsledkov. Pre

toto sa používa usporiadanie v tvare vybalancovaného binárneho stromu, kde hĺbky ľavého a pravého podstromu sa môžu líšiť maximálne o 1.



Obr. 1 Binárny strom súčtu viacerých čísel

Ukážka takéhoto stromu pre zistenie, koľko z piatich alternatív je splnených, je na obr. 1. Splnenie jednotlivých alternatív je reprezentované pravdivosťnými hodnotami symbolov a_i – jeden symbol pre každú z alternatív. Čím je uzol v strome na vyššej úrovni, tým viac alternatív pokrýva, tým väčšia môže byť jeho hodnota a preto tým väčší počet symbolov pre reprezentáciu hodnoty daného uzla. Najväčším počtom symbolov je vždy reprezentovaná hodnota koreňového uzla.

D. Komparátor

Netreba zabúdať na to, že cieľom celého uvedeného postupu je zabezpečiť, aby počet splnených alternatív zodpovedal kardinalitnému ohraničeniu. Typicky takéto ohraničenie máva jeden z tvarov

$$\geq_b (X_1, X_2, \dots, X_n) \quad \leq_b (X_1, X_2, \dots, X_n) \quad (10)$$

obmedzujúcich počet splnených alternatív limitom b zdola (ľavý tvar) alebo zhora (pravý tvar). Spôsob reprezentácie výsledného súčtu umožňuje jednoduchým spôsobom zabezpečiť platnosť takýchto ohraničení. Tak napríklad, ak by sme uvažili reprezentáciu podľa obr. 1, tak podmienka

$$v_{10} \quad (11)$$

zabezpečuje platnosť aspoň troch alternatív z piatich možných, zatiaľ čo podmienka

$$\sim v_{11} \quad (12)$$

zabezpečuje platnosť najviac troch alternatív z piatich možných.

Výhodou tohto spôsobu kódovania je aj to, že sa jednoducho dá zabezpečiť platnosť presne určitého počtu alternatív alebo dokonca použiť dva rôzne prahy a zabezpečiť tak počet platných alternatív z požadovaného intervalu.

III. VYBRANÉ EXPERIMENTY

A. Testovací problém

Pre experimentovanie s rôznymi metódami kódovania kardinalitných ohraničení bol zvolený problém ukladania predmetov do krabíc [4]. Pri tomto probléme sú zadané dve množiny objektov – p predmetov nejakého typu a h krabíc. Krabice majú rovnakú kapacitu c – do každej krabice je možné vložiť maximálne c predmetov.

Úlohou je uložiť všetky predmety do krabíc tak, aby boli dodržané nasledovné dve podmienky:

- Každý predmet má byť vložený do práve jednej krabice – žiadny predmet nemôže ostať mimo a taktiež žiadny predmet nemôže byť vložený súčasne do viacerých krabíc.
- Je potrebné rešpektovať maximálnu kapacitu krabíc – do každej krabice môže byť vložených maximálne c predmetov.

Je zrejme, že v prípade $p \leq c \cdot h$ úloha má riešenie (vlastne väčší počet riešení vzhľadom na symetriu medzi krabicami a symetriu medzi predmetmi), zatiaľ čo v prípade $p > c \cdot h$ žiadne riešenie neexistuje.

Aby bolo možné porovnať nielen výsledné logické modely generované rôznymi metódami kódovania ale aj ich vplyv na riešenie SAT solverom, bol zvolený nesplniteľný variant úlohy s neexistujúcim riešením. V tomto prípade totiž solver musí najprv prehľadať celý priestor možných riešení a až po jeho úplnom prehľadaní je schopný konštatovať nesplniteľnosť logického modelu. Pre tento variant je možné upraviť prvú podmienku na jednoduchší tvar – každý predmet musí byť vložený do minimálne jednej krabice. Táto úprava zjednoduší model a zároveň nezmení nesplniteľný model na splniteľný.

Model riešeného príkladu je teda tvorený dvomi kardinalitnými ohraničeniami. Ak symbol $X_{i,j}$ bude reprezentovať vloženie i -teho predmetu do j -tej krabice (celkovo je potrebných $p \cdot h$ takýchto symbolov), tak tieto ohraničenia budú mať tieto tvary

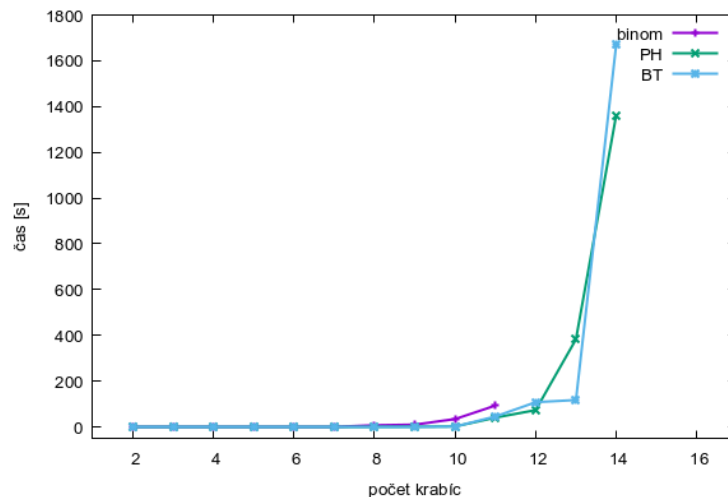
$$\forall_{i=1}^p \geq_1 (X_{i,1}, X_{i,2}, \dots, X_{i,h}) \quad (13)$$

$$\forall_{j=1}^h \leq_c (X_{1,j}, X_{2,j}, \dots, X_{p,j}) \quad (14)$$

B. Časová zložitosť generovaných modelov

Vo vykonaných experimentoch (overujúcich výsledky dosiahnuté v práci [5] so snahou ich reprodukcie) bola použitá metóda kódovania spomenutá v predchádzajúcej kapitole. Pre porovnanie dosiahnutých výsledkov boli použité ďalšie dve metódy kódovania kardinalitných ohraničení – binomické kódovanie a P-H (Pigeon-Hole) kódovanie [6]. Boli generované výsledné logické modely a tie vložené do SAT solvera, ktorý vo všetkých prípadoch dokázal nesplniteľnosť týchto modelov, čím overil ich správnosť.

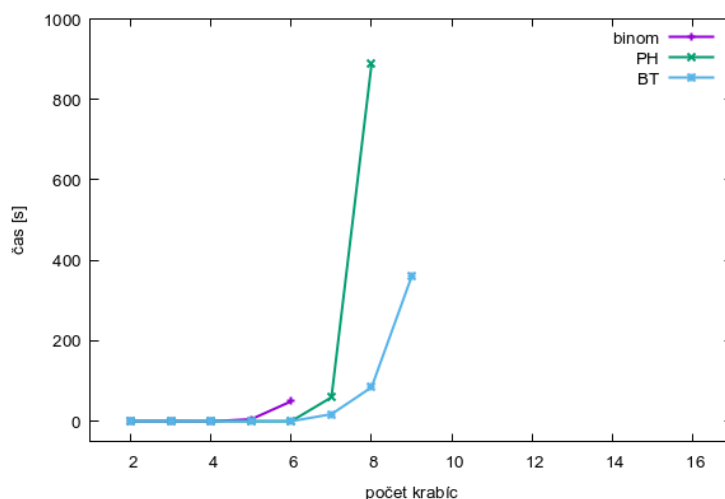
Experiment bol vykonávaný pre rôzne počty krabíc – od najmenej hodnoty 2 s postupným zvyšovaním až dovedy, kým solver bol neschopný potvrdiť nesplniteľnosť modelu v rámci doby 30 minút. Každá krabica mala kapacitu iba jedného predmetu $c=1$, počet predmetov bol vždy o jeden predmet väčší než bola absorpčná schopnosť krabíc. Výsledky sú znázornené na obr. 1.



Obr. 1 Čas potrebný pre dokázanie nespĺniteľnosti, $c=1$

Binomické kódovanie dokázalo vo vymedzenom čase zvládnuť model s 11 krabicami, časový nárast pre spracovanie modelu s 12 krabicami už presiahol vymedzený čas. Ostatné dve metódy dosiahli vo vymedzenom čase rovnaký počet krabíc s tým, že pri jedenástich krabiciach bola lepšou metóda prezentovaná v článku, zatiaľ čo v prípade štrnástich krabíc si poradie vymenili. Je zrejme, že tento prípad nefavorizuje ale ani nezavrhuje prezentovanú metódu.

O niečo iná situácia nastáva v prípade zväčšenej kapacity krabíc. Prípad $c=2$ je znázornený na obr. 2.



Obr. 2 Čas potrebný pre dokázanie nespľniteľnosti, c=2

Pri zväčšenej kapacite všetky metódy zvládli vo vymedzenom čase menší počet krabíc (avšak vďaka kapacite väčší počet predmetov). Tu už vidno prednosť prezentovanej metódy, schopnej zvládnuť najväčší počet krabíc a pri počtoch, zvládnuteľných aj metódou P-H, dosahujúcej kratších časov. Podobné pomery nastali aj pre prípad c=3.

C. Veľkosť generovaných modelov

Jednotlivé kódovacie metódy používajú jednak rôzne počty pomocných symbolov a jednak sú založené na rozdielnych princípoch. Následkom toho sú schopné generovať logické modely rôznych veľkostí. Jednou z mier pre popis veľkosti modelov je počet literálov (pozitívnych alebo negovaných symbolov) použitých v zostave modelu. Výsledky sú v tab. 1, kde je uvedený prípad pre c=2 a v tab. 2, kde je uvedený prípad pre 7 krabíc.

Tabuľka 1
 Celkový počet použitých pomocných symbolov a literálov, c=2.

Počet krabíc	BT		PH		Binom	
	Pomocné symboly	Literály	Pomocné symboly	Literály	Pomocné symboly	Literály
2	34	297	28	104	0	70
3	95	906	66	258	0	326
4	188	1977	120	480	0	1044
5	327	3674	190	770	0	2530
6	502	6065	276	1128	0	5226
7	713	9258	378	1554	0	9660
8	968	13393	496	2048	0	16456
9	1289	18666	630	2610	0	26334

Tabuľka 2
 Celkový počet použitých pomocných symbolov a literálov, 7 krabíc.

C	BT		PH		Binom	
	Pomocné symboly	Literály	Pomocné symboly	Literály	Pomocné symboly	Literály
1	328	3553	154	546	0	448
2	713	9258	378	1554	0	9660
3	1140	17189	686	3192	0	167727

Je zrejmé, že prezentovaná metóda používa najväčší počet pomocných symbolov, ktorý je približne dvojnásobný oproti metóde P-H. Rast počtu použitých literálov je ešte viac v neprospech prezentovanej metóde, iba v prípade vyšších počtov krabíc začne väčšie modely

produkovať binomické kódovanie (so zvyšujúcou sa kapacitou tento rozdiel exponenciálne narastá). Je potrebné si však uvedomiť, že pri prezentovanej metóde komparátor produkuje klauzuly jednotkovej dĺžky – a teda použitím jednoduchej jednotkovej propagácie je možné model zmenšiť. Je to možné robiť pomocou predpracovania modelu ešte pred jeho odovzdaním solveru alebo sa možno spoľahnúť na solver, keďže takýto spôsob predspracovania je v súčasnosti bežnou výbavou solverov.

IV. ZÁVER

Vykonané experimenty potvrdili použiteľnosť prezentovanej metódy kódovania kardinalitných ohraničení. Aj keď v prípade ohraničení typu \leq_1 daná metóda neprevyšuje ostatné štandardne používané metódy (v našich experimentoch reprezentované kódovaním P-H), pri ohraničeniach typu \leq_c , $c > 1$ umožňuje o niečo posunúť hranicu praktického použitia modelov, keďže dokáže spracovať aj o niečo väčšie problémy než metóda P-H. Na druhej strane je potrebné upozorniť na väčšiu veľkosť vytváraných modelov kvôli generovaniu veľkého množstva podmienok pre simuláciu aritmetických operácií pomocou logických symbolov. Vďaka jednotkovej propagácii je síce možné tieto modely zmenšiť, avšak zatiaľ kvantitatívny odhad tohto zmenšenia nie je známy.

LITERATÚRA

- [1] V. Kvasnička, *Matematická logika*, Bratislava, Slovenská Technická Univerzita, 2006
- [2] E. Wynn, “A Comparison of encodings for cardinality constraints in a SAT solver”, ArXiv, 2018.
- [3] O. Bailleux a Y. Boufkhad, “Efficient CNF Encoding of Boolean Cardinality Constraints”, *Principles and Practice of Constraint Programming, LNCS*, 2003, 108-122.
- [4] L. Rusnák a M. Mach, „SAT kódovanie kardinalitných ohraničení pomocou P-H metódy“, *Electrical Engineering and Informatics* 13, 2022, Košice, 491-495.
- [5] L. Rusnák, *Transformácia úloh s ohraničeniami na SAT problémy*, Diplomová práca, Fakulta elektrotechniky a informatiky Technická univerzita v Košiciach, 2021.
- [6] S. Jabbour, L. Sais a Y. Salhi, “A Pigeon-Hole Based Encoding of Cardinality Constraints”, *Int. Symposium on Artificial Intelligence and Mathematics*, Fort Lauderdale, FL, 2014.

Návrh výroby paliva pre palivový článok

Filip JURÍNI, Dušan MEDVEĎ

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

filip.jurini@student.tuke.sk , dusan.medved@tuke.sk

Abstrakt — Cieľom tohto príspevku je prezentovať výsledky vytvorenia jednoduchého návrhu elektrolyzéra, ktorý bude schopný vyrobiť využitím elektrolyzy palivo (v plynnom skupenstve) pre palivový článok. Nami vytvorený plyn – palivo, je alternatíva k vodíku, hoci s nižšou výhrevnosťou. Návrh zariadenia pre výrobu paliva pre palivový článok pozostáva z vytvorenia zariadenia na elektrolyzu vody. Meraním bolo zistené množstvo získaného plynu z daného objemu vody a reakčnej látky. Analýzou bola vypočítaná energetická hodnota vytvoreného plynu. Ďalej v príspevku je uvedený návrh uskladnenia vytvoreného plynu, jeho zmiešanie so zemným plynom a privádzanie do plynového kondenzačného kotla VITODES 100-W.

Príučové slová — palivový článok, vodík, elektrolyza vody, energetická hodnota plynu

I. ÚVOD

V dvadsiatom prvom storočí spôsobuje nedostatok energie a znečistenie životného prostredia veľký ekonomický a environmentálny problém. Preto sa hľadajú rôzne možnosti ako získavať elektrickú energiu z obnoviteľných (čistých) energetických zdrojov, ktoré nezaťažujú životné prostredie sploďinami alebo odpadom pri výrobe elektrickej energie. Palivové články predstavujú čistú technológiu výroby elektrickej energie s pomerne vysokou účinnosťou.

Vodík je plyn, ktorý je ako palivo najvýbúšnejší, najhorľavejší, vo vesmíre je ako prvok najjednoduchší a zároveň najrozšírenejší. Jeho spaľovaním vzniká čistá vodná para, preto je z porovnateľných palív najekologickejší. Princíp palivových článkov je o chemických reakciách, ktorý bol popísaný v roku 1838.

Vodík je možnosť, ako získavať čistú energiu a znížiť závislosť od fosílnych palív, preto sa skúmajú rôzne spôsoby jeho výroby a použitia najmä v odvetví dopravy, vykurovania domov a akumulácie elektrickej energie z obnoviteľných zdrojov a zníženia vzniku skleníkových plynov.

Existuje niekoľko typov palivových článkov, ktoré aj cez rovnaký princíp fungovania, majú rôzne možnosti využitia. Palivové články nie sú verejnosti veľmi známe, preto tento príspevok má za cieľ poukázať na ich široké možnosti využitia.

Princíp palivových článkov jednoduchý na realizáciu a použitie, pričom je však problém, že treba vyrobiť palivo (vodík) a potom tento plyn spracovať na použiteľnú energiu ako napríklad elektrinu alebo teplo bez nebezpečenstva výbuchu pri použití v praxi a nie iba v laboratórnych podmienkach.

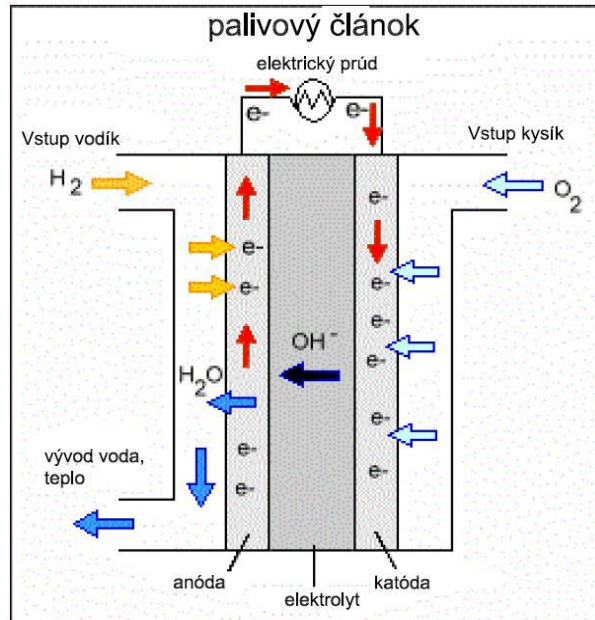
Týmto príspevkom chceme poukázať na jednoduchý princíp palivových článkov. Príspevok opisuje vlastnosti vodíka a rôzne spôsoby jeho využitia; ďalej prehľad o výrobe vodíka, technické a ekonomické aspekty jednotlivých spôsobov výroby. Zameriava sa tiež na vytvorenie jednoduchého palivového článku, pričom palivo bude vyrábané elektrolyzou vody. V príspevku je aj poukázané na vyčíslenie energetickej účinnosti elektrolyzy a možnosti využitia získaného plynu v praxi.

II. NÁVRH JEDNODUCHÉHO PALIVOVÉHO ČLÁNKU

A. Princíp palivového článku

Palivové články existujú v mnohých variantoch; všetky však fungujú rovnakým všeobecným spôsobom. Skladajú sa z troch susediacich segmentov: anóda, elektrolyt a katóda. Na rozhraniach troch rôznych segmentov prebiehajú dve chemické reakcie. Čistým výsledkom týchto dvoch reakcií je, že sa spotrebuje palivo, vytvorí sa voda alebo oxid uhličitý a generuje sa jednosmerný elektrický prúd, ktorý možno použiť na napájanie elektrických zariadení, bežne označovaných

ako záťaž. Na anóde (katalyzátor) oxiduje palivo, zvyčajne vodík, čím sa palivo mení na kladne nabitý ión a záporne nabitý elektrón. Elektrolyt je látka špeciálne navrhnutá tak, aby cez ňu prechádzali ióny. Uvoľnené elektróny prechádzajú drôtom a vytvárajú elektrický prúd. Ióny prechádzajú cez elektrolyt ku katóde. Po dosiahnutí katódy sa ióny znova spoja s elektrónmi a tieto dva reagujú s treťou chemikáliou, zvyčajne kyslíkom, za vzniku vody alebo oxidu uhličitého.



Obr. 1 Palivový článok

B. Návrh palivového článku a potrebný materiál na jeho konštrukciu

Pre zhotovenie palivového článku bola zvolená pomernej jednoduchá zostava nádoby, elektród, prírodných trubíc a ďalších komponentov.



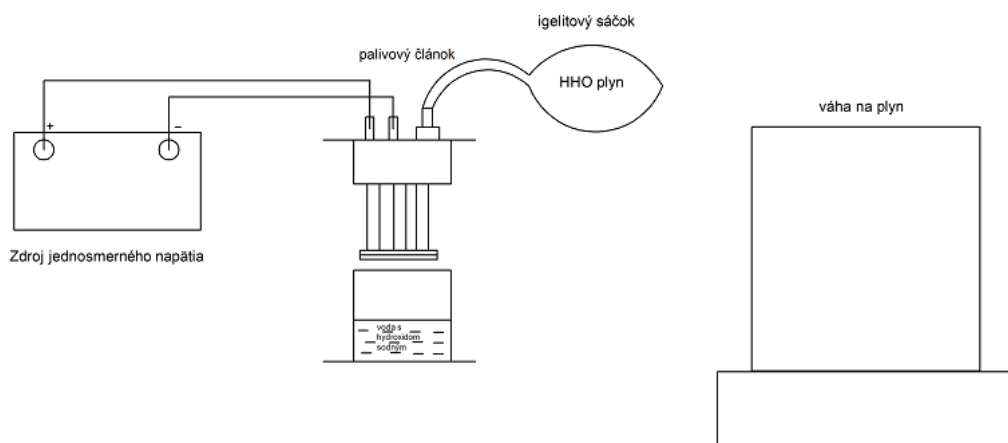
Obr. 2 Vytvorený elektrolyzátor pre generovanie paliva pre palivový článok

Popis zostavy prvkov: osem skrutiek (dĺžky 6 cm), 2 kovové podložky (priemer 1,5 cm), dve plastové podložky/gumené podložky s rovnakým priemerom ako kovové podložky, rúrka PVC (priemer 50 cm, dĺžka 10,5 cm) dve zátky na túto rúrku, jedna plastová hadička (priemer podľa mosadzného ventilu pripojeného k plastovej rúrke), 2 skrutky a 2 matice na pripojenie napájania s palivovým článkom v PVC rúrke, dva medené vodiče priemeru 1,5 mm, dĺžka podľa potreby.

A. Schéma merania na elektrolyzátori

Zdroj jednosmerného napätia pripojíme na skrutky elektrolyzátoru. Prechodom prúdu sa spustí elektrolyza vody a vytvorí sa HHO plyn, ktorý sa odvádza cez plastovú trubičku do rezervoára (PVC vrecúško). Po uplynutí 3 minút odpojíme palivový článok od napájacieho napätia.

Rezevoár (PVC vrecúško) uzavrieme izolačnou páskou, aby sme zabránili úniku vytvoreného plynu a presunieme rezervoár s plynom na váhu, kde zistíme hmotnosť získaného plynu.



Obr. 3 Schéma merania na elektrolyzéri

III. VÝPOČET ENERGETICKEJ BILANCIE ZÍSKANÉHO PLYNU

Energetická hodnota HHO plynu s hmotnosťou 0,4359 g

$$\frac{0,4359\text{g}}{\frac{2,016\text{ g}}{\text{mol}} + \frac{16\text{ g}}{\text{mol}}} \cdot \frac{2\text{ g}}{\text{mol}} = 0,0484\text{ g}$$

$$0,0484\text{ g} \cdot 120 \frac{\text{kJ}}{\text{g}} = 5,808\text{ kJ}$$

Na výpočet energetickej hodnoty vodíka pri jeho hmotnosti 0,4359 g musíme najskôr premeniť hmotnosť vodíka na jednotku v móloch. Molárna hmotnosť vodíka je 2,016 g/mol. Molárna hmotnosť vody je 16 g/mol.

Táto energia je uvažovaná v plynnom skupenstve vodíka, kde je vodík zmiešaný s kyslíkom v pomere 2/3, percentuálne je to 67% vodíka a 33% kyslíka.

IV. SPÔSOB VYUŽITIA ZÍSKANÉHO PLYNU

A. Výpočet potrebného množstva plynu

Meraním sme zistili množstvo získaného HHO plynu. Výpočtom sme získali jeho energetickú hodnotu počas horenia, ale daný plyn bolo ešte potrebné ďalej spracovať, aby sme mohli danú energiu z plynu využiť. Rozhodli sme sa preto, že daný plyn použijeme ako prímes do zemného plynu, ktorým budeme napájať plynový kondenzačný kotol VITODES 100-W, ktorý umožňuje spaľovať zemný plyn a vodík v pomere 80% (zemného plynu) a 20% (vodíka). Cena kotla je 1980,50 eur. Vybrali sme si kotol s výkonom 3,2-25 kW a spotrebou paliva 2,6 m³/h. Ďalej uvádzame výpočet pre zistenie množstva vyrobeného paliva pre zaistenie 20% spotreby vodíka pre daný kotol.

$$100\% \cdot x = 20\% \cdot 2,6\text{ m}^3/\text{h}$$

$$x = \frac{0,52\text{m}^3/\text{h}}{100\%}$$

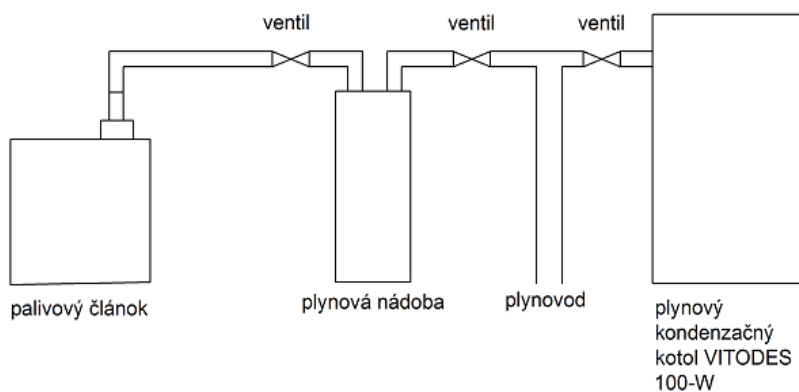
$$x = 0,52\text{ m}^3/\text{h}$$

$$0,52\text{ m}^3/\text{h} = 520\text{ l/h}$$

Zistili sme množstvo HHO plynu ako zmesi privádzaného do plynového kondenzačného kotla. Daný plyn potrebujeme aj niekde akumulovať (skladovať), preto sme si vybrali nádobu na uskladnenie získaného plynu, ktorá je z vysokotlakovej ocele z pozinkovaného materiálu. Jej cena sa pohybuje od 2733 eur po 4555 eur.

B. Analýza napájanie kotla plynom generovaným v elektrolyzéri

Návrh zapojenia palivového článku k zbernej nádobe plynu, ich prepojenie k plynovému kondenzačnému kotlu VITODES 100-W, je uvedený na nasledujúcom obrázku.



Obr. 4 Schéma zapojenia plynového kotla

C. Množstvo palivových článkov

Na získanie 520 litrov HHO plynu budeme potrebovať pri úvahe, že náš elektrolyzér pracuje konštantne 30 minút, tak vieme získať 4,359 g HHO plynu.

$$\frac{520\,000\text{ g}}{4,359\text{ g}} = 119\,294 \text{ článkov, aby sme zaistili stály chod plynového kondenzačného kotla.}$$

V. ZVÄČŠENIE ÚČINNOSTI ELEKTROLYZÉRA

V našom prípade sa účinnosť zvýši vyrobením väčšieho elektrolyzéra, pretože chceme HHO plynom napájať kotol. Existujú aj ďalšie možnosti ako zvýšiť účinnosť. Úplné ponorenie elektród tak, aby sa zväčšila aktívna plocha elektród.

Použitie plochých elektród a nie skrutiek ako v našom prípade tak, aby sme uľahčili vytváranie bublínok na elektródach a ich následné unikanie z elektrolyzéra.

Ďalší variant je využiť tzv. bipolárne usporiadanie elektród elektrolyzéra. Napätie v článku sa rovnomerne rozloží a celkové napätie je dané súčtom napätí v jednotlivých bunkách. V porovnaní s monopolárnym zapojením je napájacie napätie vyššie, no napätie medzi susednými článkami nižšie. Výhodou je vyššia efektívnosť a menšie zahrievanie vyvíjajúca. Naš elektrolyzér sme mali zapojený monopolárne, teda jedna strana bola anóda a druhá katóda.

Ďalší variant je vytvoriť nádobu pripojenú k elektrolyzérovi s dávkovaním elektrolytu počas reakcie, aby sa zamedzilo prísunu kyslíka k elektródam.

Ďalší variant je použitie nehrdzavejúcich elektród na zvýšenie životnosti elektrolyzéra.

VI. ZÁVER

Týmto príspevkom sme chceli poukázať na možnosti využitia paliva generovaného v elektrolyzéri pre palivové články. Ďalej sme poukázali na pomerne jednoduchý princíp fungovania a prečo sa nazývajú zdrojmi zelenej energie. Je to aj určitý návod, ako sa dá vytvoriť elektrolyzér doma s dostupných materiálov a k čomu sa dajú využívať. Cieľom príspevku bolo vytvoriť jednoduchý elektrolyzér pre výrobu paliva pre palivový článok, získať palivo a nájsť vhodné uplatnenie, kde by bolo možné takýto druh paliva použiť. Toto palivo je vhodné pre zníženie ekologického zaťaženia. V našom prípade zníženie skleníkových plynov a spotrebu zemného plynu na vykurovanie v domácnosti.

POĎAKOVANIE

Tento príspevok vznikol vďaka Agentúre na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0576 a APVV-21-0312, ako aj Vedeckou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu SR a SAV, podporená grantom VEGA 1/0757/21.

ZOZNAM POUŽITEJ LITERATÚRY

- [1] WR Grove Esq. MA (1839) The London and Edinburgh Philosophical Magazine and Journal of Science 1838 (online) (citované 1.11.2021) Dostupné na internete: <https://zenodo.org/record/1431021#.YYL3gFXMKUk>
- [2] Historické články z elektrochémie, (online) (citované 1.11.2021) dostupné na internete: <https://web.archive.org/web/20131005022402/http://electrochem.cwru.edu/estir/hist/hist-14-Schoenbein.pdf>
- [3] The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science (online) (citované 1.11.2021) dostupné na internete: <https://zenodo.org/record/1431031#.YYL5zVXMKUk>
- [4] Francis Thomas Bacon, biografia (online) (citované 1.11.2021) dostupné na internete: <https://www.britannica.com/biography/Francis-Thomas-Bacon>
- [5] CleanTechnica. Autá na vodíkové články (online) dostupné na internete: <https://cleantechnica.com/2021/10/14/why-are-3-automakers-still-hyping-hydrogen-fuel-cell-vehicles/>
- [6] International association for hydrogen energy, biografia Rogera E. Billingsa (online) (citované 1.11.2021) dostupné na internete: <https://www.iahe.org/advisory.asp?did=2>

- [7] Technology validation overview, Hydrogen program (online) Dostupné online: https://www.hydrogen.energy.gov/pdfs/progress10/viii_0_technology_validation_overview.pdf
- [8] Honda FC stack, Fuel cell scooter (online) (citované 1.11.2021) Dostupné online: https://web.archive.org/web/20070402035611/http://world.honda.com/news/2004/2040824_03.html
- [9] Internet archive wayback machine, Boeing successfully flies fuel cell-powered airplane (online) dostupné na internete: https://web.archive.org/web/20130509091442/http://www.boeing.com/news/releases/2008/q2/080403a_nr.html
- [10] EYE ON, Super-stealth sub powered by fuel cell (online) dostupné na internete: <http://edition.cnn.com/2011/TECH/innovation/02/22/hybrid.submarine/>
- [11] Manny Salvacion, Yibada, China presents the world s first Hydrogen-Fueled tram (online) dostupné na internete: <https://en.yibada.com/articles/21142/20150321/china-worlds-first-hydrogen-fueled-tram.htm#>
- [12] ALSTOM, unveils its zero-emission train (online) dostupné na internete: <https://www.alstom.com/press-releases-news/2016/9/alstom-unveils-its-zero-emission-train-coradia-ilint-at-innotrans>
- [13] Molly Burgess, H2 View, contract signed for first-ever hydrogen-powered train in the US (online) dostupné na internete: <https://www.h2-view.com/story/contract-signed-for-first-ever-hydrogen-powered-train-in-the-us/>
- [14] Juraj Procházka: JAXA a Honda pripravujú energetický systém pre vesmírne základne (online) 15. júna 2021 (citované 10.11.2021) dostupné na internete: <https://techbox.dennikn.sk/jaxa-a-honda-pripravuju-energeticky-system-pre-vesmirne-zakladne/>
- [15] EnergyGroove Fuel Cells (online) Dostupné na internete: <https://www.energygroove.net/fuel-cells/>
- [16] MARIÁN KŮRTHY Bakalárska práca: PALIVOVÝ CLÁNEK JAKO ALTERNATIVNÍ ZDROJ EL. ENERGIE 20. 11. 2008 (citované 28.11.2021) dostupné online: https://www.vut.cz/www_base/zav_prace_soubor_verejne.php?file_id=18463
- [17] Ing. Zdenek Porš palivové články 2002 (citované 28.11.2021) dostupné online: <https://www.cez.cz/edee/content/file/vzdelavani/palivove-clanky.pdf>
- [18] Karim Nice, Jonathan Strickland, How fuel cell work (online) dostupné na internete: <https://auto.howstuffworks.com/fuel-efficiency/alternative-fuels/fuel-cell.htm#pt2>
- [19] Ľuboš. Šárpataky, bakalárska práca, možnosti využitia palivových článkov (citované 14.12.2021) Bakalárska práca: 2018
- [20] Bc. Ján Polačík, diplomová práca, Pulzní generování vodíku (citované 20.10.2022) dostupné online: https://www.vut.cz/www_base/zav_prace_soubor_verejne.php?file_id=103248
- [21] Abdelhalim abdelnaby Zekry, ReaserchGate, výpočet vodíka elektrolyzéróm (online) (citované 28.10.2022) dostupné online: https://www.researchgate.net/post/How_to_calculate_the_amount_of_Hydrogen_produced_by_an_Alkaline_Water_Electrolyser_How_much_water_is_needed_to_produce_this_amount_of_Hydrogen
- [22] ASTRIA. VISSMAN VITODENS 100-W. [Online] [Dátum: 05.05.2023.] Dostupné online: <https://www.atria.sk/viessmann-vitodens-100-w--25-kw/>
- [23] Viessmann s.r.o. VITODENS 100-W [Online] [Dátum: 05.05.2023.] Dostupné online: https://www.abcplyn.sk/tepelna-technika/kondenzacne-kotly/pdf/Vitodens100-W_WB1C_2012_web.pdf
- [24] TŮV SŮD SLOVAKIA. Skladovanie, preprava a distribúcia vodíka. [Online] [Dátum: 05.05.2023.] Dostupné online: <https://www.tuvsud.com/sk-sk/odvetvia/energetika/konvenca-energia/vodikova-energia/skladovanie-preprava-a-distribucia-vodika>

Information System for Voting Support Using Blockchain Technology

¹Katarína CHAVKOVÁ, ²Branislav MADOŠ,

^{1,2}Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

¹katarina.chavkova@student.tuke.sk, ²branislav.mados@tuke.sk

Abstract – Voting is a fundamental aspect of democratic societies, and it is important to ensure that the process is transparent, secure, and reliable. Traditional voting systems have limitations such as centralization and susceptibility to manipulation, which can undermine the integrity of the election. Blockchain technology offers a decentralized and secure approach to voting, which can ensure transparency and immutability. In this paper, we propose an information system for voting support using blockchain technology that provides public, semi-public, and private voting options. The result of the work is a functional voting system where the use of blockchain ensures transparent voting and voting results protected from any modifications.

Keywords – blockchain, voting, online voting, smart contract

I. INTRODUCTION

The democratic process requires an efficient and secure voting system that can guarantee the integrity of the election. The traditional voting systems have been used for many years, but they are vulnerable to various forms of manipulation, which can affect the outcome of the election. Blockchain technology provides a decentralized and secure approach to voting, which can ensure transparency, immutability, and security.

The goal is to design and implement a voting system that will find use in many types of voting, not only secret but also public or semi-public.

II. ANALYSIS OF VOTING

The concept of voting according to the article [1] represents a process by which an individual can express his preference, opinion or will on a specific issue or candidate.

There are several principles that the vote should meet, such as [2] [3]:

- Universality. By universality we mean allowing everyone to vote.
- Equality. Each voter has the same right to decide the outcome, so each voter's vote has the same weight when the votes are tallied.
- Directness. Each voter casts his vote directly, he cannot transfer this right to another representative.
- Freedom of choice. The voter casts his vote for the candidate freely without influence or restrictions or sanctions.

A. Public voting

The process in which voters cast their votes publicly is called public or open voting. Not only the voters' votes, but also the voting results are available to the public with this type.

To the traditional options used in public or open voting includes voting by raising the hand and stating the voter's preference.

1) *Voting by show of hands*: Voting by show of hands is one of the most widespread voting methods. In this type of voting, individuals cast their vote by raising their hand. This method is usually used in small meetings or gatherings where a show of hands is practical and visible when votes are counted.

The benefits of a show of hands vote include simplicity, speed and transparency. On the other hand disadvantages of a show of hands includes privacy and inaccuracy.

2) *Preference Voting*: Preference voting is a method of voting in which voters express their choice by stating their preference.

The benefits of a show of hands vote include simplicity, speed and transparency. On the other hand disadvantages of a show of hands includes privacy and inaccuracy.

In both cases of voting, privacy is classified as an advantage as well as a disadvantage. The exact classification depends on the topic of the vote, where this fact can be positive for some, and negative for others.

B. Private voting

Private voting is a method of voting in which it is secret, or anonymous every vote cast. Traditional secret ballot methods include, for example: voting using a paper ballot and voting in polling stations.

1) *Paper ballot*: With ballots, voting requires a certain process. The voter marks his preferences without indicating his identity on a paper ballot, which he then throws into the ballot box. After the election, the ballot box is opened and the votes are counted. This is the most common way to vote in elections around the world.

The advantages of paper voting include simplicity and usability. On the other hand disadvantages of paper voting include time requirement and cost.

C. Semi-public voting

Semi-public voting is a method of voting where votes are collected publicly among a group of certain people, but only the result of the vote is known to the public. With this type of voting, it is possible to use all options, but it is necessary to ensure the principle of semi-public voting, where the voter's vote is known only to the voters present and only the results of the vote are known to the public.

D. Online voting

Online voting, also known as Internet voting, is a method of voting that allows individuals to cast their vote over the Internet using a computer or mobile device. Advantages of online voting include speed and credibility. Disadvantages of online voting include acces to technologies.

Election issues, or voting in general, but also the requirements that online elections should meet, have been addressed by several authors, of which we will cite a couple of examples. The book [4] provides an overview of the technical and political challenges associated with electronic voting, and the book [5] provides a comprehensive introduction to electronic voting, and the author duo [6] discusses the legal implications of blockchain technology, including potential uses in online voting systems. Another resource used in the analysis of requirements for an online voting system is the book [7], which provides a comparative analysis of online voting systems in different countries, including the challenges and opportunities they present.

From the above sources, we created a list of voting system requirements:

- Security. The voting system should be secured in such a way as to prevent manipulation, fraud and errors, which can also be ensured by using blockchain technology.
- Reliability. The voting system should be able to function efficiently and accurately without technical problems or glitches that could affect the integrity of the vote.
- Transparency. The voting system is supposed to ensure results that accurately reflect the will of the voters.
- Accessibility. The voting system should be accessible to all eligible voters regardless of location or physical ability.
- Efficiency. The voting system should be efficient, with the ability to accurately and reliably count votes in a timely manner.
- Credibility. Voters should trust the voting system.
- Equality. Each voter has an equal right to decide the outcome, where each voter's vote has equal weight in the tallying of votes.
- Freedom of choice. The voter casts his vote for the candidate freely, without influence, restrictions or sanctions.
- Ease of use. It is important that voting systems are user-friendly and easy to use, thus ensuring accessibility for all eligible voters
- Avoiding double voting. The voting system should allow each voter to vote in a particular ballot only once.

In the following subsections, you we will present a case study of an online voting system with the use of blockchain but also without its use.

Since the aim of the work is to create a voting system for smaller groups of people, and therefore smaller elections, it will not be necessary to ensure all the procedures, as with the following systems, which, on the contrary, are also used in national elections.

III. BLOCKCHAIN TECHNOLOGY

The concept of blockchain technology was first introduced in 2008 as the underlying technology for the digital currency bitcoin. The goal was to create a secure, decentralized system for recording and tracking transactions without the need for a central authority [8]. Definitions say that blockchain is a decentralized digital ledger that records transactions across a network of computers in a secure, transparent, tamper-proof manner [9].

Over the years, the blockchain concept has evolved beyond digital currencies and is now used in various industries such as supply chain management, voting systems and digital identity management. This is due to the inherent advantages of the technology, such as transparency, immutability and security.

As the technology matured, new innovations and implementations emerged, including private and consortium blockchain, smart contracts, and decentralized applications.

Blockchain consists of elements that make up its structure, where belongs blocks, transactions, chain and distributed ledger. The blockchain structure also refers to the way data is organized and stored in the blockchain network. Blockchain network is composed of a series of blocks connected in a chain. Each of the blocks contains a set of transactions and a reference to the previous block. Overall, the blockchain structure is designed to create a secure, tamper-proof record of transactions that can be trusted by all parties involved. The decentralized nature of blockchain ensures that no single entity has control over the network, making it a powerful tool for building trust in a variety of contexts.

There are several types of blockchain, which according to [10] include public, private, consortium, and hybrid. In this section, we will describe these types in more detail. Public is a fully decentralized network that is not controlled by any signaling entity or node, so anyone can join. The decision is made based on the consensus of every node in the network, and transactions are recorded with a time stamp. There are rewards and incentives [11] for network participants. A private blockchain is a semi-centralized network, meaning that nodes need consent to enter the network. Transactions in this type of blockchain are not public and only authorized authorities can work with the data [12]. Consortium is more decentralized than private, as the authority is shared by several organizations, which prevents control by one organization. A hybrid blockchain has the inherent transparency of a public blockchain to create trust in the network and the security and privacy of a private blockchain, making it easy for multiple businesses [13].

IV. FUNCTIONAL REQUIREMENTS OF A VOTING SYSTEM USING BLOCKCHAIN

In this section, we list the functional requirements for the proposed system to which they belong

- 1) Access to the system is allowed only for responsible persons, who will have access to view the created votes, to create new votes, to start and end the vote, but also to display the results of the votes.
- 2) Creating a ballot with the option of choosing the type of vote - when creating a ballot, it is necessary to enter the name of the vote, the number of voters, the options from which the voters will have a choice, and it will also be necessary to choose the type of vote, e.g.
 - secret - in this type of voting, voters will cast their votes without revealing their identity and only the voting result will be recorded in the blockchain,
 - public - every vote sent by a voter is recorded in the blockchain together with his name and also the result of the vote,
 - semi-public - voters cast their vote by stating their name, but only the result of the vote is subsequently recorded in the blockchain.
- 3) The start of voting is the same for all types of voting, the authorized person selects a vote from the list of prepared votes and then starts voting by pressing the button. Once voting starts, voters will be presented with options on their devices to cast their vote.
- 4) Voting by voters - voters have the opportunity to vote for one or more options listed.
- 5) End of voting - in secret and public voting, it is possible to end by pressing the button to end voting, which will be displayed if the number of votes cast is equal to the number of voters entered when creating the ballot, on the contrary, in semi-public voting, the user will first see the button for displaying the results and only then to end the voting.
- 6) Viewing the results of completed votes - after voting, the user is automatically redirected to the list of completed votes together with their results, which he can access at any time.

V. IMPLEMENTATION

This section of the work is devoted to the actual implementation of the system, the proposal of which was described in the previous sections. We will gradually describe the steps that had to be taken in order to achieve a functional prototype of the voting system using blockchain technology.

A. Environment Preparation

The preparation of the environment is the most essential step in the implementation, which cannot be skipped or replaced. A properly prepared environment is the main tool for successful implementation, by initially installing all the necessary technologies that will be used during the implementation.

1) *Ganache*: We start the installation of the Ganache tool using the command line by entering the command `npm install ganache`. After successful installation by entering the command `ganache` we will create a private local blockchain network for development and testing. In the output, we get a list of available accounts together with their addresses and private keys, as well as information about the network to which they belong, e.g. network identifier, block gas limit and under.

2) *Truffle*: Truffle, which will be used to connect the web application with the blockchain, will be installed using the command line, by entering the command `npm install truffle`. We can initialize the project by entering the command `truffle init`, which will create several folders.

3) *MetaMask*: We will ensure the connection of the MetaMask wallet together with the local blockchain network that we created and launched in the first step by creating a new network using the parameters that we obtained in the output of the first step. The required data to create a network includes the string ID, RPC URL, and network name.

4) *Creating a smart contract*: By initializing the project, which took place in the second step through the command `truffle init`, we created, among others, a folder called `contracts`, in which the contracts created by us, written in the Solidity language, will be stored. After writing a smart contract, or smart contract, we will create a migration in the `migrations/` folder, which was also created when the project was initialized. Through migration, we will ensure the deployment of a smart contract in the blockchain network. With the command `truffle migrate`, we will start all migrations stored in the `migrations` folder. If we get the successful message at the output, it means that our smart contract, or smart contract, is stored on the blockchain and the first account from the list of accounts obtained in the first step is considered the owner. After a successful migration, an ABI file of the given smart contract will be created in the `build/contracts` folder, which will be used for communication between the web application and smart contracts.

5) *Smart contract connection*: MetaMask injects a global API into web pages visited by users at `window.ethereum`. After instantiating the smart contract, we can start calling functions, using the ABI file that was created by the `.fd .fd truffle migrate` in the previous step.

B. System Startup

In order to ensure that only authorized persons have access to the system, we created a function `isOwner()` in the smart contract, which checks whether the address from MetaMask wallet, is the address of the authorized person.

If the function returns the value `false`, it means that the user is not authorized to enter the application. Conversely, if the function returns the value `true`, the user is redirected to a page with a list of votes that have not yet taken place. A person authorized to enter the system has the ability to create new ballots, start and end voting and display voting results.

C. Creating a ballot

When creating a ballot, the user must fill in all fields, namely the type of vote, the name of the vote, the number of voters and at least two options. If the user does not fill in all the necessary data, he will be prompted to do so. After filling in and sending all the necessary input data, the ballot data is saved in the database and the user is redirected to the list of prepared ballots. There is also database for storing ballot data, which consists of two tables. The ballots table includes data such as the name of the ballot, the maximum number of voters, the current number of voters, the type of vote and the status of the vote. This table is linked by a primary key to the options table, which stores data such as the name of the option, the primary key of the ballot, the number of voters and the voters of the given option. The function `save()` will ensure the creation of the ballot and saving the data to the database.

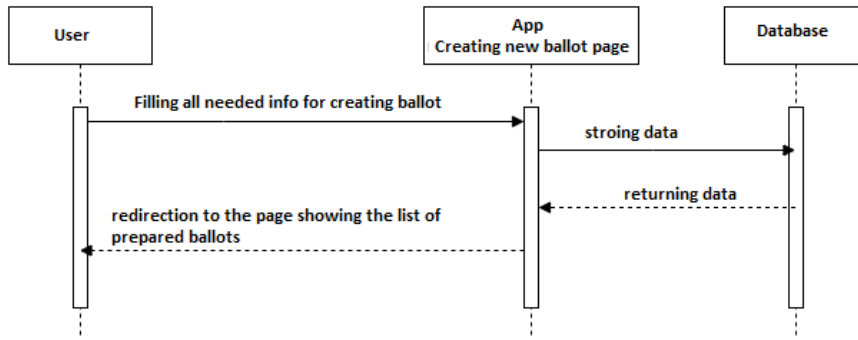


Fig. 1 Process of creating ballot.

D. Start voting

After clicking on the ballot selected by us displayed in the list of prepared ballots, the user will see a message whether he is sure about starting the given vote and a button for starting the vote. Function used to start the vote uses `socket.emit()` to announce the start of the vote. The function `socket.on()` makes changes to the event that is created by the function `socket.emit()`. In this section, the ballot data will be updated in the database, where the voting status will be changed to active. A new event is then created to announce that the election has started. When the poll is announced, the user is redirected to a page showing the number of votes already submitted, and voters see the name of the poll and all the options on their screens.

E. Vote Casting

Voting will take place on the voter's device, which is provided to him and on which he has the opportunity to choose one or more options. After clicking on the submit button, a field for entering a name/code will appear if it is a public or semi-public vote, followed by a question whether you agree with your choice. If it is a secret vote, only a consent message will be displayed to the voter.

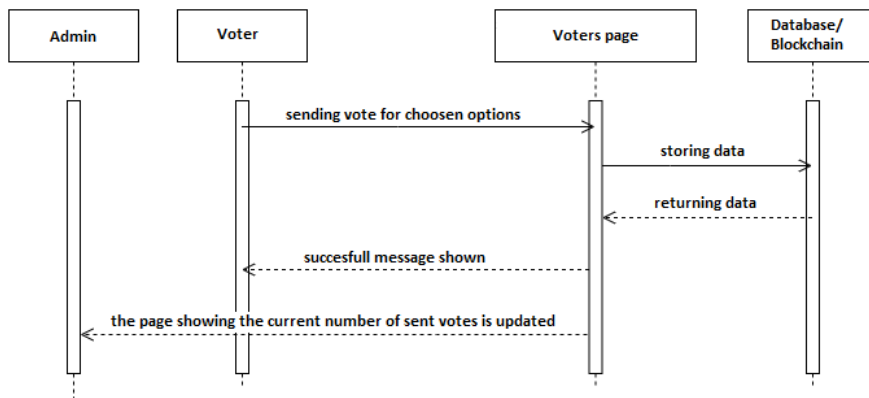


Fig. 2 Process of sending vote.

F. End of voting

If the same number of votes as entered when creating the ballot is submitted, the user will see a button to end voting on the page showing the number of submitted votes. If it is a semi-public vote, instead of the button for ending the vote, the button for displaying the results will be displayed, and only then can the vote be ended. After pressing the button to end voting, the result data is saved to the blockchain.

G. User Interface

The web application is composed of several parts, which include a home page, a list of ballots that have not yet been cast, a list of ballots that have already been cast, a page for creating a

new ballot, and a page showing the number of votes cast during voting in progress. Only the site where they cast their votes is available to voters. All the mentioned parts were implemented using the ReactJs library and the Bootstrap framework.

VI. RESULT

The result of the work is a functional voting system, which currently supports one voting administrator. The administrator has the option of creating a new ballot, where he must enter all the necessary data, namely the type votes - secret/public/semi-public, vote name, options and maximum number of voters. All this data is stored in the database. Another option is starting and ending voting, in which the data in the database for specific they update the vote by changing its status. In public voting, every single vote cast is stored in the blockchain voters and the results of public voting. In the case of semi-public and secret ballots, only voting results are stored in the blockchain. Use of blockchain will ensure transparent voting and voting results protected from any modifications. Currently, the system uses a local blockchain created using Ganache, which generated 10 test accounts for us.

VII. CONCLUSION

The aim was to design and implement a voting system using blockchain technology. The primary use of the voting system is at smaller votes, such as the resolution of various issues during municipal elections councils, in education, e.g. also when electing a school principal or deputy principal, and so on. Design and implementation are preceded by a thorough analysis of related resources. After the analysis of all possibilities, the design of the system followed. The first step of the design was the definition of functional requirements for the system, to which they belonged, e.g. creating a ballot for different types of elections, saving voting results in the blockchain, but also the possibility of voting by voters to one or more options. The implementation section describes all the steps that were needed to perform. The first was essential preparation of the environment, to which it belonged, e.g. initialization of the project, creation of a local blockchain, creation of a smart contract and its connection with the application and subsequent implementation of the functions that were necessary to fulfill functional requirements.

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: 313011W988), co-financed by the ERDF. We would like to express our sincere gratitude to all those who contributed to the development of this proposed information system for voting support using blockchain technology. We would also like to thank our colleagues who provided valuable feedback and suggestions throughout the development process.

REFERENCES

- [1] M. S. Shugart and M. P. Wattenberg, *The Effect of Voting Systems on Political Parties: A Review*.
- [2] C. R. Hughes, *The Future of Voting: How Technology is Transforming Democracy*.
- [3] D. W. Jones and B. Simons, *Voting Technology: The Not-So-Simple Act of Casting a Ballot*.
- [4] J. A. Halderman and E. W. Felten, *Electronic Voting: An Introduction*.
- [5] D. W. Jones, *Electronic Voting Systems: A Beginner's Guide*.
- [6] P. D. Filippi and A. Wright, *Blockchain and the Law: The Rule of Code*.
- [7] M. D. McDonald and D. O. Sears, *Online Voting and Democracy: A Comparative Analysis*.
- [8] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Bitcoin.org*, 2008.
- [9] X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen, "A survey on the security of blockchain systems," *Future generation computer systems*, vol. 107, pp. 841–853, 2020.
- [10] "Unfolding the blockchain era: Timeline, evolution, types and real-world applications," *Journal of Network and Computer Applications*, vol. 207, p. 103511, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1084804522001527>
- [11] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, "An overview of blockchain technology: Architecture, consensus, and future trends," in *2017 IEEE international congress on big data (BigData congress)*. Ieee, 2017, pp. 557–564.
- [12] S. Pongnumkul, C. Siripanpornchana, and S. Thajchayapong, "Performance analysis of private blockchain platforms in varying workloads," in *2017 26th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 2017, pp. 1–6.
- [13] G. Ateniese, M. T. Chiaramonte, D. Treat, B. Magri, and D. Venturi, "Hybrid blockchain," May 1 2018, uS Patent 9,959,065.

Vzdelávanie prostredníctvom hier

Roland JANSKÝ, Anton BALÁŽ, Branislav MADOŠ

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

roland.jansky@student.tuke.sk, anton.balaz@tuke.sk, branislav.mados@tuke.sk

Abstrakt - S narastajúcim počtom pripojených zariadení sa v dnešnom svete zvyšuje závislosť na digitálnych technológiách. Tento trend so sebou prináša zvýšený výskyt rizík v oblasti bezpečnosti informačných a komunikačných technológií (IKT), ktoré majú čoraz väčší dosah na fungovanie moderných spoločností. Problematika bezpečnosti IKT neustále narastá a vyžaduje si aktívne opatrenia na zlepšenie ochrany informácií a infraštruktúry. Primárnym opatrením je práve prevencia, ktorej cieľom je pripraviť používateľov na núdzové situácie, zabezpečiť ochranu súkromia a dôvernosti informácií. Prevencia sa zvyčajne zameriava na edukáciu a informovanie používateľov o bezpečnom používaní IKT. Tieto aktivity zahŕňajú tréningy a školenia na identifikáciu rizík a správne používanie bezpečnostných protokolov a nástrojov. V tejto oblasti našla uplatnenie gamifikácia, ktorá prostredníctvom herných prvkov v nehermom prostredí motivuje používateľov zvyšovať si povedomie o kybernetickej bezpečnosti.

Kľúčové slová – Gamifikácia, Hra, Bezpečnosť, IKT

I. ÚVOD

Populárne komunikačné platformy umožňujú používateľom komunikovať takmer s celým svetom, zdieľať informácie, vytvárať a udržiavať sociálne kontakty. S rozvojom sociálnych sietí narastá popularita gamifikácie, ktorá umožnila používateľom hrať hry nielen so svojimi priateľmi, ale aj s neznámymi hráčmi. Hlavnou úlohou gamifikácie je však aplikovať herné prvky do neherného prostredia s cieľom motivovať, zapájať a zlepšovať zážitky používateľov pri používaní určitých produktov alebo služieb. Gamifikácia si našla uplatnenie najmä v marketingu alebo obchodovaní. Veľký potenciál však vidíme vo vzdelávaní. Využitie našla v školskom prostredí na zvyšovanie vedomostí študentov prostredníctvom motivácie. Rôzne organizácie ju využívajú na vzdelávanie a tréning svojich zamestnancov.

Práve tak populárne sociálne siete môžu byť miestom zneužitia osobných údajov, šírenia nepravdivých informácií, kybernetickej šikany a iných problémov, ktoré ohrozujú bezpečnosť a súkromie používateľov. Bežní používatelia, ktorí nie sú dostatočne oboznámení s bezpečnosťou na internete si častokrát spôsobia problémy vlastným pričinením tým, že si nechrania heslá, klikajú na rôzne odkazy a neaktualizujú svoj softvér na najnovšiu verziu. Takéto správanie môže viesť k zneužitiu ich zariadenia, strate osobných údajov a citlivých informácií. Medzi najväčšie hrozby bezpečnosti patrí malware, phishing, identitné krádeže, hackerstvo, ransomware, zneužívanie sociálnych sietí, nezabezpečené Wi-Fi siete a pod.

Ak sa zameriame na vzdelávanie, práve gamifikácia by mohla pomôcť používateľom zvyšovať povedomie o kybernetickej bezpečnosti. Napríklad v rámci vzdelávania by sa mohli využiť herné prvky na motivovanie študentov k zlepšeniu ich bezpečnostných návykov, ako je používanie silných hesiel alebo aktualizovanie softvéru. Rôzne organizácie by mohli využiť gamifikáciu na zvýšenie angažovanosti zamestnancov v oblasti kybernetickej bezpečnosti a motivovať ich k pravidelnému aktualizovaniu softvéru, používaniu silných hesiel a iných bezpečnostných opatrení. V konečnom dôsledku by gamifikácia mohla pomôcť zlepšiť bezpečnosť IKT a naučiť používateľov zvládať kybernetické hrozby.

V našom príspevku sa budeme venovať oboznámeniu s problematikou gamifikácie a jej existujúcimi súčasnými prevedeniami. Tieto poznatky sme implementovali do vytvorenia systému praktického riešenia vzdelávania bezpečnosti IKT prostredníctvom hry v 3D priestore, ktorá obsahuje rôzne scenáre a situácie riešiteľné pomocou správnych bezpečnostných opatrení.

II. GAMIFIKÁCIA A JEJ VYUŽITIE VO VZDELÁVANÍ

Hra je neoddeliteľnou súčasťou našich životov už od narodenia a prostredníctvom nej sa učíme základným pravidlám fungovania v spoločnosti. Existuje veľa definícií pojmu hra, ktoré sa zhodujú v tom, že za hru možno považovať akúkoľvek činnosť, ktorá poskytuje človeku príjemné pocity, a ktorá je motivovaná predovšetkým radosťou z činnosti samotnej, nie nutne z uspokojenia základných potrieb.

Gamifikácia sa dá definovať ako využitie herných konceptov a metafor na ovplyvňovanie správania, motiváciu a podporu spolupráce pri plnení úloh v reálnom svete. Pri gamifikácii hovoríme o prvkoch hry, nie hrania sa. Je dôležité rozlíšiť hru od hravej interakcie, hravosti alebo hravého dizajnu[1].

Gamifikácia je technika, ktorá predstavuje aplikáciu využitia prvkov hrania hry v inej činnosti. Cieľom je urobiť túto činnosť zaujímavou. V našom prípade to bude za účelom vzdelávania. Gamifikácia implementuje využitie skórovacích prvkov videohier, ako napr. body, levely, rebríčky a úspechy, ktoré sa aplikujú do pracovného alebo vzdelávacieho kontextu mimo herného prostredia. Tieto prvky sa môžu vyskytovať napr. v organizáciách, inštitúciách, webových stránkach, mobilných aplikáciách alebo iných online systémoch[2].

Zichermann a Cunninghame, svetovo uznávaní experti na gamifikáciu pojem definujú ako proces využitia herného myslenia a herných mechaník v nie herných situáciách zapojením užívateľského publika k riešeniu úloh a problémov z rôznych oblastí ako napr. vzdelávanie, marketing, dizajn[3].

Gamifikácia je tiež definovaná ako proces využitia poznatkov o správaní a motivácii používateľov v hrách pri riešení problémov v iných oblastiach. Dôležitým prvkom gamifikácie je spätná väzba, ktorá podporuje základné ľudské potreby, ako napríklad súťaživosť, odmeny, sebvýjadrenie a podobne[4].

Digitalizácia a rýchly rozvoj informačných a komunikačných technológií (IKT) priniesli do našich životov mnoho zmien. Jednou z týchto zmien je náš vzťah k informáciám a vzdelávaniu. Súčasná generácia mladých ľudí si nevie predstaviť svoj život bez internetu a mobilných zariadení. Ako teda zabezpečiť, aby boli tieto technológie správne využívané aj v oblasti vzdelávania? Moderná doba ponúka množstvo nových spôsobov učenia, vrátane elektronického vzdelávania (eLearning). Tento spôsob vzdelávania sa využíva v školstve, vzdelávacích inštitúciách, korporáciách a obchodných spoločnostiach ako alternatíva pre prípad, keď študenti alebo zamestnanci nemôžu byť prítomní na prednáškach. E-learning sa často spája s gamifikáciou. Študenti pri riešení ťažkých otázok alebo problémov získavajú okamžitú spätnú väzbu, stávajú sa samostatnejšími a silnejšími, keď sú schopní urobiť alebo pochopiť niečo samostatne. V súčasnosti sa gamifikácia a eLearning stávajú obľúbenými spôsobmi učenia, a preto sa niektoré firmy a inštitúcie snažia zlúčiť tieto dva spôsoby, aby zlepšili spôsob učenia pomocou internetu a zvýšili efektivitu zakomponovaním viac herných prvkov. Výskumy ukázali, že hry pri učení uvoľňujú dopamín, testosterón, endorfíny a dokonca môžu podporovať rast mozgu, a preto o eLearning a gamifikáciu rýchlo rastie záujem na trhu[5].

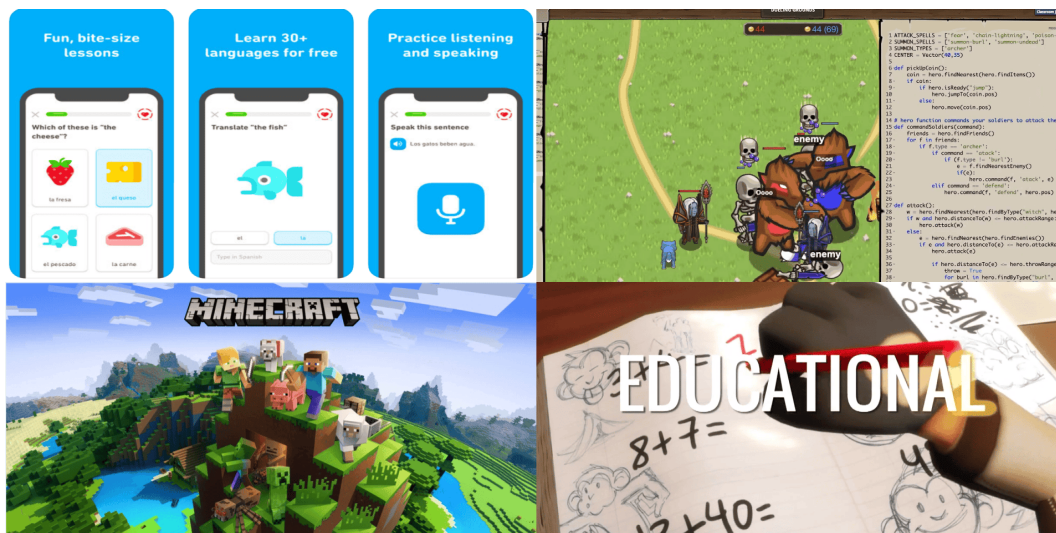
Vzdelávacie inštitúcie v minulosti používali rôzne herné prvky na zatriktívnenie vyučovania a zvýšenie motivácie študentov k lepším výkonom. Tradičné školské prostredie zvyčajne vedie k pasívnosti, podvádzaniu, naučenej bezmocnosti a niekedy aj k odchodu zo školy. Z toho vyplýva, že prítomnosť herných prvkov v učebnom pláne automaticky nezaručuje angažovanosť a úspech žiakov. Úlohou gamifikácie je z rôznych hier vyňať najužitočnejšie psychologické mechanizmy a využiť ich na zvýšenie motivácie a zaujatia študentov. Prináša snahu pochopiť, čím hra študenta zaujala a získané poznatky implementovať do učebného procesu. Gamifikácia umožňuje experimentovať s pravidlami, emóciami a sociálnymi rolami. Môže byť motivujúcim prvkom pre študentov a učiteľov, ktorí ju môžu využiť na zlepšenie angažovanosti a odmeňovania študentov. Avšak, existujú aj riziká spojené s gamifikáciou, ako je napríklad vyčerpanie zdrojov učiteľov alebo možnosť, že študenti sa budú učiť len kvôli externým odmenám. Ak sa hra stane povinnosťou, gamifikácia môže vytvoriť zážitky, ktoré budú pripomínať tradičné učenie v škole[1].

V Európe napriek tomu existuje málo praktických príkladov ako by sa gamifikácia dala komplexne využiť vo vzdelávaní. V súčasnosti funguje na báze vzdelávacích hier. **Monkey Tales** je edukačná hra od Larian Studios pre deti vo veku od 6 do 12 rokov, ktorá má za cieľ rozvíjať rôzne zručnosti, ako sú matematické a jazykové schopnosti, logické myslenie a pamäť. **Minecraft** je svetovo známa počítačová hra, ktorá slúži na výuku fyziky alebo občianskej náuky. Hráči sa pohybujú v otvorenom svete, kde majú neobmedzenú slobodu pohybu a aktivít. **Duolingo** je ďalším dobrým príkladom aplikácie, ktorej súčasťou je gamifikovaný učebný program. Duolingo je online platforma pre výučbu jazykov, ktorá je úplne bezplatná. Jej poslaním je sprístupniť výučbu jazykov pre všetkých ľudí. **CodeCombat** je edukačná hra, ktorá učí programovanie pomocou zábavných herných mechanizmov.

Vo svete začínajú vznikať seriózne hry, ktorých primárnym cieľom je hráča niečo naučiť, pričom na prvom mieste nie je zábava. Na tomto princípe vznikol americký portál **GlassLab**, ktorý s

podporou Gates Foundation a Mac Artur Foudation organizuje výukový proces prostredníctvom hier. Využívať ho môžu učitelia, žiaci a rodičia. Disponuje pozoruhodnou a jedinečnou funkciou nazývanou analýza výukových výsledkov. Táto funkcia ponúka učiteľovi možnosť spätnej väzby, čo sa žiak počas výučby naučil[6].

Quest to learn je škola, ktorá funguje už niekoľko rokov v New Yorku. V edukačnom procese sa používajú rôzne typy hier. Vyučovanie a prostredie sú založené na princípoch gamifikácie. Preferuje sa usmernenie študentov k samostatnému tvoreniu, inovácii, komplexnej práci, kritickému mysleniu a spoľahlivosti. Učebné osnovy sú navrhnuté podľa princípov dizajnu hier, pričom každá časť osnov je koncipovaná ako misia, ktorá zahŕňa herné stratégie, ako je spolupráca, hranie rolí a simulácia. Okrem toho, že študenti hrajú hry v triedach, učia sa ich aj vytvárať, aby preukázali svoje schopnosti v systémovom myslení. Vo vyučovacom procese využívajú rôzne typy hier. Učebný proces a prostredie sú gamifikované. Uprednostňujú princíp vedenia študentov samostatne tvoriť[7].



Obr. 1 Súhrn gamifikovaných vzdelávacích hier

III. BEZPEČNOSŤ INFORMAČNÝCH A KOMUNIKAČNÝCH TECHNOLOGIÍ

V 80. rokoch 20. storočia spisovateľ William Gibson použil po prvýkrát termín kybernetický priestor (cyberspace), ktorý vtedy ešte nemal jasný význam. Neskôr, s rozvojom internetu, sa však pojem kybernetický priestor stal označením pre globálnu technickú, informačnú a komunikačnú infraštruktúru. Informačná bezpečnosť mala široké pôsobenie a väčšina informácií sa spracovávala digitálnymi technológiami v rýchlo sa meniacom prostredí. S postupným vývojom kybernetického priestoru sa však pozornosť informačnej bezpečnosti sústredila najmä na ochranu elektronicky spracovaných informácií v digitálnom svete. Niekde tu sa začal používať pojem kybernetická bezpečnosť[8].

Termín kybernetická bezpečnosť zahŕňa technológie určené na ochranu počítačových systémov a dát pred neoprávneným prístupom. Ich súčasťou sú právne, organizačné, technické a vzdelávacie opatrenia, ktoré majú za cieľ zabezpečiť bezpečnosť informačných a komunikačných technológií v kybernetickom priestore. Hrozba môže predstavovať potenciálne ohrozenie bezpečnosti, ktoré sa nemusí nutne stať, ale ak áno, ide o bezpečnostný incident. V každodennom živote ľudí internet prináša veľké množstvo výhod, ale na druhej strane predstavuje aj hrozby zahŕňajú útoky zvonku, škodlivý softvér, neoprávnený prístup, ukradnutie dôverných informácií, phishing alebo pharming. Počítačové hrozby sa stále zdokonaľujú a spôsobujú významné škody. Môžu zablokovať celú sieť, poškodiť dáta a zamedziť prístup k nim, vydierať jednotlivcov a organizácie. Najčastejšie útoky ohrozenia bezpečnosti IKT predstavujú prípady, keď jednotlivec alebo skupina útočí na informačný systém pomocou softvéru, ktorý sa nazýva škodlivý softvér alebo malware. Jeho účelom je poškodiť, zničiť alebo získať neoprávnený prístup k informačnému systému a môže byť použitý na špionáž, zneužitie osobných údajov. Škodlivý softvér má rôzne podoby ako napr. **ransomware**, **spywe**, **adware**, **vírus**, **počítačový červ** alebo **trójsky kôň**[9].

Sociálne inžinierstvo je považované za jeden z najúčinnějších netechnických spôsobov, ako narušiť kybernetickú bezpečnosť. Jeho úspešnosť nezávisí na komplexnosti softvérového vybavenia, ani hlbokých znalostiach z oblasti počítačových systémov. Tento typ kybernetického

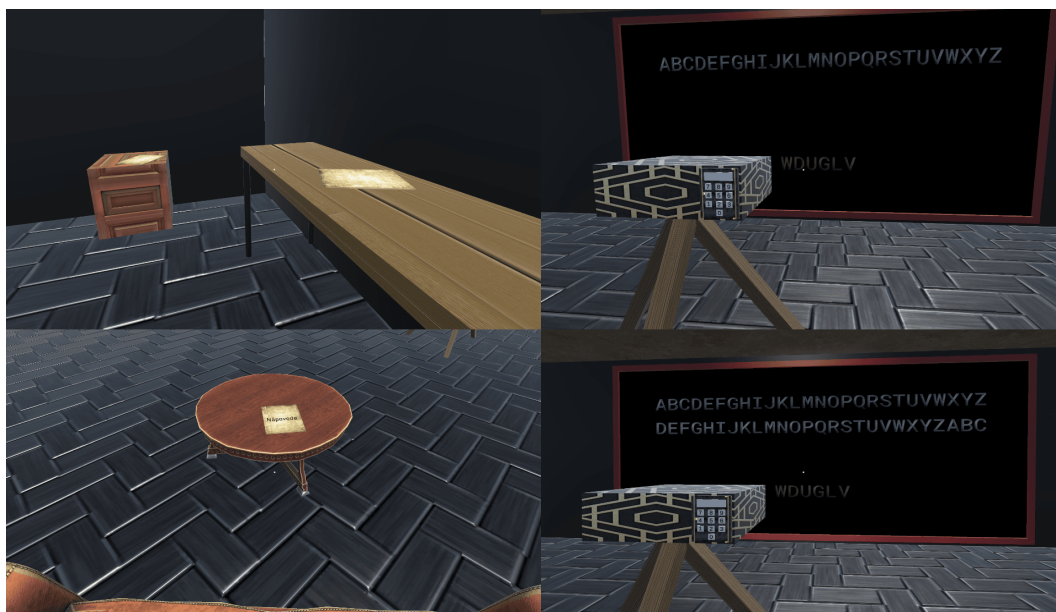
útoku závisí od ľudskej nerozvážnosti, naivity, nevedomosti a neopatrnosti. Na dosiahnutie cieľa sa využívajú rôzne manipulačné techniky, ktoré majú za cieľ získať citlivé informácie alebo prístup k informačným systémom. Aby sme sa pred týmto druhom hrozby ochránili, je potrebné použiť zdravý rozum, nezverejňovať citlivé údaje a zvyšovať si povedomie o informačných a komunikačných technológiách. Medzi najväčšie hrozby patria: **phishing**, **pharming**, **baiting** a **reverzné sociálne inžinierstvo**[10].

IV. VYUŽITIE NAVRHNUTÉHO RIEŠENIA NA GAMIFIKÁCIU VÝUČBY BEZPEČNOSTI V OBLASTI IKT

Našou úlohou bolo prostredníctvom diplomovej práce vytvoriť hru, ktorá implementuje teoretické poznatky o bezpečnosti IKT a integruje rôzne prvky a techniky na ochranu informačných systémov a údajov pred neoprávneným prístupom, zneužitím alebo poškodením. Hlavným cieľom hry bolo vytvoriť zaujímavé a motivujúce herné prostredie a zapojiť hráča do hry pomocou rôznych herných mechanizmov, ako sú hádanky, úlohy alebo výzvy a zároveň mu poskytnúť vedomosti o IKT bezpečnosti, ktoré budú pre neho užitočné.

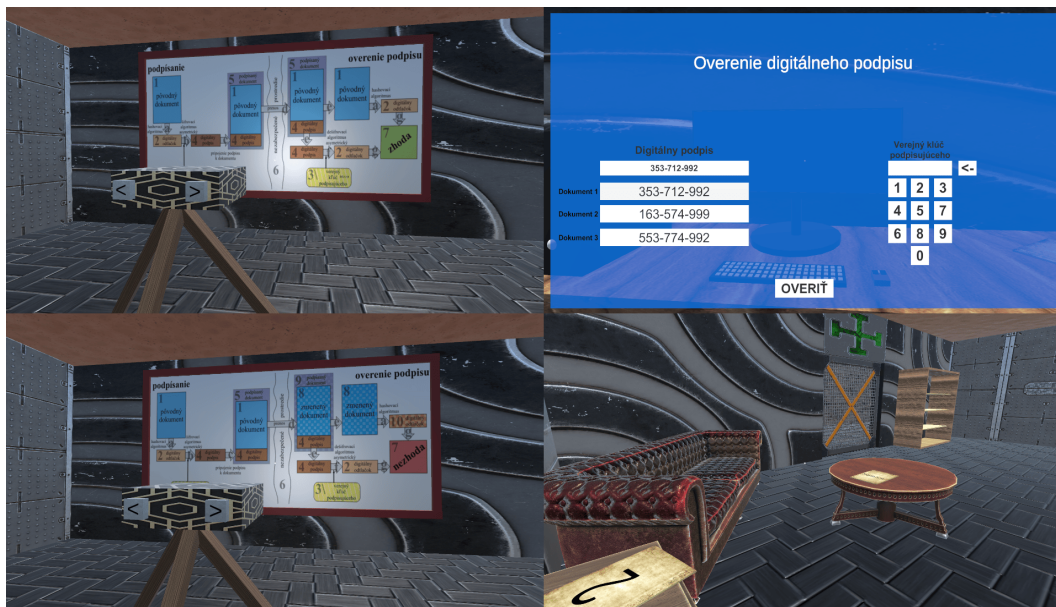
Hra **Kybernetický labyrint** je založená na koncepte Escape Room, kde je hráč uväznený v šiestich miestnostiach a musí riešiť rôzne hádanky, úlohy a výzvy z oblasti bezpečnosti IKT, aby sa postupne dostal von z miestností. Labyrint nepredstavujú jednotlivé miestnosti, ale riešená problematika v nich. Hráč musí využívať svoje schopnosti, kreativitu a nadobudnuté znalosti z miestnosti, ktoré následne aplikuje na problematiku, ktorá sa v nej nachádza, ako napr. dešifrovanie hesla a zistenie kódu, ktorý keď zadá, tak sa mu otvoria dvere. Dostane sa do ďalšej miestnosti, kde sa tento kolobeh opakuje, ale hráč rieši inú bezpečnostnú problematiku. Hra **Kybernetický labyrint** poskytuje hráčom zábavný spôsob, ako sa učiť a rozvíjať svoje znalosti z oblasti bezpečnosti IKT. Tento prístup umožňuje hráčom lepšie pochopiť a aplikovať riešenia na reálne problémy, ktoré sa môžu vyskytnúť v praxi.

V **prvej miestnosti** sme sa venovali teoretickým poznatkom o základných kryptografických nástrojoch - šifrách RC4, RSA a Cézarovej šifre. Tieto nástroje sú dôležité pre zabezpečenie komunikácie a uchovanie dôležitých dát v súkromí. Hlavným cieľom bolo prostredníctvom gamifikácie vysvetliť jednu z najjednoduchších šifier, teda Cézarovu šifru.



Obr. 2 Miestnosť 1 - Cézarova šifra

V **druhej miestnosti** sme sa zameriavali na digitálny podpis, ktorý predstavuje dôležitý nástroj pre zabezpečenie dôveryhodnosti, autenticity a nedotknuteľnosti dokumentov a správ. V digitálnom svete kde sa čoraz viac komunikácií a transakcií uskutočňuje online, je digitálny podpis nevyhnutnosťou. Umožňuje rýchle, efektívne a bezpečné overovanie identity osoby, ktorá dokument podpísala. Tiež zaisťuje, že dokument nebol zmenený po podpísaní a jeho obsah zostal originálny. Digitálny podpis predstavuje nevyhnutnosť pre každého, kto chce zabezpečiť dôveryhodnosť a bezpečnosť svojich digitálnych dokumentov.



Obr. 3 Miestnosť 2 - digitálny podpis

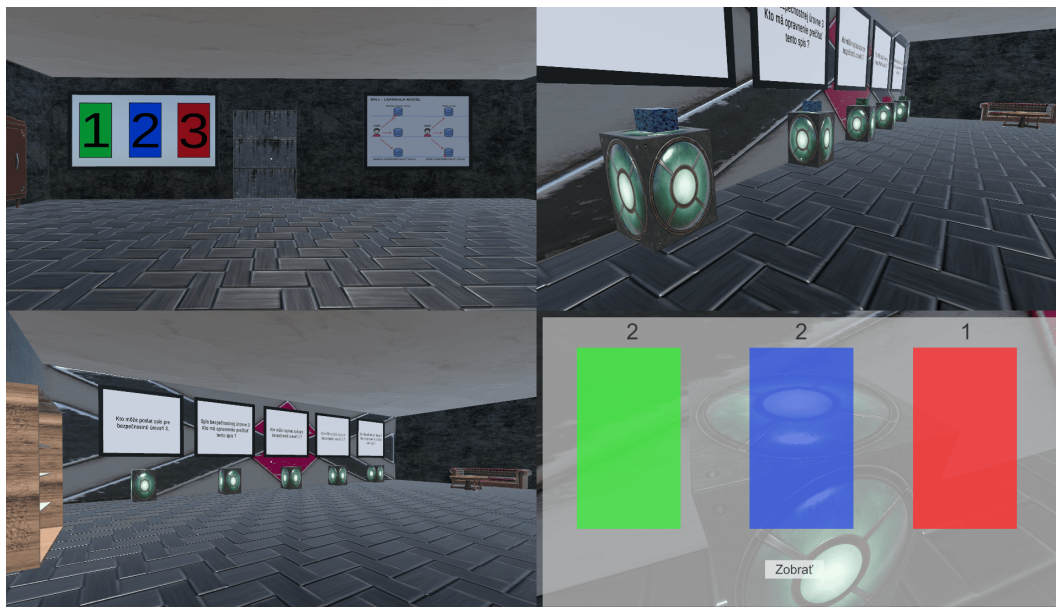
V **tretej miestnosti** sme sa zaoberali bezpečnostným problémom pretečenia vyrovnávacej pamäte. Ide o závažnú chybu, ktorá sa objavuje v programoch napísaných v jazykoch C a C++. Útočník využíva túto chybu na to, aby prepísal obsah pamäťového miesta s inými dátami, ktoré môžu zničiť program alebo umožniť útočníkovi spustiť vlastný kód. Útočník môže tiež získať prístup k súkromným informáciám, ktoré sú uložené v pamäti, ako napríklad heslá alebo iné citlivé dáta.



Obr. 4 Miestnosť 3 - Hands-on

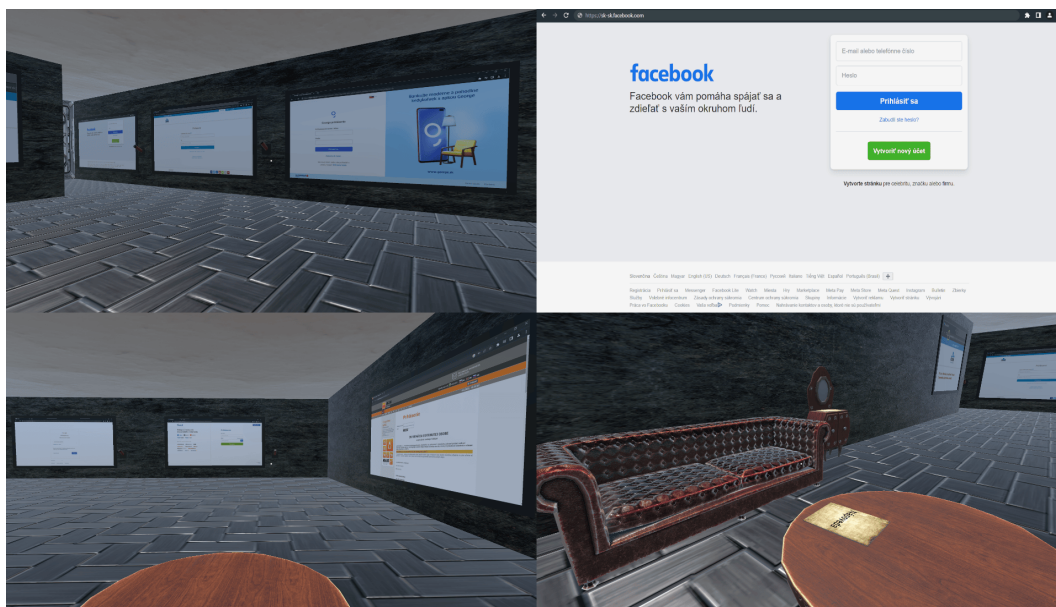
Štvrtá miestnosť sa zameriavala na Bell-LaPadula model. Ide o bezpečnostný model, ktorý slúži na ochranu tajných informácií v počítačových systémoch. Jeho hlavným cieľom je minimalizovať riziko úniku informácií a zabezpečiť, že informácie sú prístupné iba osobám, ktoré k nim majú oprávnenie. Tento model sa často využíva na uchovanie tajných informácií medzi zamestnancami tajného objektu a zabezpečuje, že určité bezpečnostné úrovne vedia zapisovať a čítať informácie len od určitých bezpečnostných úrovní. Každý pracovník má svoju vlastnú bezpečnostnú úroveň, ktorá určuje, k akým informáciám môže mať prístup. V praxi to znamená, že pracovník s vyššou bezpečnostnou úrovňou môže čítať informácie od pracovníka s nižšou bezpečnostnou úrovňou, ale nemôže mu odovzdať informácie jeho bezpečnostnej úrovne. Týmto

spôsobom sa zabezpečuje, aby sa tajné informácie nedostali do nesprávnych rúk.



Obr. 5 Miestnosť 4 - Bell-LaPadula model

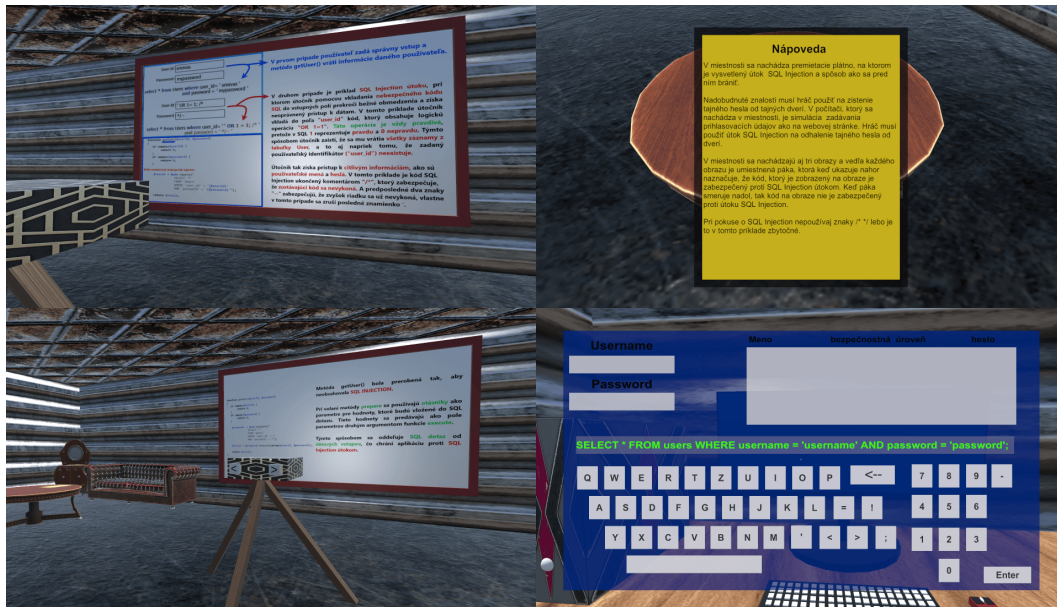
V **piatej miestnosti** sme sa zaoberali riešením závažnej problematiky kybernetických útokov typu phishing, ktoré predstavujú nebezpečenstvo pre bezpečnosť citlivých informácií vlastných obeťami. Tento typ útoku spočíva v získavaní citlivých informácií od obete útočníkom, ktorý sa vydáva za dôveryhodnú osobu alebo organizáciu. Hlavnými prostriedkami realizácie phishingových útokov sú podvodné e-mailové správy, SMS správy a falošné webové stránky, ktoré sú navrhnuté tak, aby vyzerali autenticky s cieľom získať citlivé informácie od obete. Dôsledky phishingových útokov môžu byť veľmi vážne a sprevádzajú sa finančnými stratami, krádežou identity alebo inými formami podvodu. Preto je dôležité, aby sa jednotlivci a organizácie riadne chránili pred týmto druhom kybernetických útokov.



Obr. 6 Miestnosť 5 - Phishing

V **šiestej miestnosti** sme sa venovali bezpečnostnej problematike SQL Injection. Ide o typ útoku na webové aplikácie, ktorý môže byť veľmi nebezpečný pre bezpečnosť a dôvernosť citlivých informácií. Pri tomto type útoku útočník vkladá do prihlasovacích formulárov nebezpečný kód SQL. Tento kód môže spôsobiť neočakávané správanie aplikácie a umožniť útočníkovi získať citlivé informácie, ako sú používateľské mená a heslá alebo dokonca nadobudnúť úplnú kontrolu

nad databázou aplikácie.



Obr. 7 Miestnosť 6 - SQL Injection

V. ZÁVER

Hlavným zámerom našej diplomovej práce bolo navrhnúť a implementovať vzdelávaciu hru, ktorá využíva herné prvky pre podporu výučby bezpečnosti IKT. Štúdiom dostupnej odbornej literatúry a analýza problematiky nám umožnili identifikovať kľúčové oblasti bezpečnosti, ktoré sme potom gamifikovali v hre. Implementácia hry bola realizovaná s využitím programovacieho jazyka C# a vývojového prostredia Unity. Následne bola implementovaná a testovaná na platforme operačného systému Windows. V budúcnosti plánujeme rozšíriť hru aj na platformy Mac a Linux. Hra umožňuje hráčom zábavnou formou rozvíjať svoje schopnosti v oblasti bezpečnosti, čím sa stáva užitočným nástrojom pre učiteľov, ktorí chcú svojim žiakom poskytnúť zaujímavý a interaktívny spôsob výučby. Intuitívne ovládanie a zaujímavé 3D prostredie hry motivujú hráčov dohrať ju do konca, čo v konečnom dôsledku prispieva k zlepšeniu ich vedomostí a zručností v oblasti bezpečnosti IKT. Celkovo sa dá považovať táto hra za úspešný prínos v oblasti gamifikácie výučby bezpečnosti IKT a jej potenciál ako užitočného nástroja pre učiteľov a žiakov je veľký. V budúcnosti sa môže hra rozšíriť o ďalšie miestnosti zamerané na rôzne problematiky rozširovania povedomia o kybernetickej bezpečnosti. Hra má možnosť teleportovať sa do ktorejkoľvek miestnosti prostredníctvom kódu, ktorý môže využívať učiteľ k vysvetleniu výučby v konkrétnych miestnostiach, ktoré si určí.

PODĎAKOVANIE

Článok bol podporený v rámci projektu - 002TUKE-4/2021 Implementácia moderných metód a foriem výučby v oblasti kyberbezpečnosti k požiadavkám praxe.

LITERATÚRA

- [1] M. Rončák, "Gamifikácia ako nástroj marketingovej podpory," Master's thesis, Univerzita Tomáše Bati, Zlín, 2015. [Online]. Available: https://digilib.k.utb.cz/bitstream/handle/10563/34608/ron%C4%8D%C3%A1k_2015_dp.pdf?sequence=1&isAllowed=1
- [2] K. Buzová, "Implementácia gamifikácie do knižníc s využitím sociálnych médií." [Online]. Available: https://itlib.cvtisr.sk/wp-content/uploads/docs/30_inplementacia.pdf
- [3] G. Zichermann, *Gamification by Design: implementing Game Mechanics in Web and Mobile Apps*. Sebastopol: O'Reilly Media, 2011, vol. 300.
- [4] P. Brečka, *Stratégie výučby s podporou IKT v technickom vzdelávaní*. Nitra: PF UKF v Nitre, 2018, vol. 198.
- [5] M. Púpavová, "Gamifikácia a elearning vo vzdelávaní: Čo tieto pojmy znamenajú?" 2023. [Online]. Available: <https://eduworld.sk/cd/martina-pupavova/2170/gamifikacia-a-elearning-vo-vzdelavani-co-tieto-pojmy-znamenaju>
- [6] V. Hnojská, "Herní výukový systém glasslab [online], [cit 2022-19-6]," 2015. [Online]. Available: <https://spomocnik.rvp.cz/clanek/19481/HERNI-VYUKOVY-SYSTEM-GLASSLAB.html?nahled=>
- [7] R. Lukáš, "Co může využití herních principů přinést vzdělávání? [online], [cit 2022-19-6]," Praha, 2015. [Online]. Available: <https://www.eduin.cz/clanky/co-muze-vyuzivani-hernich-principu-prinest-do-vzdelavani/>
- [8] D. Olejár, "Základy kybernetickej a informačnej bezpečnosti [online]," Univerzita Komenského v Bratislave, 2020. [Online]. Available: <https://kinit.sk/wp-content/uploads/2022/06/PistekP-2020-Zaklady-Infomacnej.pdf?fbclid=IwAR1MUwFwHP8IEdfB5OFTWsy9yZhEsSC513bRHsGZPg7FTRIjPjyBZIAM>

- [9] L. Luděk, *Konvergovaná bezpečnost*. Zlín: Radim Bačuvčík - VeRBuM, 2019, vol. 206.
- [10] J. Štofán, "Dopady kybernetických hrozeb na vybranú organizáciu," Master's thesis, Univerzita Tomáše Báti, Zlín, 5 2020.

Gamifikácia v oblasti bezpečnosti IKT

¹Jozef MAJKUT, ²Anton BALÁŽ, ³Emília PIETRIKOVÁ

¹Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach

¹jozef.majkut@student.tuke.sk, ²anton.balaz@tuke.sk, ³emilia.piatrikova@tuke.sk

Abstrakt – Cieľom tejto práce bolo vytvoriť funkčnú počítačovú hru, ktorá bude v sebe obsahovať úlohy z počítačovej bezpečnosti. Začali sme analýzou pojmu gamifikácia, pozreli sme sa nato, aké má princípy a ako sa využíva v praxi. Analyzovali sme existujúce riešenia použitia gamifikácie, spolu s počítačovou bezpečnosťou. Vybrali sme si témy bezpečnosti, ktoré sme následne spracovali do hry. Navrhli a implementovali sme hru v programe Unity. Hru sme podrobili používateľskému testovaniu, vyhodnotili výsledky a na základe toho navrhli vylepšenia.

Kľúčové slová – gamifikácia, Unity, počítačová bezpečnosť, 2D hra

I. ÚVOD

Učenie nie je niekedy zábava. Často sú to hodiny a hodiny krvi a potu. Najhoršie je, ak nás vec, ktorú sa chceme naučiť vôbec nebaví. Čo ak by sa dalo nudné učivo spraviť zábavnejším? Stále sa učíme spôsobom, že si prečítame knihu alebo poznámky z kníh. V niektorých prípadoch sa však hodí aj nejaká vizuálna pomôcka alebo reálne vyskúšanie. V bezpečnosti je mnoho tém, ktoré je ťažké sa naučiť iba z čítania. Niektoré témy je užitočné si aj reálne vyskúšať.

Hry sú momentálne veľmi rozšírené. Čo ak by sa dalo spojiť nudné alebo náročné učivo s hrou, ktorá nás bude baviť? To by predsa znamenalo, že nás bude baviť aj učenie. Naším cieľom bude využiť gamifikačné prvky hier pre vytvorenie vlastnej hry.

Analyzujeme témy počítačovej bezpečnosti a gamifikujeme ich. Výsledkom bude hra, ktorá na prvý pohľad vyzerá ako obyčajná hra. Rozdiel bude v tom, že dokončenie hry sa nezaobíde bez naučenia sa tém, ktoré sú v úlohách riešené. V ideálnom prípade, po dokončení hry bude mať hráč viac vedomostí o bezpečnosti ako pred hraním hry.

II. GAMIFIKÁCIA A JEJ PRINCÍPY

Gamifikácia je definovaná ako použitie dizajnových elementov z hier v kontexte inom ako hernom. Tieto elementy sa dajú zovšeobecniť, aby boli využiteľné pre široké spektrum oblastí [1]. Môže sa použiť ako nástroj na zlepšenie spolupráce a motivácie ľudí, ktorí vykonávajú úlohy, ktoré nie sú veľmi atraktívne. Znamená to teda, že neatraktívne úlohy sa dajú podať takým spôsobom, ktorý ich zmení na viac atraktívne [2]. Použitím týchto princípov môže nastať k zvýšeniu atraktivity riešenej úlohy. Pri vytváraní našej hry použijeme princípy gamifikácie nato, aby sme zatriktívniili témy bezpečnosti.

Gamifikácia nie je o tom pretvoriť každodenné aktivity na hru, ale vytvoriť aktivity pomocou herných mechaník, pre zábavnejší a záživnejší zážitok. To sa môže týkať rôznych aktivít z rôznych oblastí.

A. Stavbné časti gamifikácie

Hry majú svoje špecifické časti, ktoré hrajú významnú rolu pri gamifikácii [3]:

- používatelia - zamestnanci, zákazníci, študenti,
- výzvy a úlohy, ktoré používateľ musí splniť, aby dosiahol cieľ,
- body, ktoré používateľ zbiera po splnení úlohy,
- úrovne, ktoré postupne prechádza a dostáva body,
- medaily za splnenie úlohy,
- hodnotenie hráčov na základe ich dosiahnutých bodov.

B. Tvorba gamifikovaného systému

Hlavné kroky k vytvoreniu gamifikovaného systému [3]:

- **determinácia charakteristík používateľa** - ak chce učiteľ implementovať nové učebné metódy. Je nevyhnutné definovať profil používateľa, aby sa zistilo, či použité techniky budú fungovať [3]. Základom je určiť, aké schopnosti musia používatelia mať, aby splnili úlohy. Ak budú úlohy príliš jednoduché alebo príliš ťažké, vyvolá to demotiváciu u používateľov [4],
- **definícia cieľov** - ciele musia byť presne a jasne definované. Účel vzdelávania je dosiahnuť ciele, pretože inak by bola snaha zbytočná. Ciele určujú aké učivo sa zaradí do procesu a aké aktivity treba zahrnúť. Podľa toho sa vyberú aj príslušné herné mechaniky a techniky [3],
- **vytvorenie vzdelávacieho obsahu a aktivít gamifikácie** - obsah by mal byť interaktívny, záživný a bohatý na multimediálne elementy. Mali by byť vytvorené tréningové aktivity za účelom dosiahnutia stanovených cieľov [5], Aktivity by mali dovoliť:
 - **viac pokusov** - aktivity musia byť vytvorené tak, aby ich mohol študent opakovať v prípade neúspešného pokusu. Je dôležité vytvoriť podmienky a príležitosti na dosiahnutie stanoveného cieľa. Opakovaním si študent zlepšuje svoje schopnosti [3],
 - **realizovateľnosť** - aktivity majú byť splniteľné, musia byť prispôbené znalostiam a schopnostiam študenta [3],
 - **zvýšenie náročnosti** - každá nasledujúca úloha je zložitejšia ako predošlá. Vyžaduje teda viac úsilia od študenta primerane ich novozískaným znalostiam [3],
 - **viac ciest** - nato, aby sa študenti naučili rôzne znalosti, musí im byť umožnené dopracovať sa k výsledku rôznymi cestami. To im umožní vytvoriť si vlastné stratégie, čo je jednou z kľúčových charakteristík v aktívnom učení sa [3].
- **pridanie herných elementov a mechaník** - kľúčovým elementom gamifikácie je zaradenie úloh, ktoré má študent vykonať. Splnenie úlohy vedie k zbieraniu bodov, posúvanie sa do ďalších úrovní a vyhrávanie cien. Všetky tieto akcie sú zamerané na splnenie dopredu vytvoreného cieľa. Aké elementy budú zaradené do tréningu závisí od určených cieľov. Aktivity, ktoré vyžadujú samostatnú prácu odmeňujú študenta individuálnymi odmenami, napríklad medaily. Aktivity, ktoré vyžadujú spoluprácu s ostatnými študentmi, zaradia študenta do komunity študentov. Ich výsledky sú zverejnené a viditeľné v tabuľke najlepších riešiteľov [4].

III. VYBRANÉ TÉMY BEZPEČNOSTI

V tejto časti si popíšeme, ktoré témy bezpečnosti spracujeme do hry. Krátko si definujeme každú tému.

A. Symetrické šifrovanie

V symetrickom šifrovaní sa na šifrovanie a dešifrovanie používa iba jeden kľúč. Táto metóda spotrebuje značnú časť výpočtových zdrojov, ako CPU a operačná pamäť. Poznáme Caesarovu šifru, AES, DES, 3DES, RC6 a iné. Pre vysvetlenie symetrického šifrovania bohato postačí Caesarova šifra, ktorú spracujeme v hre. Hráč dostane slovo a kľúč, ktorým bude šifrovať alebo dešifrovať.

1) *Asymetrické šifrovanie*: V asymetrickej šifre sa používajú dva kľúče, a to súkromný a verejný. Verejný kľúč sa používa na šifrovanie a súkromný kľúč sa používa na dešifrovanie, napríklad RSA šifra a ECC. Šifrovanie je založené na matematických funkciách, je náročné na výpočty a nie je príliš efektívne. V hre implementujeme jednoduché počítanie kľúča RSA šifry.

2) *Digitálny podpis*: Digitálny podpis je matematická schéma, ktorá zabezpečuje súkromie konverzácie, integritu údajov a pravosť digitálnej správy odosielateľa. V hre vytvoríme digitálny podpis pomocou hašovacej funkcie md5.

3) *Malvér*: Malvér je škodlivý softvér, ktorý útočí na iný softvér. Všeobecne škodlivo znamená, že sa správa inak ako sa prezentuje. V hre bude zjednodušene ukázaný útok, ktorý sa spustí kliknutím na odkaz v podvodnom maile.

4) *Ransomware*: Ransomware je typ malvéru, ktorý šifruje súbory obete a následne požaduje platbu výmenou za kľúč, ktorý dokáže súbory dešifrovať. Keď je Ransomware prvýkrát nainštalovaný na počítači obete, zvyčajne sa zameria na citlivé súbory, ako napríklad dôležité finančné údaje, obchodné záznamy, databázy, osobné súbory a ďalšie. Osobné súbory, ako sú fotografie a domáce filmy, môžu mať sentimentálnu hodnotu pre obeť. V hre implementujeme, podobne ako pri malvéri, podvodný email a odkaz, ktorý stiahne vírus a bude požadovať výkupné.

5) *Phishing*: Phishing je popisovaný ako pokus oklamať používateľa, aby odhalil citlivé informácie, ako sú bankové údaje a čísla kreditných kariet, odoslaním škodlivých odkazov používateľovi, ktoré vedú k falošnému zriadeniu webu. Hráč dostane podvodný email od banky s odkazom na zmenu PIN kódu. Po zmene bude účet zneužitý a budú ukradnuté peniaze. Tieto tri typy malvérov ukážeme v hre. Hráč bude musieť tieto škodlivé softvéry označiť ako škodlivé, podľa vymysleného správania v hre.

6) *Silné heslo*: Silné heslo pozostáva z dĺžky hesla aspoň 8 znakov, čísel, veľkých písmen a špeciálnych znakov. Hráč si v určitej časti bude musieť takéto heslo vymyslieť a zapamätať.

7) *Multifaktorová autorizácia*: Multifaktorová autentifikácia poskytuje vyššiu úroveň bezpečnosti ako jednofaktorová. Je založená napríklad na biometrii alebo formou jednorázového kódu. V hre použijeme metódu jednorázového kódu, ktorý je zaslaný na mobilné zariadenie.

IV. NÁVRH HRY

Ak chceme vytvoriť hru, ktorá má hráča oboznámiť alebo naučiť istú vec z počítačovej bezpečnosti, je potrebné si určiť, ktoré témy bezpečnosti do hry zakomponovať. Z predošlých riešení a literatúry sme vybrali kryptografiu a internetovú bezpečnosť. Z kryptografie sme vybrali symetrické, asymetrické šifrovanie a hašovanie. Z internetovej bezpečnosti sme vybrali Malware, Ransomware, Phishing, multifaktorovú autorizáciu a silné heslo.

S týmito témami sa bude hráč oboznamovať postupne počas hrania. Bude sa pohybovať v hernom svete až narazí na úlohu, ktorú musí splniť, aby sa dostal ďalej. Na to, aby hráč splnil úlohu, musí ovládať tému, na ktorú je úloha zameraná. Hra bude mať aj jednoduchý príbeh, ktorý bude navádzať hráča k ďalším krokom.

V hre vytvoríme úlohy, ktoré budú zamerané na určitú tému. Tieto úlohy budú zakomponované do herného sveta tak, aby ich hráč plnil postupne ako sa dostáva ďalej v hre. V hernom svete budú fyzické prekážky, cez ktoré sa nedá prejsť. Tieto prekážky budú odstránené splnením úlohy hráčom. Tento účel budú plniť zatvorené dvere v miestnosti. Za splnenú úlohu hráč dostane body, ktoré sa mu budú počítať do celkového hodnotenia pri skončení úrovne. Bude meraný aj čas, za ktorý bude úroveň dokončená a započítava sa do celkového bodovania.

Úlohy začnú témou symetrického a asymetrického šifrovania. Hráč sa stretne s pár úlohami na túto tému, aby si ju lepšie zapamätal. Následne sa zoznámia so silným heslom a multifaktorovou autorizáciou, potom sa stretne s hašovaním a digitálnym podpisom pomocou hašovania a nakoniec mu budú prezentované počítačové vírusy rôzneho typu.

A. Použitie gamifikácie

Už sme popisovali postup gamifikácie a teraz hlavné kroky gamifikácie aplikujeme pri návrhu hry. Poslúži nám to ako vzor, ktorého sa budeme držať pri tvorbe úloh a úrovní.

1) *Determinácia charakteristík používateľa*: Používateľmi našej hry budú aj ľudia, ktorí nemajú s počítačovou bezpečnosťou veľké skúsenosti, preto do hry spracujeme základné témy bezpečnosti. Bude pre nich výhodou, ak už niekedy hrali počítačové hry, ale nič sa nestane ak hry nikdy nehrali.

Hra nebude mať náročné ovládanie natoľko, aby sa nedala zvládnuť. Naším cieľom je prilákať ľudí, ktorí sa chcú dozvedieť viac o šifrovaní a bezpečnosti na internete.

Musíme dodať, že hra nebude vhodná pre zrakovovo alebo sluchovo znevýhodnených, pretože bude využívať audiovizuálne prvky, ktorých vnímanie je dôležité pre zážitok z hry.

2) *Definícia cieľov*: Cieľom je vytvoriť úlohy na nami vybrané témy. Úlohy budú väčšinou vytvorené ako praktické použitie princípov šifrovania a bezpečnosti na internete.

Hráč si zjednodušene prakticky vyskúša šifrovať a dešifrovať a vyskúša si aj symetrické a asymetrické šifrovanie. Prakticky uvidí ako by prebiehalo obdržanie podvodného emailu a následné napadnutie vírusom. Reálne si vytvorí heslo, ktoré spĺňa podmienky pre silné heslo. Zjednodušene si vyskúša ako prebieha multifaktorová autorizácia.

3) *Vytvorenie vzdelávacieho obsahu a aktivít*: Vytvoríme viac úloh na jednu tému s postupne zvyšujúcou sa náročnosťou. Povinné úlohy budú mať neobmedzený počet pokusov na splnenie. Bude to tak preto, lebo v týchto úlohách sa hráč s témou stretne často po prvýkrát, preto mu chceme dať priestor na chyby.

Bonusové úlohy by mal hráč plniť až keď už trochu tému ovláda, preto má iba jeden pokus. Po zadaní odpovede dostane informáciu, či zadal odpoveď správne alebo nesprávne. Všetky úlohy budú splniteľné, dokonca aj tie náročnejšie, avšak zaberú možno viac času. Hráč bude mať na výber, v akom poradí chce niektoré úlohy splniť alebo akou cestou sa vydá. Žiadne z jeho rozhodnutí nezapríčiní, že hra sa nebude dať dohrať. Dostane sa do konca nech si zvolí akúkoľvek cestu.

4) *Pridanie herných elementov a mechaník:* V hre implementujeme herné mechaniky ako bodovací systém. Hráč zbiera body za plnenie úloh a zbieranie mincí. Celkové bodové hodnotenie záleží aj od času, za ktorý zvládne úroveň. Povedzme, že hráč nazbiera 200 bodov a úroveň ukončí. Splnenie úrovne mu trvalo 400 sekúnd. Časový limit pre úroveň je 1000 sekúnd. Odpočítame čas od limitu a pripočítame body, vtedy vznikne celkové hodnotenie. V hre vytvoríme prvky, ktoré budú premiestňovať hráča na iné miesto, v ktorom sa nachádza. Pre herný zážitok nebude príjemné, ak sa hráč potrebuje dostať na iné miesto v hernom svete, ktoré je vzdialené a je potrebné sa dlho presúvať. Na to poslúžia portály, ktoré hráč buď použije alebo sa spustia samé po dokončení úlohy.

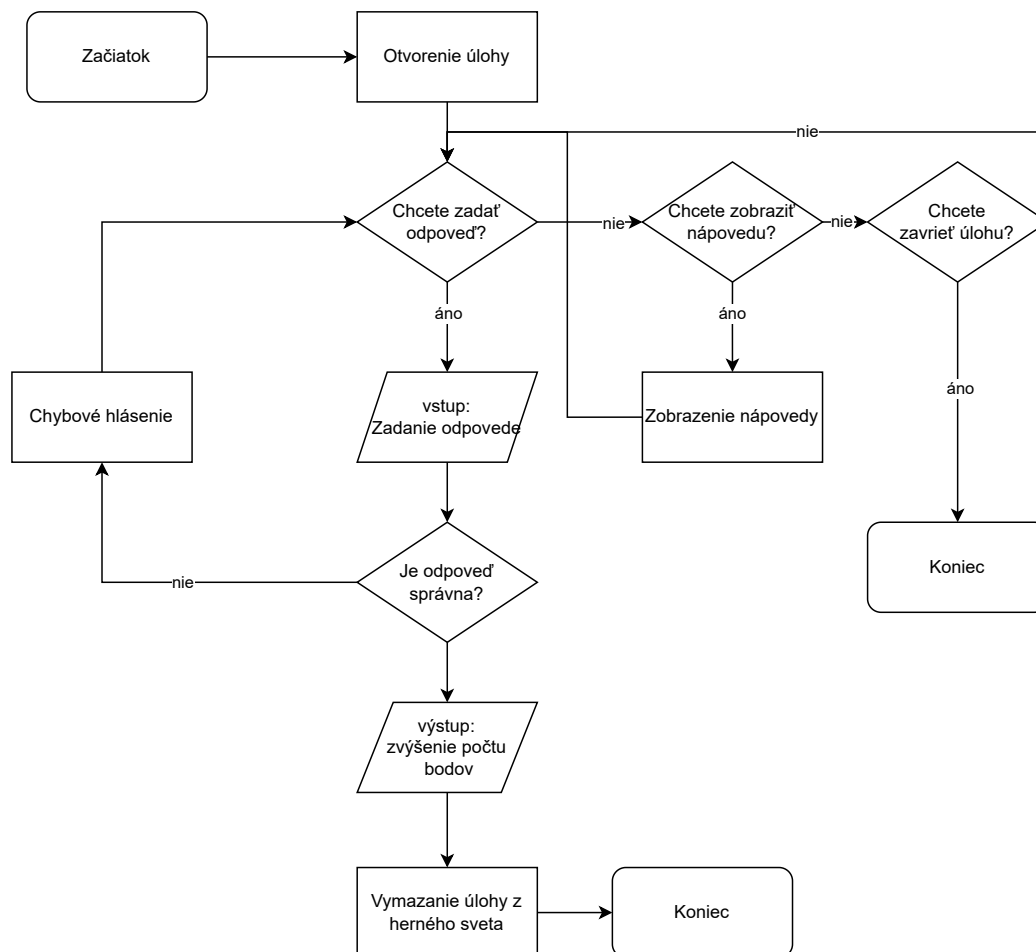
Ďalšou mechanikou budú záchytné body. Hráč napríklad spadne mimo mapu alebo príde do kontaktu s objektom, ktorý ho napríklad zabije. Vtedy sa spustí mechanika záchytného bodu a vráti hráča do pozície v hernom svete, kde už predtým bol a môže skúsiť prejsť pasáž zнова.

B. Scenár vyriešenia úlohy

Hráč sa bude pohybovať po hernom svete a narazí na úlohu. Zobrazí sa mu istá inštrukcia, že má stlačiť tlačidlo, aby úlohu zobrazil. Ak tak urobí, zobrazí sa mu znenie úlohy, ktoré si prečíta. Ak nevie hneď zadať odpoveď, môže si zobraziť nápovedu.

Po zadaní odpovede, časť kódu prislúchajúca úlohe vyhodnotí správnosť a upraví bodový stav. Ak je odpoveď nesprávna, zobrazí sa chybová hláška. Hráč tak môže zadať odpoveď zнова až kým nebude správna.

Tento prípad sme si navrhli aj pomocou stavového diagramu, ktorý je na obrázku 1. Hráč začne



Obr. 1 Proces vyriešenia úlohy

otvorením úlohy. Tu sa môže rozhodnúť, či si zobrazí nápovedu, zadá odpoveď alebo zavrie úlohu.

Ak si vyberie zobrazenie nápovedy, tak sa mu zobrazí nápoveda. Môže sa vrátiť späť a tam skúsiť zadať odpoveď alebo ukončiť úlohu. Ak si zvolí ukončenie úlohy, tak sa úloha zavrie, ale ostane v hernom svete a môže sa k nej vrátiť. Ak si zvolí zadanie odpovede, na vstup zadá odpoveď, ktorej správnosť sa overí systémom. Keď systém vyhodnotí odpoveď ako nesprávnu, zobrazí sa chybové hlásenie a odpoveď sa dá zadať zнова.

Nebude to platiť pri bonusových úlohách. V bonusovej úlohe sa pri zadaní zlej odpovede úloha uzavrie, zmizne z herného sveta a nebude už prístupná. Bude znova dostupná pri opakovaní úrovne.

Ak systém vyhodnotí úlohu ako správnu, pridá sa určený počet bodov hráčovi a úloha sa vymaže z herného sveta. V niektorých prípadoch sa stanú aj udalosti navyiac ako:

- otvorenie dverí a sprístupnenie tak novej oblasti v hre,
- zmiznutie prekážky, ktorá bránila v ďalšom postupe,
- presunutie hráča na iné miesto,
- otvorenie ďalšej úlohy,
- ukončenie úrovne, pretože bola splnená posledná úloha.

V. IMPLEMENTÁCIA HRY

Hru budeme tvoriť na spôsob 2D platformovej hry. Hráč sa pohybuje po hernom svete, prekonáva prekážky a interaguje s objektami. Použijeme na to herný engine Unity. Základné stavebné prvky tvorby hier v Unity, ktorá sme použili sú:

- scény, do ktorých sa umiestňujú jednotlivé komponenty,
- kamera, ktorá je zodpovedná za to, čo budeme v hre vidieť,
- objekty, ktoré sa umiestňujú do scény a tvoria tak herný svet,
- zvuky, ktoré priradzujeme rôznym udalostiam v diani hry,
- elementy používateľského rozhrania, napríklad tlačidlá, textové polia a iné.

V hre budeme mať tri scény:

- scéna menu, ktorá obsahuje všetky objekty herného menu,
- scéna úroveň 1, ktorá obsahuje všetky objekty prvej úrovne,
- scéna úroveň 2, ktorá rovnako obsahuje všetky objekty druhej úrovne.

V každej scéne bude bežať vlastná funkcionálna, ktorú si stručne popíšeme.

A. Scéna herného menu

V scéne sa nachádza kamera. Kamera určuje, čo bude na obrazovke vidieť po spustení hry. Má niekoľko vlastností, ale dôležité je nastaviť jej veľkosť a pozíciu. Pre zobrazenie menu používame plátno a kameru umiestnime tak, aby zaberala celé plátno. Plátno má jednu dôležitú vlastnosť, a to je veľkosť. Veľkosť plátna sa zvyčajne nastaví rovnako ako je rozlíšenie obrazovky počítača. Na toto plátno sa potom umiestňujú prvky používateľského rozhrania.

Pri tvorbe menu sme použili plátno, do ktorého sme vložili prázdny obrázok a nastavili jeho veľkosť na veľkosť plátna. Nový obrázok v Unity je vždy biely, čiže prázdny. Je potrebné mu nastaviť zdroj a v našom prípade sme mu nastavili zdroj pozadia hry.

Na plátno sme umiestnili textový element a jeho hodnotu zmenili na názov hry. Ďalej sme pridali tri tlačidlá, ktorých textovú časť sme zmenili tak, aby tvorili tieto tri tlačidlá:

- hrať, ktoré po kliknutí zobrazí úrovne na výber, maximálne dosiahnuté skóre a typ medaily získanej za dokončenie úrovne,
- návod, ktoré v prvom snímku popíše princíp hry a v druhom snímku popíše ovládanie, úlohy a mince,
- zavrieť, ktoré jednoducho slúži na vypnutie hry.

B. Scéna úroveň

Scény úroveň1 a úroveň2 sú si podobné, preto ich tvorbu popíšeme spoločne. Scény fungujú rovnako, len majú zmenené objekty, pretože sa v scéne odohráva iný dej hry.

Zmena oproti scéne s menu je, že kamera nebude statická. Implementovali sme, aby kamera sledovala hráča. V kóde zistíme pozíciu hráča v hernom svete a nastavíme rovnakú pozíciu pre kameru. To sa deje v cykle, aby sme mali vždy aktuálnu pozíciu hráča.

Mince sú vytvorené ako objekty, ktoré sme si nakreslili a vložili do scény. Objekty majú svoje komponenty. Komponentov je veľa, ale spomenieme len tie, ktoré sú pre túto funkcionálnu dôležité. Pre fungovanie mince je potrebná trieda s funkcionálnou prídavania bodov, kolízny komponent a animácia. Trieda sa vytvorí a pridá ako komponent k minci. Obsahuje metódu, ktorá k aktuálnym bodom pripočíta hodnotu.

Kolízny komponent s touto triedou spolupracuje tak, že deteguje kedy príde minca do kontaktu s hráčom a vtedy zavolá metódu na pridanie bodov.

Animácia beží v cykle a transformuje pozíciu mince vertikálne z jedného bodu do druhého a späť.

1) *Úlohy*: V hre sú jasne vyznačené úlohy, za ktoré hráč získava body. Úlohy nasledujú za sebou a niektoré na seba nadväzujú. Narazí na 2 typy úloh:

- povinné úlohy, ktoré hráč musí splniť, aby pokračoval v hre,
- bonusové úlohy, ktoré slúžia iba na získanie vyššieho počtu celkových bodov.

Značky v hre slúžia ako tlačidlá, ktoré aktivujú konkrétny panel s úlohou. Majú svoju animáciu vertikálneho pohybu rovnako ako mince. Animácia naznačuje hráčovi, že môže s tlačidlom interagovať. Majú aj kolízny komponent pre detekciu či je hráč pri úlohe, aby ju mohol otvoriť.

Povinné aj bonusové úlohy sú bodovo hodnotené a prispievajú do celkového bodového zisku. Ak hráč otvorí úlohu, zobrazí sa mu panel, ktorý už predstavuje samotnú úlohu a požaduje jej vyriešenie. Pri úlohách sú nápovedy, ktoré vysvetľujú tému konkrétnej úlohy a pomôžu tak s jej splnením. Zobrazia sa kliknutím na tlačidlo, ktoré znova len aktivuje ďalší panel.

Povinné úlohy môže hráč opakovane skúšať až kým odpoveď nebude správna. Bonusové úlohy majú iba jeden pokus. Pri prvej zlej zadanej odpovedi sa zobrazí chybová hláška, deaktivuje sa tlačidlo pre zadanie odpovede a úloha sa už dá iba zavrieť a zmizne zo sveta. Pri zadaní správnej odpovede sa zobrazí hláška, ktorá oznámi, že odpoveď je správna.

Úlohy sú implementované nasledovne. Úlohu, ako sme už spomínali, tvorí panel. V tomto prípade má v sebe aj pole pre zadávanie vstupu od používateľa. Pre každú úlohu sa vytvorí vlastná trieda s vlastnou funkcionalitou, metódami a referenciami herného sveta. Trieda sa prideli značkám pre úlohy. Konkrétna značka bude mať v sebe funkcionalitu jednej úlohy.

V triede potrebujeme priradiť referencie na vstupné pole, panel a tlačidlá. Ak hráč stojí na značke úlohy, môže úlohu otvoriť, pretože kolízny komponent zaznamená, že hráč stojí na značke. Panel s úlohou sa zavolanou metódou zmení na aktívny. Hráč si prečíta úlohu, a keď zadá odpoveď do vstupného textového poľa, klikne na tlačidlo, ktoré zavolá metódu pre vyhodnotenie odpovede.

V triede, ktorá je v úlohe pridelená sa získa hodnota textu, ktorý hráč zadal a porovná sa s textom, ktorý je v kóde zadaný ako správny. Ak je odpoveď správna, spustí sa akcia špecifická pre úlohu. Takýchto akcií máme v hre niekoľko, napríklad:

- otvorenie dverí do miestnosti,
- zobrazenie textu, ktorý oznámi, že odpoveď je správna,
- presunutie hráča na iné miesto.

Otvorenie dverí je implementované ako spustenie animácie, ktorá transformuje pozíciu dverí tak, aby už nebránili v ceste. Zobrazenie textu funguje na už známom princípe aktivovania elementu, a tým sa text objaví na obrazovke. Presunutie hráča funguje na princípe transformácie pozície postavy v hernom svete.

Hráč neplní iba úlohy, ale na to, aby sa dostal ďalej v hre, musí prekonať pripravené prekážky ovládaním hlavnej postavy, napríklad v prípade platformovej hry to môže byť preskakovanie z platformy na platformu. V našom riešení sme do hry implementovali časť, ktorú musí hráč s postavou preskakať. Podobne sme vytvorili aj časť, kde sa hráč vyhýba kamerám a kopijám, ktoré vystupujú zo zeme. Tieto časti viac popisujeme v kapitole.

C. Stav celej implementácie

Vytvorili sme hru pre operačný systém Windows 10. Vývojové prostredie Unity dovoľuje hru spracovať aj pre iné operačné systémy, herné konzoly a mobily. Pre nás stačí jeden, pretože hra sa nezmení, ak ju vytvoríme napríklad pre mobil.

V hre sú dva úrovne. Hráč začne v hernom menu, kde je návod so stručným opisom hry a jej ovládania. Hrať sa môže akákoľvek úroveň chce, ale doporučame začať prvou.

D. Prvá úroveň

V prvej úrovni začne pred budovou, v ktorej pracuje. Tu sa hneď nachádza prvá bonusová úloha o Caesarovej šifre. Do budovy vstúpi splnením úlohy na dešifrovanie Caesarovej šifry, ktorá mu otvorí dvere. Pokračuje k pracovnému stolu, kde musí nastaviť nové heslo pre dvere do budovy. Použije na to šifrovanie Caesarovou šifrou.

Hráč postupuje ďalej do časti, kde sa musí preskakať k počítaču, ktorý hlási chybu. Počas skákania zbiera mince a narazí na dve bonusové úlohy, ktoré sú o symetrickom a asymetrickom šifrovaní. Počítač má v sebe úlohu na výpočet verejného kľúča RSA šifry. Splnením tejto úlohy otvorí dvere do serverovne. V serverovni sú 4 servery, ktoré potrebujú nové pamäťové disky. Disky sa nachádzajú v sklade a majú pri sebe ďalšie úlohy na výpočet verejného kľúča RSA šifry. Tento vypočítaný kľúč je potrebné zadať pri serveri, a tým sa server opraví.

Po opravení všetkých štyroch serverov hráč použije portál do kancelárie, kde pre neho čakajú ďalšie úlohy.

Pri osobnom počítači je úloha pre nastavenie silného hesla. Hráč má k dispozícii textové pole, do ktorého zadáva heslo.

Ak je heslo nastavené, presunieme sa ďalej v miestnosti, kde sa nachádza mobilný telefón. V mobile sa prihlásime a zapneme multifaktorovú autorizáciu. Presunieme sa späť k počítaču. Zadáme znova heslo a zobrazí sa nám mobil s ikonou prijatej správy. Správu otvoríme a zobrazí sa štvormiestny kód, ktorý zadáme do textového poľa a dokončíme tak prihlásenie.

Zobrazí sa mu päť textových súborov. Každý súbor ma na svojom konci podpis tvorený hašovacou funkciou md5. Hráč má k dispozícii tlačidlo, ktoré zo súboru vytvorí nový haš. Hráč porovná novovytvorený haš s tým, ktorý je priložený pri súbore. Rozhodne tak, či bol súbor narušený alebo je v poriadku. Tu úroveň končí a zobrazí sa koncová obrazovka. Na obrazovke je zobrazený počet nazbieraných bodov, čas hrania, celkový počet bodov, najvyššie dosiahnuté skóre a typ medaily podľa počtu bodov. V zobrazenom okne má tri tlačidlá:

- menu,
- ďalšia úroveň,
- hrať znova.

Táto úroveň má v sebe 12 povinných a 5 bonusových úloh s možným bodovým ziskom 170 bodov. Obsahuje aj 28 mincí, ktoré sú každá za tri body. Dokopy je možné získať 84 bodov za mince.

E. Druhá úroveň

Druhá úroveň začína pred domom hlavnej postavy hry. Náchádzame sa pred bránou, ktorá je zabezpečená heslom. Hráč tu má dve možnosti:

- zadať heslo, ktoré je zašifrovanie slova kľúčom o veľkosti 20,
- prejsť záhradou, v ktorej sú kamery a kopije, cez ktoré musí preskákať.

Ak správne uhádne heslo, bude odmenený 50-timi bodmi a dostane sa domov. Ak zvolí prechod záhradou, čakajú na neho štyri povinné a tri bonusové úlohy. Úlohy sú formou kvízu. Na výber sú štyri možnosti a iba jedna je správna. Hráč si tak postupne otvára brány v záhrade a dostáva sa bližšie k domu.

V záhrade sú pre neho pripravené aj iné prekážky ako úlohy. Celá záhrada je pokrytá bezpečnostnými kamerami, ktoré majú svoje zorné pole. Ak sa hráč dostane do zorného poľa kamery, hra ho vráti naspäť na začiatok. Ďalej má pred sebou kopije, ktoré sa vysúvajú z podlahy a stropu. Ak s nimi príde do kontaktu, bude vrátený do polovice záhrady. Do polovice preto, lebo prvé testovania hry ukázali túto pasáž ako veľmi náročnú a frustrujúcu. Na konci sa dostane do domu rovnako ako pri prvej možnosti.

V dome zamieri hráč hneď k posteli. Postava zaspí a ocitne sa v sne. V sne sa objaví počítač, ktorý potrebuje pomoc. Potrebné je nájsť 6 programov, preklikať sa obrázkami z ich prostredia a určiť, či je program bezpečný alebo nebezpečný. Po každom určení sa zobrazí okno s vysvetlením o aký program išlo a prečo je bezpečný alebo nebezpečný.

Splnením všetkých úloh sa dostávame von zo sna. Tu úroveň končí a zobrazí sa rovnaká obrazovka ako pri prvej úrovni.

Táto úroveň obsahuje 12 povinných úloh a 5 bonusových úloh s možným bodovým ziskom 170 bodov. Úroveň obsahuje 10 mincí, ktoré majú dokopy hodnotu 30 bodov.

VI. ZÁVER

Na základe existujúcich riešení sme navrhli koncept našej hry. Určili sme si konkrétne témy počítačovej bezpečnosti, ktoré sme spracovali do úloh pre hru. Navrhli sme hru, jej príbeh, hrateľnosť, bodovací systém a grafické prevedenie.

V nami vybranom hernom engine sme tento návrh implementovali. Vytvorili sme tri scény ako menu, úroveň1 a úroveň2. Hráč sa stretne s 24 povinnými úlohami a 10 bonusovými úlohami. V úlohách sú obsiahnuté témy ako symetrické šifrovanie, asymetrické šifrovanie, silné heslo, multifaktorová autorizácia, malvér, ransomware, phishing a antivírus.

Hráč sa pohybuje v hernom svete s postavou. Počas hrania rieši úlohy a zbiera mince, aby získaval body. Stretne sa s prekážkami ako skákacia časť v prvej úrovni a prechod cez záhradu, ktorá je strážená kamerovým systémom a kopijami, ktoré vystupujú zo zeme a stropu.

Našu hru sme podrobili používateľskému testovaniu a výsledky sme popísali vo vyhodnocovacej časti. Hráči pozitívne hodnotili spracovanie tém do úloh. Pre všetkých hráčov bola hra zaujímavá a vďaka nej sa niečo naučili. Všetci hráči by túto hru doporučili ľuďom, ktorí sa chcú s riešenými témami zoznámiť iným spôsobom ako čítaním literatúry. Hráči ohodnotili hru ako stredne ťažkú, čo bolo aj naším cieľom. Hru sme chceli vyvážiť, aby nebola príliš ľahká, ale ani príliš ťažká.

Spätnú väzbu, ktorú nám hráči zanechali sme vyhodnotili a navrhli sme vylepšenia. Popísali sme vylepšenie používateľského rozhrania v jednom konkrétnom prípade.

Hru je možné rozšíriť o ďalšie úrovne, ktoré môžu byť zamerané na ďalšie ešte nespracované témy bezpečnosti. Kód a herné objekty sa dajú znova použiť. Zaberie to menej práce ako v našom prípade, keď sme všetko vytvárali od nuly.

Mohlo by byť zaujímavé pridať prvok súťaživosti, ktorý je jeden zo základných princípov gamifikácie a v tejto hre nie je implementovaný ideálne. Týka sa to skóre, ktoré by sa mohlo zdieľať pomocou pripojenej databázy. Hráč by tak nevidel iba svoje najlepšie dosiahnuté skóre, ale aj skóre iných hráčov, ktorí by svoje výsledky nahrali do databázy. Zvýšilo by to motiváciu získať lepšie skóre, ak by sa hráči porovnávali medzi sebou.

PODĎAKOVANIE

Článok bol podporený v rámci projektu - 002TUKE-4/2021 Implementácia moderných metód a foriem výučby v oblasti kyberbezpečnosti k požiadavkám praxe.

LITERATÚRA

- [1] T. I.-C. Gaming, "Chi 2011 workshop gamification: Using game design elements in non-game contexts."
- [2] A. F. Aparicio, F. L. G. Vela, J. L. G. Sánchez, and J. L. I. Montes, "Analysis and application of gamification," in *Proceedings of the 13th International Conference on Interacción Persona-Ordenador*, 2012, pp. 1–2.
- [3] G. Kiryakova, N. Angelova, and L. Yordanova, "Gamification in education." *Proceedings of 9th International Balkan Education and Science Conference*, 2014.
- [4] W. H.-Y. Huang and D. Soman, "Gamification of education," *Report Series: Behavioural Economics in Action*, vol. 29, pp. 11–12, 2013.
- [5] J. Simoes, R. D. Redondo, and A. F. Vilas, "A social gamification framework for a k-6 learning platform," *Computers in Human Behavior*, vol. 29, no. 2, pp. 345–353, 2013.

Vzdialená správa hardvéru - IP KVM

¹Matej VYSOKAI, ²Anton BALÁŽ

¹Katedra elektroniky a multimediálnych telekomunikácií, FEI TU v Košiciach

²Katedra počítačov a informatiky, FEI TU v Košiciach

¹matej.vysokai@student.tuke.sk, ²anton.balaz@tuke.sk

Abstrakt – Vzdialená správa zariadenia je rozšírený úkon, ktorý je vykonávaný IT administrátormi vo firmách, ale aj používateľmi v domácnostiach a preto je dôležité poznať jej možnosti a spôsoby implementácie. Cieľom tejto práce je popísať problematiku vzdialenej správy zariadení rôznych operačných systémov, ktoré nie je možné spravovať pomocou softvéru na vzdialenú správu. Taktiež táto práca obsahuje analýzu problematiky vzdialenej správy, jej výhody ako aj nevýhody a bezpečnostné riziká. Ďalej sa venuje návrhu a implementácii špecializovaného hardwaru na báze OS Linux.

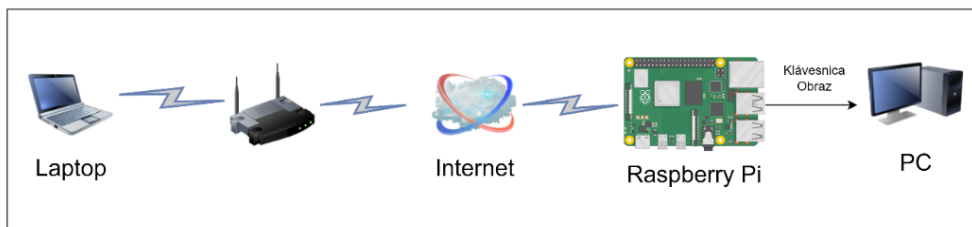
Kľúčové slová – IP KVM, Raspberry Pi, VNC, HID

I. ÚVOD

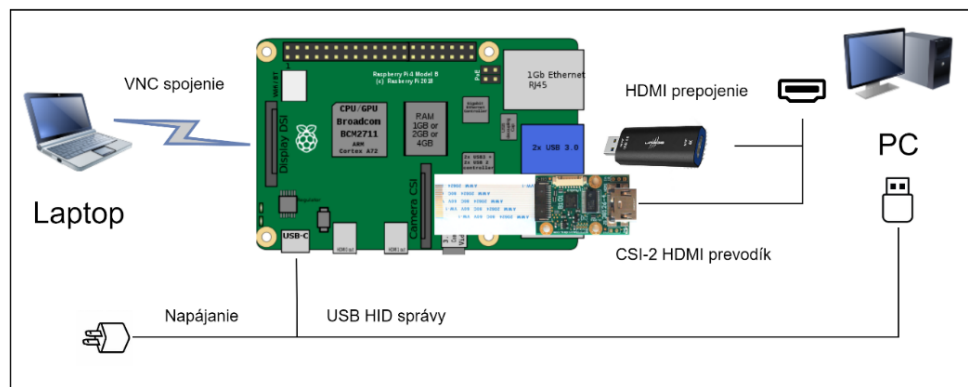
Špecializovaného Táto práca sa zaoberá problematikou vzdialenej správy systémov cez sieť za pomoci špecializovaného hardvéru. Cieľom tejto práce bolo vytvoriť KVM IP prepínač, ktorý možno použiť na ovládanie vzdialeného zariadenia, ako je server alebo akékoľvek iné zariadenie, ktoré má USB vstup a HDMI, VGA alebo DVI výstup. Cieľom práce je vytvoriť architektúru využívajúcu dedikovaný hardvér založený na operačnom systéme Linux a zostaviť KVM prepínač s pomocou veľmi obľúbeného mikropočítača Raspberry Pi a komponentov na získanie grafického výstupu z cieľového zariadenia a odosielanie signálov na zabezpečenie ovládateľnosti vzdialeného zariadenia. V práci sme sa zaoberali analýzou súčasných existujúcich riešení pre IP KVM prepínače pre vzdialenú správu zariadení. V druhej časti analýzy porovnávame a analyzujeme rôzne komponenty použité na vytvorenie IP KVM prepínača a riešenia používané na prenos obrazu a prenos signálov na ovládanie zariadenia. Následne navrhujeme architektúru IP KVM prepínača z hardvérového aj softvérového hľadiska. Ďalej v práci podrobne popisujeme vytvorenie fyzického prototypu KVM pomocou Raspberry Pi.

II. ANALÝZA PROBLEMATIKY

Stredné a veľké firmy potrebujú spravovať zariadenia na úrovni hardvéru na diaľku a v súčasnej dobe to už nie je žiadnym luxusom, ale práve naopak, nevyhnutnosťou. Mnohé podniky, ktoré vykonávajú správu serverov v dátových centrách, musia prekonávať niekoľkokilometrovú vzdialenosť. Je veľmi pravdepodobné, že firma má svojich klientov a zariadenia na druhom kontinente a správu zariadení vykonáva z Európy. V tomto prípade by to nebolo finančne, ale ani prakticky realizovateľné, aby firma posielala svojho správcu na takúto vzdialenosť. Softvérových riešení pre problematiku vzdialenej správy je na trhu mnoho, avšak je mnoho prípadov, kedy táto správa nestačí, a je potrebné zasiahnuť o úroveň nižšie. Správa hardvéru, vrátane kontroly a zmeny systému BIOS (Basic Input/Output System), reštartovanie a prepínanie napájania, správa sériových zariadení si vyžaduje priamy prístup ku klávesnici, videu a myši (KVM). Práve tu prichádzajú na rad KVM prepínače, ktoré umožňujú správcovi spravovať či už viacero zariadení naraz z jedného miesta nachádzajúc sa v danej firme, alebo taktiež na druhej strane zemegule a to s takmer rovnakým komfortom, ako keby fyzicky sedel za počítačom. Avšak nevýhodou KVM prepínačov je ich cena a častokrát aj zastaralý systém. Pre veľké firmy by cena KVM prepínačov nemala byť problém. Častokrát bežný používateľ technológií, ktorý sa o ne zaujíma, a chce mať vzdialený prístup ku svojmu serveru alebo počítaču, ale nechce to riešiť len pomocou čisto softvérového riešenia, si nemôže zadovážiť plnohodnotné KVM riešenie [1].



Obr. 1 Schéma IP KVM



Obr. 2 Návrh riešenia RPi IP KVM

III. NÁVRH RIEŠENIA

Na základe požiadaviek užívateľov bolo uvažované o zariadení, ktoré bude poskytovať zákazníčkovi vzdialenú správu jeho zariadení bez nutnosti zasahovania do vnútra zariadení, iba za pomoci ich vonkajších vstupov. Takéto zariadenie by dokázalo emulovať vstup z klávesnice a myši, prenášať výstupný obraz do mikropočítača a následne riadiť vzdialenú správu za pomoci softwaru. Navrhované riešenie by malo byť cenovo dostupné, a hlavne technicky spoľahlivé a pre bežného používateľa jednoduché na používanie. Návrh práce využíva model Klient-Server, kde mikropočítač Raspberry Pi predstavuje server, ktorý zabezpečuje prenos vstupov a grafického výstupu medzi správcom a zariadením za pomoci lokálnej siete - Obr. 1. Správca zariadenia sa dokáže vzdialene pripojiť na mikropočítač za pomoci programu VNC Viewer. Návrh riešenia sa skladá z troch základných blokov. Prvým z nich je koncové zariadenie (PC), ktoré sa snažíme vzdialene ovládať. Týmto zariadením môže byť hocikaké zariadenie ktoré disponuje aspoň jedným grafickým výstupom typu VGA, DVI, alebo HDMI a aspoň jedným USB portom. Taktiež zariadenie musí podporovať ovládanie za pomoci myši a klávesnice. Ďalším zariadením je mikropočítač Raspberry Pi (RPi) pomocou, ktorého dokážeme emulovať vstup klávesnice a taktiež prenášať obraz pre človeka, ktorý sa pokúša dané koncové zariadenie spravovať. Hostovské zariadenie (Laptop), ktoré používame na vzdialenú správu musí obsahovať iba aplikáciu VNC Viewer a disponovať pripojením na internet. Následne pomocou softvéru VNC sa dokážeme pripojiť k mikropočítaču Raspberry Pi z ktoréhokoľvek miesta na internete.

Za pomoci redukcie následne bude možné prenášať napájanie pre mikropočítač a zároveň prenášať dáta, ktoré reprezentujú znaky klávesy. Obraz bude možné zachytávať dvoma spôsobmi a to za pomoci aplikácie OBS Studio a USB nahrávacej karty alebo pomocou programu raspivid a CSI-2 HDMI prevodníku. Schému hardvérového prepojenia komponentov môžeme nájsť na Obr. 2.

Zo softvérového hľadiska sa riešenie skladá z Aplikácie VNC Viewer, ktorá nám umožní vzdialenú správu mikropočítača. Následne je možné vybrať možnosť zachytávania obrazu za pomoci aplikácie OBS Studio alebo programu raspivid. V ďalšom kroku je možné spustiť web server obsahujúci HTML stránku na zachytávanie kláves, preklad znakov na HID kód a následné preposlanie na rozhranie pripojené na koncové zariadenie.

A. Prenos grafického výstupu

Pre prenos obrazu z koncového zariadenia je potrebné implementovať zariadenie na zachytávanie obrazového signálu a daný signál následne spracovať. Na danú úlohu sme využili dva spôsoby

spracovania a to USB nahrávaciu kartu alebo rozhranie priamo na doske Raspberry Pi CSI-2, ktoré primárne slúži na zachytávanie obrazu z kamery. Spoločne s USB nahrávacou kartou bolo potrebné využiť softvér na spracovanie obrazu. V našom prípade sme zvolili aplikáciu OSB Studio pre jej širokú ponuku možností úpravy grafického výstupu a jednoduchej inštalácie aplikácie. Ďalšou možnosťou je využitie zbernice CSI-2. Raspberry Pi disponuje rozhraním MIPI CSI-2 na jednoduché pripojenie kamery pre grafický záznam vstupu. Zbernica MIPI CSI-2 podporuje až štyri dátové linky s maximálnou šírkou pásma 1 Gb/s na prepojenie s celkovou šírkou pásma 4 Gb/s. Okrem toho rozhranie využíva minimálny počet elektrických pripojení na zníženie zložitosti PCB. Dátová komunikácia je jednosmerná, od kamery resp. grafického vstupu k procesoru. Tento adaptér je špeciálne navrhnutý pre Raspberry Pi a podporuje vstup HDMI až do 1080p@30fps. Pri prenose vzniká len malé oneskorenie a podporuje kódovanie videa H.264. Dané rozhranie nám ponúka viacero výhod ako je veľmi nízka odozva s hodnotami cca. 100ms. Taktiež ponúka výhody ako natívna zmena rozlíšenia bez nutnosti využitia nahrávacieho softwaru, zmenu farieb a farebného priestoru a video kompresie [3]. Spolu s rozhraním CSI-2 sme využili na zachytávanie zabudovaný program pre snímanie a nahrávanie grafického vstupu s názvom raspivid. Pomocou prepínačov sa nám podarilo dosiahnuť stabilný prenos obrazu bez artefaktov. Najlepší výsledok dosahovala úprava vstupu s nižšie uvedenými parametrami *raspivid - o - -t0 - w1100 - h700 - fps30*. Po úspešnom zachytení obrazu je potrebné koncové zariadenie ovládať pomocou vstupno-výstupných periférií. Jednou z možností je využitie štandardu USB OTG. USB On-The-Go (USB OTG) je špecifikácia prvýkrát použitá koncom roku 2001, ktorá umožňuje zariadeniam, ako sú smartfóny alebo mikropočítače fungovať ako hostiteľ, čo umožňuje pripojenie iných USB zariadení ako sú klávesnica alebo myš. OTG Dual Role Device (DRC) je zariadenie, ktoré môže fungovať ako hostiteľ pre iné zariadenia alebo ako periférne zariadenie, čím umožňuje dve rôzne funkcie. V tomto režime sa OTG DRC chová ako USB host alebo ako USB zariadenie podľa potreby. Pri použití ako hostiteľ OTG DRC sa správa rovnako ako USB port na počítači a umožňuje pripojenie iných USB zariadení, ako sú napríklad klávesnice, myši, externé pevné disky a tlačiarne. OTG Dual Role Device využíva špeciálne zariadenie USB, ktoré sa nazýva "OTG controller", ktoré je schopné meniť svoju funkcionálnosť medzi hostiteľom a periférnym zariadením podľa toho, ako je potrebné. Tento OTG kontrolér umožňuje zariadeniu pracovať v režime duálnej úlohy a zvyšuje tak jeho flexibilitu a využiteľnosť [4].

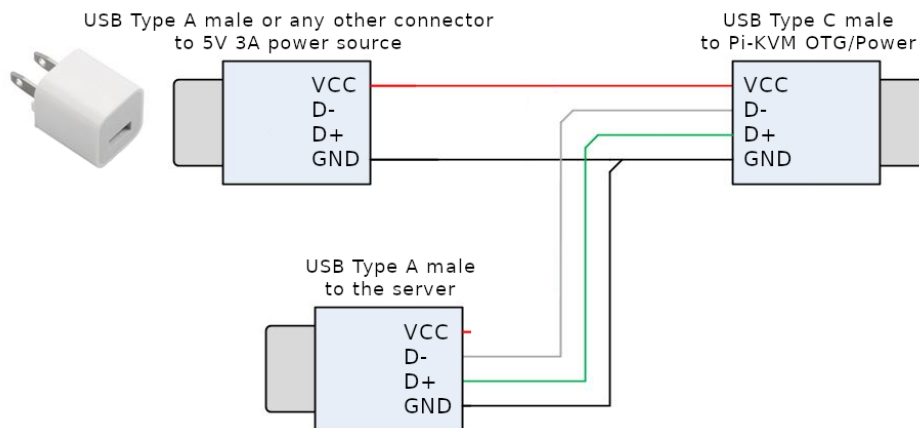
Skutočné vylepšenie však pochádza z rozhrania Human Interface Device (HID). To umožňuje aplikáciám v užívateľskom režime interagovať s operačným systémom, ako keby to boli USB zariadenia. Protokol HID definuje spôsob, akým zariadenia ako klávesnice, myši, joysticky a podobne komunikujú so systémom. HID zariadenia poskytujú užívateľské rozhranie pre interakciu s počítačom alebo iným zariadením. Emulácia klávesnice a myši v protokole HID znamená, že zariadenie vysiela dáta, ktoré simulujú pohyb myši alebo stlačenie kláves na klávesnici. Tieto dáta sú potom prijaté ako vstup do systému, akoby boli generované skutočnou klávesnicou alebo myšou [5]. Za pomoci skriptu je možné čítať a posilať inštrukcie, ktoré operačný systém prekladá na dáta pre klávesnicu. Daný ovládač využíva cestu */dev/hidg0* na komunikáciu s koncovým zariadením. Pri každom stlačení klávesa odošle klávesnica 8-bitovú správu s názvom report. Pre signalizáciu ukončenia zadávania dokážeme odoslať 8 bitovú správu pozostávajúcu zo samých núl. Aby sme túto funkcionálnosť mohli využiť, je potrebné vyriešiť problém s napájaním. Keďže Raspberry Pi disponuje len jedným USB C portom, bolo potrebné vytvoriť USB redukciu – spiltter, ktorá sa skladá z USB C konektora a dvoch USB A konektorov. Konektor bielej farby slúži na prívod napájania z napájacieho zdroja a USB port čiernej farby bude slúžiť ako vysielač emulovaných signálov pre koncové zariadenie, ktoré sa snažíme vzdialene ovládať. Na Obr. 3 je možné vidieť schému danej redukcie.

Za pomoci upravenej verzie softvéru *KeyMimePi*, sa nám úspešne podarilo vytvoriť aplikáciu vo webovom frameworku Flask, ktorý je napísaný v programovacom jazyku Python. V prvom kroku bolo potrebné povoliť podporu USB gadget služby pomocou vytvoreného skriptu nachádzajúceho sa v mikropočítači Raspberry Pi. Po reštarte sme vytvorili virtuálne prostredie, vybrali sme adresár, do ktorého sme ho chceli umiestniť - v našom prípade venv a spustili sme modul venv ako skript s cestou k adresáru. Ďalej sme za pomoci príkazu nainštalovali potrebné balíky z textového súboru requirements.txt. Celý program je možné nájsť v priečinku app.

V danom priečinku sa nachádzajú python súbory main.py, hid.py a js_hid.py a webová stránka pozostávajúca z html súboru index.html a css súboru style.css a javascript súboru app.js. Stránka sa skladá z indikátora stavu programu. Pod spomínaným indikátorom môžeme vidieť históriu stlačených kláves. Náhľad web stránky je možné vidieť na Obr. 4.

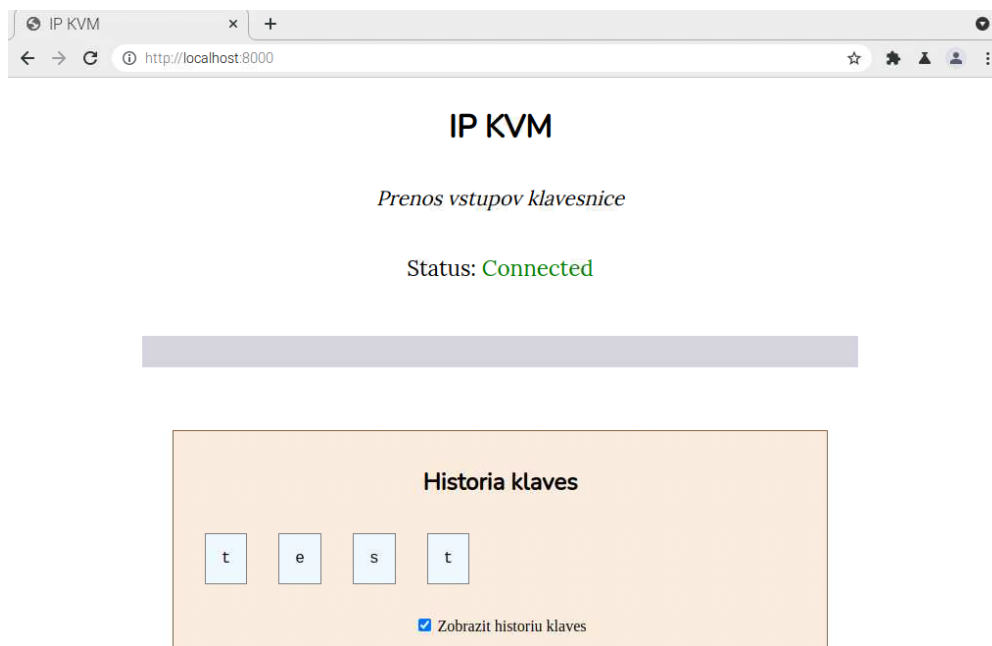
IV. ZÁVER

Vzdialená správa počítačov sa stáva čoraz obľúbenejšou vo všetkých typoch organizácií, preto aktuálnosť danej témy je v súčasnosti veľmi vysoká. Vďaka rozvoju technológií sa vzdialená



Obr. 3 USB Redukcia

správa počítačov stáva ľahko dostupnou aj pre malé podniky a dokonca aj pre jednotlivcov. Táto práca opisuje metódy implementácie KVM-over-IP riešení, na základe ktorých navrhuje modulárny KVM prepínač, ktorý distribuuje príkazy do bezdrôtovej siete, v ktorej sa nachádzajú koncové zariadenia pripojené k hosťovskému systému (systém, ktorý chceme vzdialene ovládať). V súčasnosti je potreba vzdialenej správy vyššia ako kedykoľvek predtým. Dôvodom bola pandémia COVID-19, ktorá viedla k väčšiemu využívaniu práce z domu a vzdialenej spolupráce. Mnoho spoločností muselo prejsť na vzdialenú správu rýchlo, aby mohli zabezpečiť pokračovanie svojej činnosti počas pandémie. Vzdialená správa sa stala pre mnoho spoločností a zamestnancov nevyhnutnosťou a mnoho z nich ju bude pravdepodobne používať aj po skončení pandémie. Môže to byť kvôli úsporám na nákladoch na prevádzku, väčšej flexibilitě pre zamestnancov alebo kvôli možnosti získať prístup k širšiemu množstvu uchádzačov, keďže nebude potrebné aby pracovali na jednom mieste. Vzdialená správa však nie je vhodná pre všetky spoločnosti a všetkých zamestnancov. Niektoré typy práce môžu byť ťažšie vykonávateľné na diaľku a niektorí zamestnanci si môžu ťažšie zvykať na prácu z domu. Preto je dôležité, aby spoločnosti zvažili výhody a nevýhody vzdialenej správy a rozhodli sa, či je pre nich naozaj vhodná. Dané riešenie by predstavovalo konkurenciu pre súčasne hardvérové riešenia v cenovej relácii 100-150€. Oproti tomu súčasné hardvérové riešenia na báze IP KVM switchov sa pohybujú v cenových reláciách od 500 do 3000€ [6]. Takéto riešenie by malo len zanedbateľnú spotrebu v jednotkách wattov a ponúkalo by nepretržitú prevádzku. Taktiež architektúra mikropočítača Raspberry Pi umožňuje



Obr. 4 Ovládanie klavesnice

jednoduchú zmenu a prispôsobiteľnosť a možnosť doplnenia funkcií po nasadení do prevádzky. Prácu by bolo taktiež možné rozšíriť o vzdialené ovládanie napájania koncového zariadenia. Pomocou GPIO pinov dostupných na Raspberry Pi s vytvorením splitteru pre dvojpinový konektor ATX s označením „Power SW“ by bolo možné vyslať signály pre zapnutie alebo vypnutie zariadenia. V takomto prípade by nebola potrebná žiadna obsluha koncového zariadenia. Taktiež by bolo možné vytvoriť plánovač pre zapnutie a vypnutie zariadenia.

PODĎAKOVANIE

Článok bol podporený v rámci projektu - 002TUKE-4/2021 Implementácia moderných metód a foriem výučby v oblasti kyberbezpečnosti k požiadavkám praxe.

LITERATÚRA

- [1] Vertiv Group Corp. (2022). What Is a KVM Switch?. [cit. 2023-04-20] URL: (<https://www.vertiv.com/en-emea/about/news-and-insights/articles/educational-articles/what-is-a-kvm-switch/>)
- [2] RealVNC Limited. (2023). What is vnc remote access technology. [cit. 2023-04-20] URL: (<https://discover.realvnc.com/what-is-vnc-remote-access-technology>)
- [3] Waveshare Electronics. (2023). HDMI to CSI Adapter. [cit. 2023-04-20] URL: (https://www.waveshare.com/wiki/HDMI_to_CSI_Adapter)
- [4] Texas Instruments Incorporate. (2007). USB On-the-Go Solutions. [cit. 2023-04-20] URL: (<https://www.ti.com/lit/ml/sprp605/sprp605.pdf>)
- [5] USB Implementers Forum. (2023). Human Interface Devices (HID) Specifications and Tools. [cit. 2023-04-20] URL: (<https://www.usb.org/hid>)
- [6] ATEN International Co. (2023). ATEN CS1798. [cit. 2023-04-20] URL: (<https://www.aten.com/global/en/products/kvm/rack-kvm-switches/cs1798/?p=1>)

Možnosti využitia elektrolyzéra pre elektroenergetické účely

Vladimír SZOMOSI, Dušan MEDVEĎ

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

vladimir.szomosi@student.tuke.sk, dusan.medved@tuke.sk

Abstrakt — Tento príspevok sa zaoberá možnosťami využitia elektrolyzéra pre energetické účely. Jedná sa o navrhnutý a skonštruovaný model elektrolyzéra, ktorý vyrába zmes plynov vodíka a kyslíka. Na konkrétnom modeli boli realizované experimentálne merania za účelom zistenia jeho účinnosti a strát. Výsledná zmes plynov, nazývaná HHO plyn, bola aplikovaná ako palivo pre plynový horák. Tento príspevok je rozdelený do niekoľkých častí, pričom sa zaoberá návrhom a konštrukciou elektrolyzéra, experimentálnymi meraniami a aplikáciou výstupnej zmesi plynov.

Kľúčové slová — Brownov plyn, elektrochemická premena, elektrolyza, elektrolyt, HHO, HHO generátor, HHO plyn, kyslík, vodík

I. ÚVOD

V posledných desaťročiach sa ľudstvo stáva čoraz závislejšie od zdrojov elektrickej energie. Rastúci dopyt po elektrine urýchľuje hľadanie spôsobov ako efektívne využiť a skladovať zdroje, ktoré sú teraz dostupné. Jedným zo spôsobov akumulácie elektrickej energie je premena nestabilného energetického zdroja na spoľahlivejší a dlhodobo uskladniteľný zdroj. Využitím elektrochemickej premeny vody napájanej z neregulovateľných obnoviteľných zdrojov. Produktom elektrochemickej premeny vodných roztokov je vodík. Vodík je uskladniteľným zdrojom energie, môže byť použitý ako palivo pre palivové články, pre spätnú výrobu elektrickej energie, alebo pre mechanický pohon spaľovacích motorov.

II. ELEKTROLÝZA

Elektrochemický proces, pri ktorom látkou prechádza jednosmerný elektrický prúd, ktorý spôsobí chemickú zmenu látky na atómovej úrovni (oxidáciu alebo redukciu), je elektrolyza. Proces prebieha v elektrolytickom článku pozostávajúceho z kladných, neutrálnych a záporných elektród, ktoré sú od seba mechanicky oddelené a ponorené do vodivého roztoku (elektrolytu) obsahujúceho kladné a záporne nabité ióny. Pôsobením zápornej polarita jednosmerného prúdu privedeného na katódu putujú zložky elektrolytu na povrch elektródy, kde sa spájajú s elektrónmi, kde prebieha proces redukcie. Oxidácia (proces odovzdávania valenčných elektrónov) prebieha na kladnej elektróde (anóde). Produkty oxidácie alebo redukcie môžu byť nové atómy alebo látky, to závisí od toho, či chceme aby chemické procesy prebehli v roztoku alebo na elektródach [12].

III. BROWNOV PLYN (HHO)

Yull Brown bol bulharský chemik, získal patent na elektrolyzér v roku 1977. Brownov plyn, HHO alebo oxyhydrogen, je stabilná zmes plynov vodíka a kyslíka (v pomere 2:1), ktorá vzniká pri elektrochemickej premene elektrolytu. Vzniká tak plyn s vyšším energetickým potenciálom ako má vodík. Elektrolyt sa prechodom elektrického prúdu rozkladá na vodík a kyslík, rozdielom je, že v elektrolyzéroch na výrobu vodíka je použitá polopriepustná membrána zabráňujúca zmiešaniu jednotlivých plynov, pričom v HHO elektrolyzéri membrána nie je, vzniká zmes plynov. Z jedného litra elektrolytu je možné získať elektrolyzou 1860 litrov Brownovho plynu. Brownov plyn je skladovateľný, v kontakte s iskrou alebo ohňom imploduje (vzniká malé množstvo vodnej pary), pri atmosférickom tlaku horí chladným plameňom o teplote do 130°C plameň pri správnych podmienkach môže dosahovať maximálnu teplotu až 2800°C (zaujímavé je, že svojou relatívne nízkou teplotou dokáže sublimovať volfrám) [12].

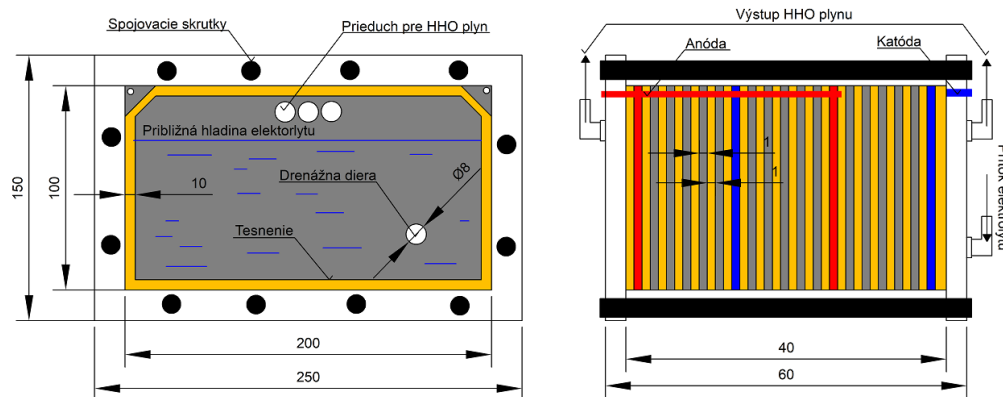
IV. NÁVRH A KONŠTRUKCIA ELEKTROLYZÉRA

Samotný modelový aparát je zložený z dvoch častí, hlavnej časti, ktorou je elektrolyzér a z obslužných zariadení, ktoré zabezpečujú jeho chod a bezpečnosť. Celý aparát nazývame HHO generátor. Elektrolyzér produkuje už spomínaný HHO plyn elektrochemickou reakciou medzi elektrolytom a elektródami, v prostredí elektrolyzéra typu dry cell (suchý článok). Elektrolyzér je napájaný zo spínaného zdroja jednosmerného prúdu. Nádržka na elektrolyt plní funkciu zásobníka elektrolytu pre elektrolyzér a zároveň do nej vstupuje aj HHO plyn vo forme peny s obsahom elektrolytu, pena sa rozptýli a elektrolyt zostáva v nádržke. Do čističky plynu vstupuje čistý plyn bez obsahu elektrolytu, plyn je zbavený vodnej pary a nečistôt. Tlakovaná nádoba zabezpečuje horenie s nižším rizikom spätného šľahnutia plameňa pri horení HHO plynu. Spätnému šľahnutiu predchádzajú aj jednocestné ventily, suché predlohy a zachytávače spätného vzplanutia. Spätné šľahnutie plameňa môže nastať spaľovaním HHO plynu a tiež pri aplikáciách spojených s horením nastane, ak je malá výstupná rýchlosť plameňa, pri prehriatí nastavca horáka, alebo pri jeho nedokonalom uchytení [8].

A. Elektrolyzér

Elektrolyzér typu dry cell je zložený z elektródových zväzkov, článkov zložených z anódy, katódy a neutrálnych elektród. Jednotlivé články sú paralelne spojené so zdrojom jednosmerného prúdu. Hlavným problémom návrhu je správny výber materiálu elektród, vzhľadom na silné oxidačné účinky elektrolytu použitého ako katalyzátor elektrolyzy. V našom prípade bol ako elektrolyt zvolený 15 % vodný roztok hydroxidu sodného. Vybraným materiálom, ktorý vyhovoval týmto podmienkam je oceľ s antikoróznou úpravou, ľudovo povedané nehrdzavejúca oceľ. Je vysoko legovaná chrómom, niklom a mangánom, ktoré zabezpečujú vysokú odolnosť a mechanickú pevnosť [10].

Články je potrebné zložiť tak, aby sa jednotlivé elektródy nedotýkali. Na oddelenie elektród sme vybrali gumené tesnenie, ktoré je schopné odolávať tlaku a elektrolytu. Prúdenie HHO plynu zabezpečíme prieduchmi v hornej časti elektród, ktoré prepoja jednotlivé komory. Prítok elektrolytu do komôr je realizovaný drenážnymi dierami v spodnej časti každej elektródy. Elektródy a tesnenia sú spojené akrylovými panelmi, ktoré slúžia ako vývodové plochy pre HHO plyn a prívodové plochy pre elektrolyt [5].



Obr. 1 Návrh elektrolyzéra

B. Elektródy

Elektródy sú vyrobené z antikoróznej ocele bez povrchovej úpravy s hladkým lesklým povrchom. V budúcnosti zvažujeme opieskovanie povrchu, ktoré zväčší aktívnu plochu elektród na mikroskopickej úrovni. Týmto krokom môžeme zväčšiť produkciu plynu a zabezpečiť rýchlejší odvod bubliniek plynu z povrchu elektród. V elektrolyzéri sú použité dva typy elektród, napájané elektródy anóda, katóda a medzi nimi sa nachádzajúce neutrálne elektródy [9].

Minimálne napätie potrebné pre začatie procesu elektrolyzy je 1,24 V medzi dvomi neutrálnymi elektródami. Pri vyššom napätí sa nadbytočné napätie mení na tepelné straty, ale množstvo elektrolytu potrebného na zníženie odporu vody zapríčiňuje to, že napätie 1,24 V nie je dostatočné pre elektrolyzu vodného roztoku, preto minimálne napätie pre začatie elektrolyzy medzi dvomi elektródami závisí od použitého elektrolytu; napríklad pri NaOH (hydroxid sodný) je minimálne napätie 1,69 V a pri KOH (hydroxid draselný) 1,67 V. V našom prípade používame kombináciu paralelne spojených sériových článkov, ktoré majú medzi sebou spoločné anódy alebo katódy v usporiadaní $(- n n n n n + n n n n n - n n n n n +, n - \text{neutrálna elektróda})$, pri výpočtoch musíme zohľadniť to, že sa nejedná o samostatný článok, ale o tri články spojené paralelne. Kompletné usporiadanie tvorí 19 elektród zložených do troch článkov zapojených paralelne za sebou [2].

C. Výpočet množstva generovaného HHO plynu

Výpočet produkcie pomocou Faradayových vzťahov nezahŕňa typ elektrolyzéra, veľkosť aktívnej plochy elektród, počet elektród, článkov a typ elektrolytu využívaného pri elektrolyze. Tieto výsledky sú len orientačné a slúžia na určenie približnej produkcie pri zvolenom usporiadaní elektrolyzéra. Budú porovnávané s meraniami na reálnom modelovom aparáte. Množstvá vyprodukovaných plynov vodíka a kyslíka vypočítame odvodením Faradayových vzťahov elektrolyzy.

Celkové množstvo vyprodukovaného HHO plynu V je súčtom objemu plynu vodíka V_H a objemu plynu kyslíka V_O v HHO plyne, ktoré sa vyprodukuje na všetkých článkoch po dobu jednej minúty. Jednotlivé množstvá objemu získavame dosadením do upravených Faradayových vzťahov, pri štandardných atmosférických podmienkach a teplote okolia 20°C.[6].

Výpočet objemu vodíka:

$$V_H = \frac{R \cdot T \cdot Q}{F \cdot p \cdot z_H} = \frac{8,314 \cdot 293 \cdot 600}{9,648 \cdot 10^4 \cdot 101,325 \cdot 2} = 0,0748 \text{ l/min} \quad (1)$$

kde: R plynová konštanta, 8,314 J·K⁻¹·mol⁻¹

T teplota okolia

Q elektrický náboj

z_H počet voľných elektrónov vodíka

F Faradayova konštanta

p atmosférický tlak

Výpočet objemu kyslíka:

$$V_O = \frac{R \cdot T \cdot Q}{F \cdot p \cdot z_O} = \frac{8,314 \cdot 293 \cdot 600}{9,648 \cdot 10^4 \cdot 101,325 \cdot 4} = 0,0374 \text{ l/min} \quad (2)$$

kde: z_O počet voľných elektrónov kyslíka

Celkové množstvo vyprodukovaného HHO plynu:

$$V_{\text{článok}} = V_H + V_O = 0,0748 + 0,0374 = 0,112 \text{ l/min} \quad (3)$$

$$V = n \cdot V_{\text{článok}} = 3 \cdot 0,112 = 0,336 \text{ l/min} \quad (4)$$

kde: n počet článkov elektrolyzéra

Zo zvyšujúcim sa prúdom I dodaným na svorky elektrolyzéra rastie elektrický náboj Q , dochádza tak k zvyšovaniu produkcie HHO plynu. Vyšší prúd môže zapríčiniť zahrievanie elektrolytu, to má za následok vyšší podiel vodných výparov v zmesi HHO plynu, spôsobuje to tiež oxidáciu na povrchu elektród. Do HHO plynu sa týmto vylučujú oxidy chrómu, ktoré znečisťujú výsledný HHO plyn [4].

Tab. 1 Porovnanie nameraných a vypočítaných predpokladaných výsledkov produkcie HHO plynu

Prúd zdroja [A]	Nameraná produkcia HHO plynu [l/min]	Predpokladaná produkcia HHO plynu [l/min]	Chyba výpočtu [%]
10	0,731	0,336	54,04
15	1,03	0,504	51,07
20	1,11	0,675	39,19
25	1,429	0,841	41,15
29	1,714	1,008	41,19

Vypočítané hodnoty produkcie HHO plynu pomocou upravených Faradayových vzťahov na prvý pohľad nie sú porovnateľné s nameranými hodnotami. Dôvodom môže byť obmedzenie vstupných parametrov pri výpočte predpokladanej produkcie, napr. aktívna plocha elektród alebo pracovný prúd a napätie pre jednotlivé články. Ďalším dôvodom spôsobujúcim rozdiel môže byť čistota vyprodukovaného HHO plynu, napätie na elektródach sa premieňa na tepelné straty, ktoré môžu spôsobiť vyparovanie elektrolytu, vzniká vodná para produkcia HHO plynu narastá, ale nezvyšuje sa obsah koncentrácie vodíka a kyslíka. Koncentráciu množstiev vodíka a kyslíka v zmesi HHO plynu vieme odmerať napríklad pomocou plynových snímačov alebo kalorimetrickým meraním. To je jeden zo spôsobov, ako vieme overiť správnosť vypočítaných hodnôt.

Pre dodatočné overenie produkcie plynu použitím aj iných parametrov elektrolyzéra sme použili voľne dostupný softvér HHO Hydrogen Generator Cell Configurator [2]. Je schopný zo vstupných parametrov, ako sú aktívna plocha elektród (časť elektród, ktorá nie je pokrytá tesnením), pracovný prúd a napätie pre jednotlivé články, počet neutrálnych elektród a počet článkov v elektrolyzéri, vypočítať predpokladanú produkciu plynu.

Tab. 2 Porovnanie nameraných a softvérom vypočítaných hodnôt produkcie HHO plynu

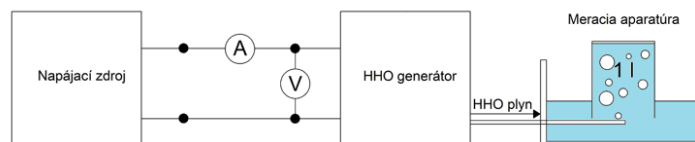
Prúd zdroja [A]	Nameraná produkcia HHO plynu [l/ min]	Predpokladaná produkcia HHO plynu [l / min]	Chyba výpočtu [%]
10	0,731	0,626	14,36
15	1,03	0,94	8,74
20	1,11	1,25	-12,61
25	1,429	1,565	-9,52
29	1,714	1,879	-9,63

Hodnoty produkcie HHO plynu vypočítané využitím vyššie uvedeného softvéru, sú porovnateľné s nameranými hodnotami produkcie. Overili sme si tak vierohodnosť softvéru a správnosť jeho výpočtu. Porovnaním nameraných a vypočítaných hodnôt sme určili chybu výpočtu produkcie HHO plynu.

V. MERANIE A APLIKÁCIA

A. Meranie produkcie HHO plynu

Cieľom predmetného merania bolo zistiť produkciu HHO plynu a overiť predpokladané vypočítané hodnoty navrhovaného HHO generátora na už funkčnom modeli zhotoveného podľa návrhu. Princípom merania je odmerať produkciu HHO plynu pri rôznych hodnotách napájacieho prúdu spínaného jednosmerného zdroja a pozorovať nárast produkcie plynu, ktorý bude meraný v jednotkách liter za minútu. Zvyšovanie prúdu je priamoúmerné výške produkcie plynu a zároveň rýchlosti rastu teploty elektrolytu. Očakávané je, že výsledky merania budú porovnateľné s vypočítanými hodnotami. Meranie sme realizovali v odmernom valci s objemom jedného litra, ktorý naplníme vodou. HHO plyn postupne vytláča vodu z valca, časovačom určíme presný čas vyprodukovania jedného litra HHO plynu.



Obr. 2 Meranie produkcie HHO plynu

Tab. 3 Meranie č. 1 produkcia plynu HHO generátora

Prúd zdroja [A]	Napätie zdroja [V]	Čas produkcie 1 l HHO plynu [s]	Produkcia HHO plynu [l/min]
10,1	12,89	82	0,731
15,3	13,32	58	1,03
19,8	13,45	54	1,11
25,2	13,68	42	1,429
29,1	14,54	35	1,714

Tab. 4 Meranie č. 2 produkcia plynu HHO generátora

Prúd zdroja [A]	Napätie zdroja [V]	Čas produkcie 1 l HHO plynu [s]	Produkcia HHO plynu [l/min]
10,3	12,89	94	0,638
15,1	12,97	65	0,923
20,05	13,2	44	1,36
25,2	13,44	37	1,621
29,3	16,67	35	1,714

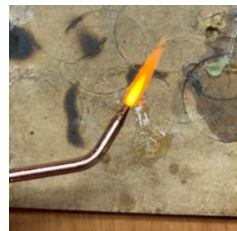
Predmetom merania bolo odmerať produkciu HHO na realizovanom fyzickom modeli. Výsledky prvého merania korešpondovali s predpokladom, preto sme pre skúšku správnosti merania zopakovali. Porovnaním výsledkov vidíme, že výsledky oboch meraní sú porovnateľné s priemerným rozdielom v čase vyprodukovaného jedného litra HHO plynu $t = \pm 6$ s. Rozdiel v produkcii mohol vzniknúť nepresným nastavením veľkosti elektrického jednosmerného prúdu, ktorý nie je konštantný, ale mení sa v závislosti od výšky hladiny v elektrolyzéri. Pri meraní sme pozorovali lineárne zvýšenie produkcie HHO závislé od veľkosti priloženého prúdu. Pri veľkom množstve bublínok sme pozorovali vytlačenie elektrolytu späť do elektrolytickej nádržky, znižovala sa tak aktívna plocha elektród a zároveň aj produkcia HHO plynu [11].

B. Aplikácia HHO plynu ako paliva pre horák

Pri spaľovaní vzniká vysokoteplotný plameň, ktorý môže byť použitý napríklad na rezanie tenkých plechov, tavenie kovov s nízkou teplotou tavenia, ako aditívum pri spaľovacích procesoch alebo napríklad na ohrev vody. Výsledný produkt HHO generátora (HHO plyn), zásobujeme pod stálym tlakom do horáka, ktorý ho nepretržite spaľuje. Pri prevádzke HHO generátora sa plyn uskladňuje v tlakovej nádobe, stály tlak plynu vstupujúci do horáka zabezpečíme vhodným výberom horáka. Horák spaľujúci HHO plyn využíva potrebné množstvo vzduchu, teda kyslíka, obsiahnutého priamo v zmesi HHO plynu. Preto spaľovací proces závisí iba od prítoku HHO plynu. HHO plameň vďaka svojej vysokej teplote dokáže rýchlo ohriať kovové materiály. Použitie HHO plynu na rezanie a zváranie je výhodné, úzky plameň dokáže rezať plochy materiálu presnejšie a hladšie [7].



Obr. 3 Plameň HHO horáka



Obr. 4 nevyhovujúci HHO plameň

Generovaný plyn HHO bol dočasne uskladnený v tlakovej nádobe, keďže produkcia plynu za minútu je väčšia ako spotreba horáka, pričom bol kontrolovane vypúšťaný do prostredia tak, aby vzniknutý tlak nespôsobil poruchu na elektrolyzéri. Ak použijeme injektor s väčším priemerom, v konkrétnom prípade na Obr. 4, kde je použitý injektor s priemerom 0,8 mm, potrebujeme vyvinúť väčší tlak v tlakovej nádobe pre dosiahnutie rovnakej kvality spaľovania HHO plynu ako pri injektore s priemerom 0,23 mm zobrazený na Obr. 5. Vyšší tlak je pre konkrétny HHO generátor nežiadúci, preto sme zvolili použitie horáka s menším priemerom injektora. Výkon plameňa horáka spaľujúceho HHO plyn bol približne v rozmedzí 60 až 70 W. Tento výkon je v porovnaní so vstupným napájacím výkonom stratový. Straty pozostávajú z tepelných strát, kde sa energia stráca do okolia a z úmyselného vypúšťania HHO plynu z tlakovej nádoby. V budúcich vylepšeniach generátora plánujeme so zvýšením tlaku v akumuláčnej nádobe. Tlak dokáže upraviť zvýšením mechanickej pevnosti elektrolyzéra a to konkrétne tesnením. Zvýšením tlaku zabezpečíme lepší prítok plynu horákom, dosiahneme tak zvýšenie výkonu HHO plameňa [7].

VI. ÚČINNOSŤ A STRATY HHO GENERÁTORA

A. Účinnosť

Účinnosť samotného procesu elektrolýzy, bez ohľadu na straty pri stanovených podmienkach, bola uvažovaná pri štandardnej teplote okolia $T = 20^{\circ}\text{C}$ a pri atmosférickom tlaku $p = 101,235$ kPa. Tlak stanovuje, koľko HHO plynu vyprodukujeme za jednotku času pri konštantnom napájaní jednosmerným elektrickým prúdom a napätím. Uvažovali sme, že v jednom litri vyprodukovaného HHO plynu, zloženého z vodíka a kyslíka v pomere 2:1, nachádza sa 0,8 litra vodíka a 0,2 litra kyslíka.

$$\eta = \frac{n \cdot \Delta H \cdot V}{U \cdot I \cdot t} \cdot 100 \quad [\%] \quad (5)$$

Počet molov vodíka n [-] v jednom litri vyprodukovaného HHO plynu je 0,03254 mol, vypočítaný z obsahu vodíka v jednom litri zmesi HHO plynu [10]. Teploto spáleného vodíka $\Delta H = 285$ kJ/mol, napájaného napätím U [V] a prúdom I [A], pri objeme vyprodukovaného plynu V [l] za určitý čas t [s] [1].

Z Tab. 5 pre výpočet účinnosti elektrochemickej premeny elektrolyzéra zisťujeme, že so zvyšujúcou sa produkciou HHO plynu klesá účinnosť elektrolyzéra. Pri priebehu elektrolýzy sa na elektródach produkujú plynové bubliny, pri pomalom odvode plynu z komôr elektrolyzéra pokrývajú povrch elektród. Dochádza k nárastu odporu, teda k tzv. „nadpotenciálu“. Nadpotenciál spôsobuje ťažší priebeh reakcie v elektrolyzéri, výsledkom je zvýšenie napätia zdroja.

Tab. 5 Účinnosť HHO generátora

Prúd zdroja [A]	Napätie zdroja [V]	Produkcia HHO plynu meranie č.1 [l/ min]	Účinnosť HHO generátora [%]
10,1	12,89	0,731	86,668
15,3	13,32	1,03	78,118
19,8	13,45	1,11	64,424
25,2	13,68	1,429	64,07
29,1	14,54	1,714	62,613

Straty, ktoré vznikli pri elektrochemickej premene, sú spôsobené Joulovým teplom, ktoré vzniká pri prechode elektrického prúdu generátorom. Vyšší vstupný prúd zapríčiňuje klesanie účinnosti HHO generátora, zvýšenie odporu v dôsledku prechodových strát závisí rovnako od teploty elektrolytu ako aj od koncentrácie. K celkovým stratám taktiež patrí aj znečistenie HHO plynu chemickými stratami. Chemické straty vznikajú kontamináciou HHO zmesi, z elektród z nehrdzavejúcej ocele, kde sa pri vysokých teplotách elektrolytu vylučujú napríklad oxidy chrómu a oxidu železitého. Chemické straty vieme znížiť zvolením optimálnej pracovnej teploty elektrolytu.

VII. ZÁVER

Cieľom príspevku bolo preskúmať možnosti využitia elektrolyzéroov na energetické účely a z teoretických poznatkov navrhnúť a skonštruovať konkrétny elektrolyzér pre generovanie HHO plynu. Podľa návrhu sme skonštruovali a uviedli do prevádzkového stavu elektrolyzér, spoločne s obslužnými zariadeniami, ktoré potrebuje k práci. Pomocou HHO generátora sme experimentálne overili predpokladané výpočty produkcie HHO plynu v čase a výstupný HHO plyn použili na konkrétnu aplikáciu.

POĎAKOVANIE

Tento príspevok vznikol vďaka Agentúre na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0576 a APVV-21-0312, ako aj Vedeckou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu SR a SAV, podporená grantom VEGA 1/0757/21.

ZOZNAM POUŽITEJ LITERATÚRY

- [1] BALAJKA, J. Vodík a iné nové nosiče energie. Bratislava: Alfa, 1982.
- [2] BIGGS, D. Cell Configurator. In [online]. [cit. 2023-03-08]. Dostupné na internete: <http://www.hho4free.com/configurator/cell_configurator.html>.
- [3] EL SOLY, A.K., El Kady, M. A., Farrag Ahmed El Fatih., Gad M. S. Comparative experimental investigation of oxyhydrogen (HHO) production rate using dry and wet cells. In International Journal of Hydrogen Energy [online]. 2021. Vol. 46, no. 24. [cit. 2023-03-20]. Dostupné na internete: <<https://www.sciencedirect.com/science/article/pii/S0360319921002056>>.
- [4] JABALLA, K., Elmasri, G., Alqaed, S., Maatug, B. Design of Dry HHO cell to Produce Hydrogen using Electrolysis. In [online]. 2022. s. 49–58. Dostupné na internete: <https://www.researchgate.net/profile/Khalid-Jaballa-3/publication/358969408_Design_of_Dry_HHO_cell_to_Produce_Hydrogen_using_Electrolysis/links/621fccbd4c4fa27cd21a21e/Design-of-Dry-HHO-cell-to-Produce-Hydrogen-using-Electrolysis.pdf>.
- [5] MARTON, D. Kalkulačky a vzorce. In [online]. [cit. 2023-03-04]. Dostupné na internete: <http://hho.sk/HHO_HOME_files/HHO_kalkulacky.htm#Kalkula%C4%8Dky%20-%20C3%9A%C4%8Dinno%C5%A5%20HHO%20gener%C3%A1tor>.
- [6] MORENO-SORIANO, R., Soriano-Moranchel, F., Flores-Herrera, L., Sandoval-Pineda, J., González-Huerta, R. Thermal Efficiency of Oxyhydrogen Gas Burner. In Energies [online]. 2020. Vol. 13, no. 20. [cit. 2023-04-24]. Dostupné na internete: <<https://www.mdpi.com/1996-1073/13/20/5526>>.
- [7] NABIL, T., KHAIRAT DAWOOD, M.M. Enabling efficient use of oxy-hydrogen gas (HHO) in selected engineering applications; transportation and sustainable power generation. In Journal of Cleaner Production [online]. 2019. Vol. 237. [cit. 2023-04-29]. Dostupné na internete: <<https://www.sciencedirect.com/science/article/pii/S0959652619326587>>.
- [8] OLIVARES-RAMÍREZ, J.M., Campos-Cornelio, M.L., Uribe Godínez, J., Borja-Arco, E., Castellanos, R. Studies on the hydrogen evolution reaction on different stainless steels. In International Journal of Hydrogen Energy [online]. 2007. Vol. 32, no. 15, s. 3170–3173. [cit. 2023-04-04]. Dostupné na internete: <<https://www.sciencedirect.com/science/article/pii/S036031990700064X>>.
- [9] PUBCHEM Sodium Hydroxide. In [online]. [cit. 2023-05-02]. Dostupné na internete: <<https://pubchem.ncbi.nlm.nih.gov/compound/14798>>.
- [10] WANG, M., Wang, Z., Gong, X., Guo, Z. The intensification technologies to water electrolysis for hydrogen production – A review. In Renewable and Sustainable Energy Reviews [online]. 2014. Vol. 29, s. 573–588. [cit. 2023-04-12]. Dostupné na internete: <<https://www.sciencedirect.com/science/article/pii/S1364032113006308>>.
- [11] Electrolysis | Definition, Uses, & Facts | Britannica. In [online]. [cit. 2022-12-15]. Dostupné na internete: <<https://www.britannica.com/science/electrolysis>>.
- [12] What Is Brown's Gas? In Eagle-Research [online]. [cit. 2023-04-02]. Dostupné na internete: <<https://eagle-research.com/what-is-browns-gas/>>.

Porovnanie AC a DC ostrovnej prevádzky nn siete

Miroslav ANDRAŠ, Dušan MEDVEĎ

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

miroslav.andras@student.tuke.sk, dusan.medved@tuke.sk

Abstrakt — Tento článok sa venuje porovnaniu AC a DC ostrovnej prevádzky v sieti nízkeho napätia. Porovnanie týchto dvoch prevádzok prebiehalo predovšetkým v simulačnom programe MATLAB/Simulink. Prebiehali tam simulácie, ktoré ozrejmujú rozdiely medzi využitím buď AC alebo DC ostrovnej prevádzky. Navrhnuté prevádzky obsahujú dve batériové stanice, ktoré ako hlavný zdroj energie využívajú fotovoltaický systém, ktorý následne dobíja akumulátory. Odberateľmi elektrickej energie sú domácnosti, konkrétne v modeli sa nachádzajú tri odberné miesta, ktoré majú simulovať spotrebu elektrickej energie v domácnostiach. Okrem návrhu modelov v Simulinku článok sa venuje aj praktickému návrhu týchto prevádzok, predovšetkým z dôvodu porovnania z hľadiska finančných nákladov.

Kľúčové slová – AC ostrovná prevádzka, DC ostrovná prevádzka, fotovoltaický systém, batériové úložisko

I. ÚVOD

V súčasnosti prevažne používame v distribučnej elektrickej sieti striedavý elektrický prúd, ktorý bol predstavený pred viac ako 100 rokmi Georgom Westinghousom. Striedavý elektrický rozvod aktuálne postačuje elektroenergetickým požiadavkám, no napriek tomu by sme sa mali pravidelne zaoberať možnými efektívnejšími náhradami za striedavú elektrickú sústavu. Jednou z možností by mohla byť náhrada AC systému za DC systém. A práve tento článok sa bude zaoberať porovnaním AC a DC systému v ostrovnej prevádzke.

Hlavnou myšlienkou tohto článku je porovnanie AC a DC ostrovnej prevádzky nn siete. V súčasnosti vo svete prevláda AC prenos rovnako aj AC ostrovné prevádzky, ktoré sú napájané z obnoviteľných zdrojov, akými sú napr. fotovoltaika, zatiaľčo DC prevádzky sú vo svete pomerne veľmi málo využívané.

V prvej časti sa článok venuje predovšetkým teoretickými poznatkami striedavého a jednosmerného napätia a prúdu, pričom táto kapitola obsahuje aj príklady jednosmerného prenosu elektrickej energie vo svete. Obsahom druhej kapitoly je energetická bilancia spotrebiteľa elektrickej energie v AC a DC ostrovnej prevádzke, pričom v prípade tohto článku ide o spotrebu elektrickej energie v rámci domácnosti. Tretia kapitola sa zaoberá samotným návrhom AC a DC ostrovnej prevádzky v simulačnom programe MATLAB/Simulink, ale aj skutočným návrhom týchto ostrovných prevádzok. Ďalej nasleduje porovnanie týchto ostrovných prevádzok z hľadiska úbytkov napätia, počiatočných finančných nákladov a porovnania batériových systémov v závislosti od použitej ostrovnej prevádzky.

II. ÚVOD DO PROBLEMATIKY

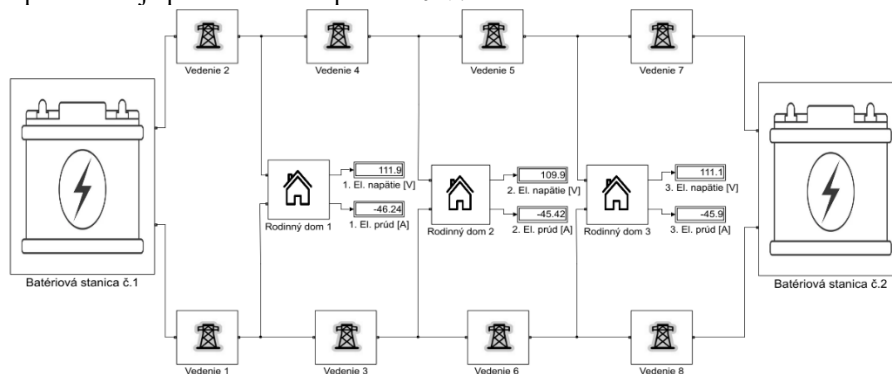
Tento článok otvára diskusiu, ktorej iniciovanie možno pripísať konkurencii medzi Thomasom Edisonom a Georgom Westinghousom, pokiaľ ide o relatívne výhody jednosmerných a striedavých distribučných systémov. Zatiaľčo koniec 19. storočia je známy najinovatívnejším inžinierstvom na svete, táto súťaž bola zjavne ovplyvnená aj pripisovaním jednosmerného alebo striedavého prúdu na použitie v kontroverznom novom vynáleze – elektrickom kresle. Odvtedy si striedavá distribúcia získala priazeň a stala sa neoddeliteľnou súčasťou našich systémov distribúcie elektrickej energie.

Obnoviteľné zdroje energie, ako napr. fotovoltaika, sú zdrojom jednosmerného prúdu. Aj veterná energia môže byť lepšie optimalizovaná, ak je aspoň časť kapacity veterného generátora prepojená prostredníctvom výkonovej elektronickej konverzie, ktorá často zahŕňa jednosmernú zbernicu. Konverzia z jednosmerného prúdu na striedavý a neskôr späť na jednosmerný, spôsobí

počas konverzií viaceré straty, ktorým by sa dalo predísť. Náš súčasný systém distribúcie striedavého prúdu nemá prakticky žiadne skladovanie energie, ale jednosmerná energia sa celkom jednoducho ukladá v batériách.

III. NÁVRH MODELOV AC A DC OSTROVNEJ PREVÁDZKY V PROGRAME MATLAB/SIMULINK

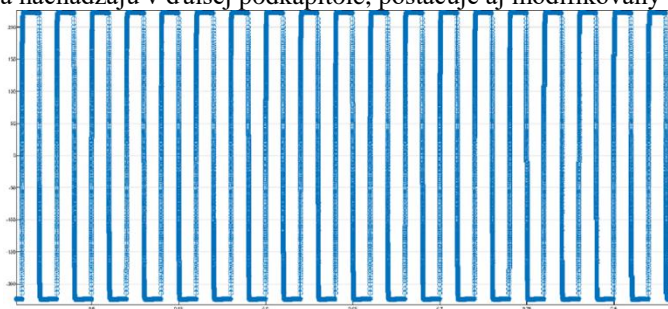
Ostrovná prevádzka pre účely tohto článku bola navrhnutá v programe MATLAB/Simulink. Samotný model ostrovnej prevádzky má poukázať na rozdiely medzi jednotlivými možnosťami distribúcie elektriny a to buď striedavým alebo jednosmerným napätím. Keďže ide o ostrovnú prevádzku (off-grid), tak odberatelia elektrickej energie sú napájaní len z batériových staníc. V modeli sú použité dve batériové stanice. Tieto stanice pozostávajú z batérií, ktorých konfigurácia je rovnaká pri oboch prevádzkach. Batérie sú nabíjané pomocou MPPT regulátora nabíjania, ktorý je napájaný z fotovoltaických panelov. Celkový výkon fotovoltaických panelov v oboch batériových staniách je 17,2 kWp. V modeli sú použité polykrystalické fotovoltaické panely, ktorých maximálny výkon na jeden panel je 320 W. V batériovej stanici pri AC ostrovnej prevádzke okrem fotovoltaických panelov, MPPT regulátora nabíjania a batérií je aj menič napätia, ktorý slúži na konvertovanie z DC na AC prenos. Prenos elektrickej energie je zabezpečený pomocou medeného káblového vedenia z batériových staníc k odberateľom elektrickej energie. Prevádzkové napätie AC ostrovnej prevádzky je 230 V o frekvencii 50 Hz. Pri DC prevádzke je prevádzkové napätie 110 V.



Obr. 1 Model DC ostrovnej prevádzky navrhnutý v programe MATLAB/Simulink

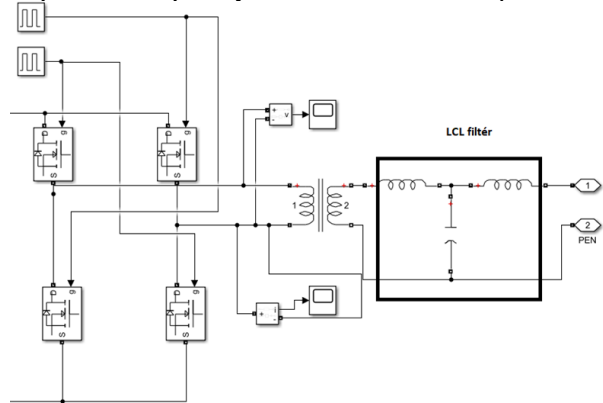
Keďže v programe SIMULINK je nasimulovaná ostrovná prevádzka a odberatelia nie sú napojení na distribučnú sieť, tak ako hlavný zdroj energie sa využíva fotovoltaický systém, ktorý následne dobíja batérie, z tohto dôvodu je celé elektrické napájanie založené na jednosmernom napätí. Preto je potrebné konvertovanie z jednosmerného na striedavé napätie (DC/AC). V praxi sa na to využívajú DC/AC meniče napätia. Úlohou DC/AC meničov je vytvoriť sínusový priebeh a upraviť hodnotu napätia v určitom rozsahu. Väčšinou výstupná hodnota napätia na meničoch je v rozsahu pri fázovom napätí od 220 do 230 V o frekvencii 50 Hz. Táto výstupná hodnota napätia nie je náhodná, keďže ide o normalizovanú hodnotu striedavého napätia podľa STN EN 50160, a rovnako aj väčšina spotrebičov, či už v domácnostiach alebo v priemysle, je napájaná takýmto napätím. Veľkou výhodou konvertovania z DC na AC, je výstupne napätie, ktoré je 230/400 V pri frekvencii 50 Hz, čo je hodnota napätia, pri ktorom pracuje väčšina zariadení, či už v domácnostiach alebo v priemysle. Z tohto hľadiska je preto výhodné využiť menič (striedač) v ostrovnej prevádzke.

Na prvom modeli bola simulovaná ostrovná prevádzka pri prevádzkovom striedavom napätí (AC). Hodnota napätia je 230 V. Pričom frekvencia napätia je 50 Hz. Pre konvertovanie z DC na AC slúži v prevádzke menič napätia (striedač). Na obrázku nižšie je možné vidieť modifikovaný sínusový priebeh, ktorý je výstupom z meniča napätia. V tomto prípade nebol použitý za transformátorom LCL filter, ktorý slúži na vyhladenie sínusového priebehu. Napriek tomu, pre merania, ktoré sa nachádzajú v ďalšej podkapitole, postačuje aj modifikovaný sínusový priebeh.

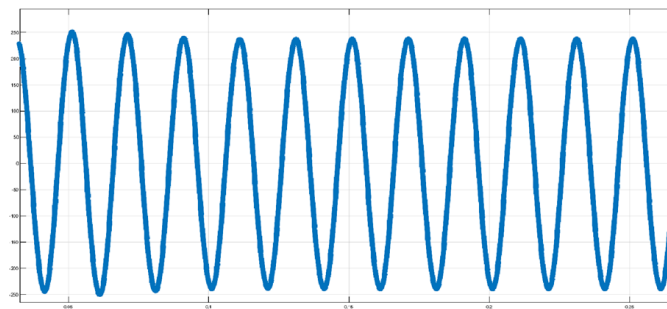


Obr. 2 Modifikovaný sínusový priebeh napätia a prúdu AC ostrovnej prevádzky

V prípade, pokiaľ sa použije LCL filter za transformátorom, v meniči napätia je možné výstupný sínusový priebeh vyhladiť do podoby, ako je na obrázku nižšie (Obr. 4). Nastavenie parametrov LCL filtra prebehlo empiricky. Z tohto dôvodu nebol pri meraní LCL filter použitý.

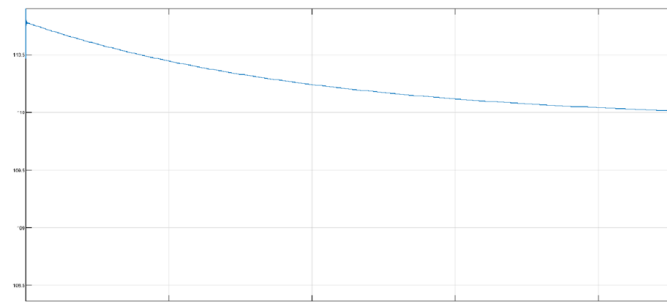


Obr. 3 Menič napätia s LCL filtrom



Obr. 4 Čistý sínusový priebeh napätia a prúdu v AC ostrovnej prevádzke

Druhý model nám simuluje ostrovnú prevádzku pri jednosmernom napätí (DC). Hodnota napätia je v rozsahu od 105–115 V DC.



Obr. 5 Časový priebeh DC napätia na DC ostrovnej prevádzke v programe MATLAB/Simulink

V tomto prípade nie je potrebné konvertovanie z DC napätia na AC, keďže hodnota výstupného napätia je 110 V DC, čo znamená, že bez DC/AC meniča sa systém podstatne zjednodušil a vďaka tomu je viac prehľadný. Ďalšou výhodou je zníženie elektrických strát oproti AC prevádzke, kedy bolo potrebné použiť menič; v tomto prípade nebol DC/AC menič použitý.

IV. POROVNANIE ÚBYTKOV NAPÄTÍ V AC A DC OSTROVNEJ PREVÁDZKE

Keď vodičom tečie elektrický prúd, tak vodič sa zahrieva. Rovnako môžu v ňom nastať chemické zmeny a v jeho okolí vzniká magnetické pole. Z tohto dôvodu je potrebné správne nadimenzovať prierezy vodičov použitých pri distribúcii elektrickej energie. Toto meranie má poukázať na rozdielne hodnoty elektrického prúdu tečúceho vodičmi pri distribúcii v jednosmernej a striedavej ostrovnej prevádzke.

Hodnota elektrického prúdu pri AC, tak aj pri DC ostrovnej prevádzke, závisí od druhu záťaže, kedy pri AC prevádzke ide o kombináciu odporovej, induktívnej a kapacitnej záťaže a v prípade DC ide o kombináciu odporovej a induktívnej záťaže o výkone $3 \times 5\,000$ W, keďže v modeli sú 3 odberné miesta.

Pri meraní základných veličín, akými sú elektrické napätie a prúd pri AC a DC ostrovnej prevádzke, bol pre prírodné vodiče k odberným miestam zvolený prierez 35 mm^2 . Meranie prebiehalo v programe MATLAB/Simulink, kedy sa menil odoberaný výkon zo siete. Spolu prebehlo 7 meraní, a to od 3 kW až po 18 kW (Tab. 1).

Tab. 1 Hodnoty odoberaného výkonu v priebehu meraní

Číslo merania	Prierez prírodných vodičov [mm ²]	Rodinný dom č.1 odoberaný výkon [kW]	Rodinný dom č.2 odoberaný výkon [kW]	Rodinný dom č.3 odoberaný výkon [kW]	Celkový odoberaný výkon [kW]
1.	35	0	0	0	0
2.	35	1	1	1	3
3.	35	2	2	2	6
4.	35	3	3	3	9
5.	35	4	4	4	12
6.	35	5	5	5	15
7.	35	6	6	6	18

Tab. 2 Namerané hodnoty prúdu a napätia pri zmene odoberaného výkonu pre AC prevádzku

AC prevádzka				
Meranie č.1	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	236,3	236,3	236,3	236,30
El. prúd v mieste prípojky [A]	0	0	0	0,00
Meranie č.2	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	232,3	232,1	232,3	232,23
El. prúd v mieste prípojky [A]	4,4	4,473	4,47	4,45
Meranie č.3	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	226,7	226,3	226,7	226,57
El. prúd v mieste prípojky [A]	8,57	8,59	8,61	8,59
Meranie č.4	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	220,5	219,8	220,4	220,23
El. prúd v mieste prípojky [A]	12,5	12,4	12,43	12,44
Meranie č.5	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	216,1	215,3	216,1	215,83
El. prúd v mieste prípojky [A]	16,34	16,24	16,3	16,29
Meranie č.6	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	211,4	210,4	211,3	211,03
El. prúd v mieste prípojky [A]	19,98	19,82	19,91	19,90
Meranie č.7	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	208,4	207,1	208,2	207,90
El. prúd v mieste prípojky [A]	23,64	23,56	23,68	23,63

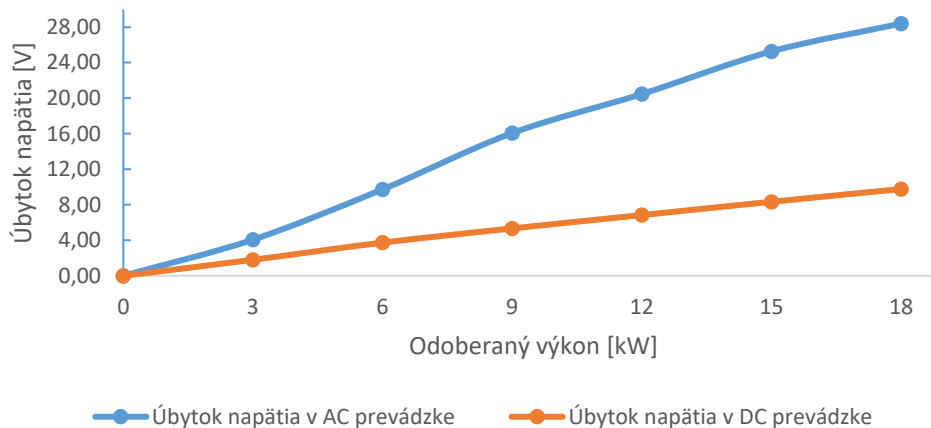
Tab. 3 Namerané hodnoty prúdu a napätia pri zmene odoberaného výkonu pre DC prevádzku

DC prevádzka				
Meranie č.1	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	116,3	116,3	116,3	116,30
El. prúd v mieste prípojky [A]	0	0	0	0,00
Meranie č.2	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	114,6	114,2	114,7	114,50
El. prúd v mieste prípojky [A]	9,47	7,43	9,47	8,79
Meranie č.3	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	112,9	111,9	112,9	112,57
El. prúd v mieste prípojky [A]	18,66	18,5	18,66	18,61
Meranie č.4	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	111,4	110,1	111,4	110,97
El. prúd v mieste prípojky [A]	27,65	27,31	27,65	27,54
Meranie č.5	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	110,1	108,3	110	109,47
El. prúd v mieste prípojky [A]	36,38	35,79	36,38	36,18
Meranie č.6	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	108,7	106,5	108,7	107,97
El. prúd v mieste prípojky [A]	44,93	44,01	44,91	44,62
Meranie č.7	Rodinný dom č.1	Rodinný dom č.2	Rodinný dom č.3	Priemerná hodnota
El. napätie v mieste prípojky [V]	107,4	104,8	107,4	106,53
El. prúd v mieste prípojky [A]	53,28	51,98	53,26	52,84

Tab. 4 Úbytky napätia v závislosti od zmeny odoberaného výkonu

Odoberaný výkon [kW]	Úbytok napätia pri AC prevádzke [V]	Úbytok napätia pri DC prevádzke [V]
0	0,00	0,00
3	4,07	1,80
6	9,73	3,73
9	16,07	5,33
12	20,47	6,83
15	25,27	8,33
18	28,40	9,77

Úbytok napätia v závislosti od zmeny odoberaného výkonu



Obr. 6 Grafická závislosť úbytkov napätí od zmeny odoberaného výkonu

Z meraní vyplýva, že úbytky napätia, vzniknuté zmenou odoberaného výkonu záťaže pri AC ostrovej prevádzke, sú väčšie než pri DC ostrovej prevádzke. Je to spôsobené predovšetkým vyššou hodnotou prevádzkového napätia pri AC prevádzke, ktorá pracuje pri napätí 230 V, zatiaľ čo pri DC prevádzke je napätie 110 V. To znamená, že AC prevádzka o vyššom prevádzkovom napätí je viac náchylnejšia na úbytok napätia spôsobený zmenou odoberaného výkonu záťaže.

V. ZÁVER

Tento článok sa zaoberal porovnaním AC a DC ostrovej prevádzky, pričom pri porovnávaní týchto dvoch prevádzok sa použil simulačný program MATLAB/Simulink, v ktorom sa navrhnuté ostrovné prevádzky odsimulovali, čím bolo možné ich porovnať. Pri prvom meraní sa porovnávali tieto dve prevádzky z hľadiska úbytkov napätia v závislosti od zmeny odoberaného výkonu. Pri tomto meraní bol pri oboch prevádzkach použitý rovnaký prierez prívodných vodičov. Celkovo prebehlo 7 meraní, kedy sa výkon menil v rozsahu od 0 do 18 kW. Z merania vyplynulo, že AC prevádzka o vyššom prevádzkovom napätí je viac náchylnejšia na úbytok napätia spôsobený zmenou odoberaného výkonu záťaže.

Väčšina nevýhod DC prevádzky, ktoré vyplývajú z výsledkov, uvedených v tomto článku, je nižšie prevádzkové napätie v porovnaní s AC prevádzkou, a to aj pri rovnakom prenášanom výkone. To znamená, že v DC prevádzke je podstatne vyšší prúd a aj straty, ktoré sa v prípade prenosu elektrickej energie dajú vyriešiť len väčším prierezom vodičov, čo vplyva na cenu takejto prevádzky. Ďalšou nevýhodou DC prevádzky je vznik elektrického oblúka pri spínaní, napr. vypínačov, z tohto dôvodu je potrebné v prevádzkach používať podstatne nižšiu hodnotu napätia než je to pri AC napätí. No veľkou výhodou DC prevádzky je jednoduchosť, vďaka ktorej je systém menej náročný na údržbu a je viac spoľahlivý a preto do budúcnosti považujeme DC ostrovné prevádzky využívané napr. v domácnostiach, ako jednu z možností pri výbere vhodného napájacieho a rozvodného systému.

POĎAKOVANIE

Tento príspevok vznikol vďaka Agentúre na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0576 a APVV-21-0312, ako aj Vedeckou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu SR a SAV, podporená grantom VEGA 1/0757/21.

ZOZNAM POUŽITEJ LITERATÚRY

- [1] Shopdelta, „Jednosmerný prúd (DC)“, [online], [cit. 2023-02-04], dostupné na internete: <https://shopdelta.eu/jednosmerny-prud-dc-19-aid761.html>
- [2] TME, „Elektrický prúd, definície a typy“, [online], [cit. 2023-03-04], dostupné na internete: <https://www.tme.eu/sk/news/events/page/51506/elektricky-prud-definicia-a-typy/>
- [3] M. Andraš, „Porovnanie AC a DC ostrovej prevádzky nn siete“. Diplomová práca. Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2023.

Design and Implementation of a CASE Tool for Reverse Engineering

Martin SEDLAČKO, Branislav MADOSŤ

Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

`martin.sedlacko@student.tuke.sk`, `branislav.mados@tuke.sk`

Abstract – This paper deals with the development of a desktop application in the JavaFX programming language for the analysis and visualization of C programming language source code using UML activity diagram generation. The paper starts with an introduction providing a brief overview of the topic. It then discusses reverse engineering, unified modeling language, and CASE tools that are utilized in the development process. The paper further delves into the application design and implementation, explaining how the parsing algorithm works and how the final diagram is created. Finally, the paper concludes by summarizing the main findings and highlighting the significance of the developed application in facilitating effective code review.

Keywords – CASE tools, Reverse Engineering, Java, SrcML, XML

I. INTRODUCTION

Programming is a complex task that involves writing lines of code to create a working software application. It requires not only a deep understanding of the programming language but also a clear understanding of the problem that needs to be solved. One important aspect of programming is to visualize how the code will execute, which can help identify potential issues and improve the overall design. In this context, Unified Modeling Language (UML) Activity diagrams are a popular tool for modeling the behavior of a software system. They allow developers to visualize the flow of control within the system, showing how activities are performed and how they interact with each other. However, creating UML Activity diagrams manually can be a time-consuming and error-prone task, especially for large and complex codebases. To address this challenge, a CASE tool has been developed within our research that takes C programming language source code as input, decomposes it through reverse engineering, and automatically generates UML Activity diagrams of how the code will be executed. This application provides the way how developers analyze and understand software systems, allowing them to gain deeper insights into the behavior of the code without the need for manual analysis.

II. REVERSE ENGINEERING

According to [1], reverse engineering can roughly be defined as taking a final product, dissecting it to understand its functionality, obtain design and other useful information. Software Reverse Engineering analyzes the system's code, documentation, and behavior to create system abstractions and design information. Helps to understand internal system complexities while having access only to external behavior. Practice with reverse engineering techniques improves the ability to understand a given system quickly and efficiently. This skill has serious economic effects when software engineers are called upon to solve problems pertaining to software evolution. As fifty to ninety percent of the time, depending upon the system under consideration is spent on program understanding. The reverse engineering process enables you to determine how a part was designed so that you can recreate it [2].

III. UNIFIED MODELING LANGUAGE

UML, short for Unified Modeling Language, is a standardized modeling language consisting of an integrated set of diagrams, developed to help system and software developers for specifying, visualizing, constructing, and documenting the artifacts of software systems, as well as for business modeling and other non-software systems [3]. As the systems developed were complicated to understand, it led to the design and analysis problems which were faced after the deployment of the system. It was difficult to explain the system to others. As soon as the UML was introduced, many game-changing experiments and approaches were made for simplifying such difficult tasks of analyzing the system. [4] The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems. The UML is a very important part of developing object oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects. Using the UML helps project teams communicate, explore potential designs, and validate the architectural design of the software. UML consists of 14 diagrams which are divided into 2 groups. All of the 14 diagrams divided into groups are shown in the picture 1.

- Behavior diagrams show the dynamic behavior of the objects in a system, which can be described as a series of changes to the system over time.
- Structure diagrams show the static structure of the system and its parts on different abstraction and implementation levels and how they are related to each other. The elements in a structure diagram represent the meaningful concepts of a system, and may include abstract, real world and implementation concepts [3].

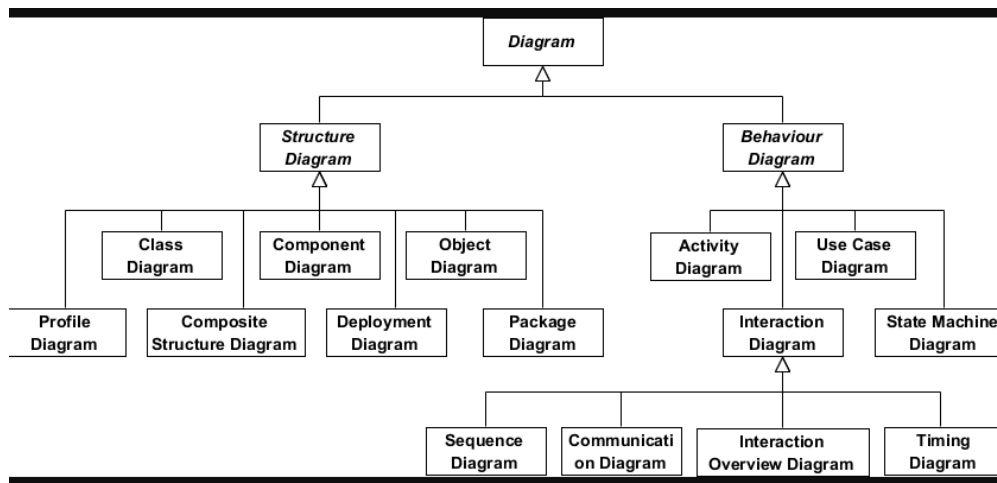


Fig. 1 UML diagrams divided into two groups [3]

IV. CASE TOOLS

CASE tools are tools that provide automated assistance in software development [5]. CASE tools are set of application programs to automate software development lifecycle activities and are used by managers in a project, engineers and analysts to build a software system is called CASE tools and the software development cycle stages can be simplified using several tools such as design, analysis, project management, database management, documentation, etc. and the use of these tools speeds up the project development to obtain desired results [6]. Most classifications of CASE tools start by considering whether the tool is upper CASE, lower CASE or integrated CASE. An upper CASE tool provides support for the early stages in the systems development life cycle such as requirements analysis and design. A lower CASE tool provides support for the later stages in the life cycle such as code generation and testing. Integrated CASE tools support both the early and later stage [7].

V. APPLICATION DESIGN

Before starting the process of implementing the software application, it was necessary to develop a detailed concept that would define the structure, design, and functionality of the code of the software solution. For this reason, a modular plan was compiled, which was followed in each stage of the development.

A. Component diagram

The application has been designed into five modules as can be seen on the picture of UML component diagram 2.

InputModule is responsible for reading the input source code in the C programming language. If the user does not insert any file, a warning message will appear in a pop-up window. The module provides the input file of the SrcMLIntegrationModule.

SrcMLIntegrationModule ensures integration with the SrcML library. The module retrieves the input source code from the InputModule and uses the SrcML library, which analyzes the input file and generates an output in the form of an XML file. The XML file is then provided to the ReadXMLDomParserModule.

ReadXMLDomParserModule allows reading the XML file, extracts all important information from it, and then creates a Java collection, which is provided to the DataCleanerModule.

DataCleanerModule goes through the collection and checks the types of records. If two or more variable declarations follow each other, these records are merged into one record. The resulting modified Java collection is provided to the last module, MainSceneModule.

MainSceneModule iterates through the modified Java collection, checks its types, and then calls the appropriate method, which creates the suitable UI element. The elements are then inserted and displayed in the resulting diagram.

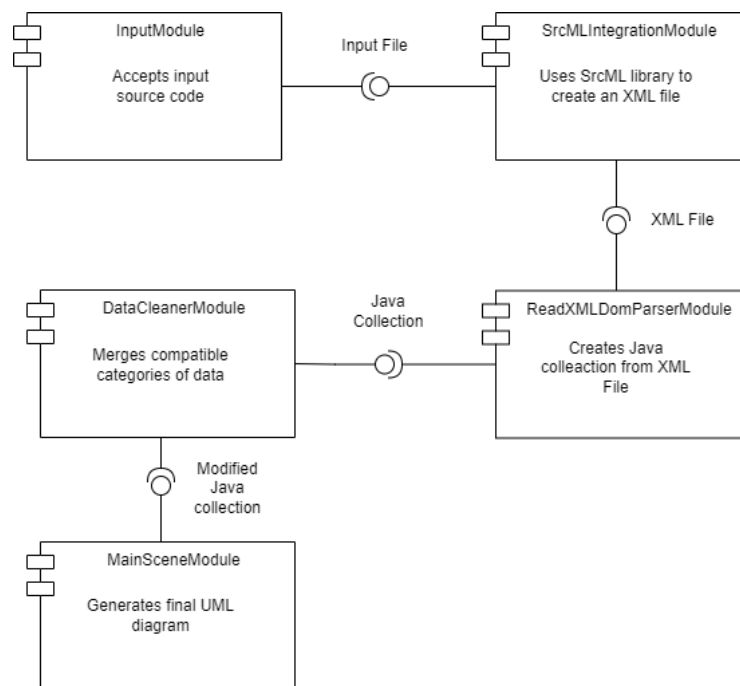


Fig. 2 UML Component Diagram of application.

B. Graphical User Interface

The application UI 3 consists of a navigation panel with three buttons, a variable sidebar toggle, an input file insertion button, and a result diagram generation button. The variable sidebar toggle allows the user to show or hide a sidebar displaying variables relevant to the application. The input file insertion button allows the user to upload a file to be used in generating the result diagram. Finally, the result diagram generation button initiates the creation of the diagram on the main screen.

The main screen of the application is where the result diagram is displayed. The diagram is a visual representation of the data contained in the uploaded file. The layout and design of the diagram may vary depending on the specific application, but the main purpose is to provide the user with a clear and concise representation of the data.

Overall, the UI is designed to be simple and intuitive, with clear labeling and easily accessible buttons. The navigation panel is located in a prominent location, allowing the user to quickly access the key features of the application. The main screen is designed to focus the user's attention on the result diagram, which is the main output of the application. By providing a clear and concise representation of the data, the UI enables the user to quickly understand the information presented and make informed decisions based on the results.

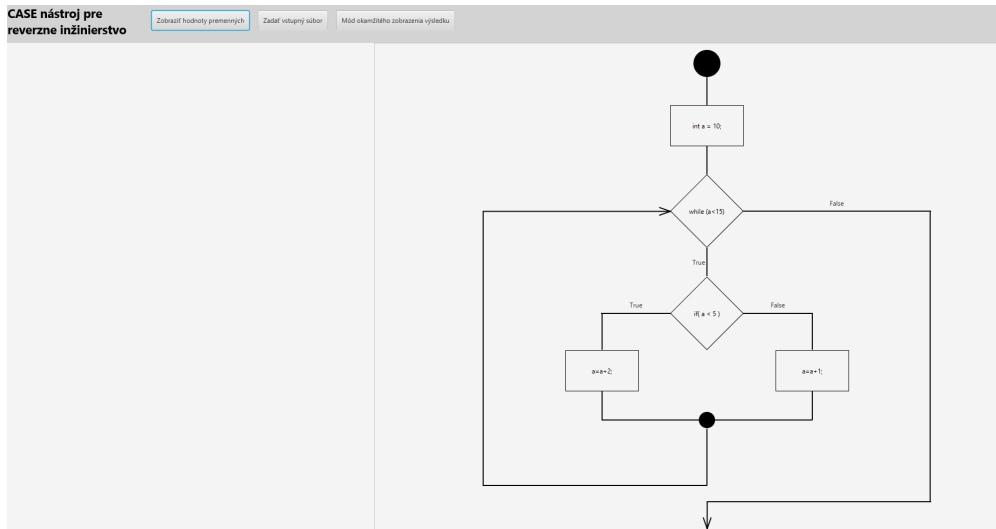


Fig. 3 Graphical User Interface of an application.

VI. IMPLEMENTATION

In software development, understanding the structure and logic of the code is crucial for effective code review, maintenance, and documentation. Overall, this section provides a detailed understanding of the different components involved in developing an application for analyzing and visualizing code.

A. Parsing input source code

To simplify working with the source code in the C programming language, it is advisable to divide it into a more readable and logical structure. SrcML tool, which parses the source code into an XML structure was chosen to parse input source code. The aforementioned structure is more suitable for navigating through individual code blocks, processing input, and displaying results.

The output XML file 1 contains the code divided into several blocks, such as the declaration block, expression block, condition block, and loop block. Thanks to this division, we can work with the code more efficiently and easily find and extract important parts of the code.

Listing 1 Example-srcml-output.xml

```

...
<function>
  <type>
    <name>int</name>
  </type>
  <name>main</name>
  <parameter_list>()</parameter_list>
  <block>
    {
      <decl_stmt>
        ...
      </decl_stmt>
      <expr_stmt>
        ...
      </expr_stmt>
    }
  </block>
</function>
</unit>

```

B. Algorithm to process XML file

The developed algorithm works by searching for parent tags in the XML output file, such as `<decl_stmt>`, `<expr_stmt>`, `<if>`, and so on. The tag is then sent as an argument to function 2, which iterates through its child tags containing all the necessary information. The algorithm recursively goes through these tags and searches for its child tags until it finds a tag that contains a value. The value is appended to the resulting string, which is inserted into a List containing the processed code.

Listing 2 setResultValue.java

```
private void setResultValue(final StringBuilder result ,
                           final Node node ,
                           final String delimiter) {
    if (node.hasChildNodes()) {
        for (int index = 0; index < node
             .getChildNodes().getLength(); index++)
        {
            this.setResultValue(result ,
                                node.getChildNodes().item(index) ,
                                delimiter);
        }
    } else {
        if (!node.getNodeValue().trim().isEmpty()) {
            result.append(node.getNodeValue().trim())
                .append(delimiter);
        }
    }
}
```

C. Java Collection consisting of parsed code

The processed code is saved in a Java collection, specifically in the List class, which contains elements of the NodeData class. The NodeData class 3 contains variables key and value. Key represents the type of the element, such as `decl_stmt`, `expr_stmt`, `if`, etc.

Listing 3 NodeData.java

```
public class NodeData {
    private String key;
    private String value;
}
```

D. Displaying the resulting diagram

The obtained collection containing the processed code is modified in the first step. The modification is carried out by iterating through the collection and determining the previous and current type of the element. If the types are identical and they are variable declarations, they are merged into a single resulting shared element. After modifying the data, it is necessary to iterate through the modified collection, check the type, and based on the type, call the appropriate method that creates a UI element. The created UI element is then inserted into the resulting diagram. The shortened source code 4 shows the method for creating the diagram.

Listing 4 MainScene.java

```
public GraphVBox drawDiagram(...) {
    VBox vbox = new VBox();
    List<NodeData> content = getModifiedContent(parsedCode);
    content.forEach(value -> {
        Node node;
        if (value.getKey().equals(
            NodeType.CONDITION.getStringValue())) {
            ...
        }
    });
}
```

```
        node = createConditionElement();
    } else if (value.getKey().equals(
        NodeType.FOR.getStringValue())) {
        ...
        node = createForElement();
    } else if (value.getKey().equals(
        NodeType.WHILE.getStringValue())) {
        ...
        node = createWhileElement();
    } else {
        ...
        node = createElement();
    }
    vbox.getChildren().add(node);
});
return new GraphVBox(vbox, parentHeightAtomic.get());
}
```

VII. CONCLUSION

This paper presents the development and implementation of a desktop application in the JavaFX programming language designed for analysis and visualization of the execution flow of C programming language source code using UML activity diagram generation. The primary goal of the application is to provide users with a clear and detailed understanding of the basic structure of the code, facilitate effective code review, and assist in creating accurate software documentation. The application is developed as a desktop solution using the JavaFX framework, a popular choice for creating rich, cross-platform desktop applications. JavaFX offers a wide range of features, making it an ideal choice for creating modern, sophisticated user interfaces. The decision to develop a desktop application was based on several factors, including excellent performance, offline capabilities, improved security, and the ability to provide users with a customized experience. The solution provides an effective tool for better understanding source code, its documentation, and communication, ultimately contributing to the improvement of software development processes.

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: 313011W988), co-financed by the ERDF and by the project KEGA 002TUKE-4/2021 Implementation of Modern Methods and Education Forms in the Area of Cybersecurity towards Requirements of Labour Market

REFERENCES

- [1] M. R. Ali, "Why teach reverse engineering?" *SIGSOFT Softw. Eng. Notes*, vol. 30, no. 4, p. 1–4, 7 2005. [Online]. Available: <https://doi.org/10.1145/1082983.1083004>
- [2] B. Hess, "What is reverse engineering and how does it work?" Online, 9 2019. [Online]. Available: <https://astromachineworks.com/what-is-reverse-engineering/>
- [3] Visual Paradigm, "What is uml?" Online, 2 2022. [Online]. Available: <https://www.visual-paradigm.com/guide/uml-unified-modeling-language/what-is-uml/>
- [4] A. Walker, "Uml diagrams: History, types, characteristics, versions, tools," *Guru99*, 2 2023. [Online]. Available: <https://www.guru99.com/uml-diagrams.html#3>
- [5] R. D. Banker and R. J. Kauffman, "Reuse and productivity in integrated computer-aided software engineering: An empirical study," *MIS Quarterly*, vol. 15, no. 3, pp. 375–401, 1991. [Online]. Available: <http://www.jstor.org/stable/249649>
- [6] S. Shivakumar, "Introduction to case tools," Online, 4 2020. [Online]. Available: <https://www.educba.com/case-tools/>
- [7] N. L. Chervany and D. Lending, "Case tools: Understanding the reasons for non-use," *SIGCPR Comput. Pers.*, vol. 19, no. 2, p. 13–26, apr 1998. [Online]. Available: <https://doi.org/10.1145/292349.292353>

Riadenie pohonu s asynchrónnym motorom na báze inverzného fuzzy modelu

¹Marek FEDOR, ²Daniela PERDUKOVÁ

^{1,2} Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹marek.fedor.4@student.tuke.sk, ²daniela.perdukova@tuke.sk

Abstrakt — Článok sa zaoberá zostavením zjednodušeného inverzného fuzzy modelu asynchrónneho motora s využitím odhadu kritických parametrov Klossovho vzťahu. Vytvorenie modelu si vyžaduje meranie výstupných veličín motora na skokové zmeny jeho vstupov, pričom nie je potrebná žiadna apriórna znalosť jeho parametrov. Získaný fuzzy model asynchrónneho motora je možné využiť pre linearizáciu jeho momentovej slučky a tým aj pre návrh jeho nadradených regulačných slučiek (rýchlostnej, polohovej). Takýmto postupom sa odstráni potreba transformácií veličín motora, čo je charakteristické napríklad pre vektorové riadenie.

Kľúčové slová — asynchrónny motor, Klossov vzťah, inverzný fuzzy model, riadenie nelineárnych systémov

I. INTRODUCTION

Zatiaľ čo riadenie lineárnych systémov je v súčasnosti detailne prepracované a vieme ho analyticky navrhnuť prakticky pre všetky reálne praktické požiadavky praxe, riadenie nelineárnych systémov (vzhladom na ich negatívnu definíciu – neplatí pre nich princíp superpozície) sa rieši väčšinou len pre určitú obmedzenú menšiu alebo väčšiu skupinu nelineárnych systémov a neexistujú analytické univerzálne metódy návrhu ich riadenia. Pre riešenie širokej triedy takýchto úloh sa v súčasnosti ukazujú veľmi vhodné metódy umelej inteligencie, medzi ktoré patria aj fuzzy systémy.

Fuzzy systémy predstavujú úspešný prístup k riadeniu mnohých zložitých nelineárnych systémov. Riadenie s fuzzy logikou možno považovať za jednu z najúspešnejších aplikácií fuzzy systémov, ktoré sa v mnohých prípadoch stali alternatívou k použitiu konvenčných techník riadenia v rôznych oblastiach vrátane: energetických systémov [1-2], mechanicko-robotických systémov [3-5], automobilových systémov [6-7] a systémov výkonovej elektroniky [8-9].

V tomto príspevku je ukázaná metodika návrhu regulácie nelineárnej sústavy pomocou jej linearizácie metódou inverzného fuzzy modelu a jej aplikácia na jednoduché riadenie pohonu s asynchrónnym motorom.

II. MATEMATICKÝ MODEL ASYNCHRÓNNEHO MOTORA

Asynchrónny motor (AM) je z hľadiska riadenia sústava piatich nelineárnych diferenciálnych rovníc prvého rádu. Ak zvolíme za stavové veličiny prúd statora a tok rotora, potom môžeme asynchrónny motor nakrátko popísať nasledujúcimi rovnicami [10]:

$$\begin{bmatrix} \frac{di_{1x}}{dt} \\ \frac{di_{1y}}{dt} \\ \frac{d\psi_{2x}}{dt} \\ \frac{d\psi_{2y}}{dt} \end{bmatrix} = \begin{bmatrix} -\omega_0 & \omega_1 & -K_{12}\omega_g & -K_{12}\omega_m n_p \\ -\omega_1 & \omega_0 & K_{12}\omega_m n_p & -K_{12}\omega_g \\ M\omega_g & 0 & -\omega_g & \omega_2 \\ 0 & M\omega_g & -\omega_2 & \omega_g \end{bmatrix} \begin{bmatrix} i_{1x} \\ i_{1y} \\ \psi_{2x} \\ \psi_{2y} \end{bmatrix} + \begin{bmatrix} K_{11} & 0 \\ 0 & K_{11} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{1x} \\ u_{1y} \end{bmatrix} \quad (1)$$

$$n_p \frac{M}{L_2} (\psi_{2x} i_{1y} - \psi_{2y} i_{1x}) - T_L = J \frac{d\omega_m}{dt} \quad (2)$$

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 $\{xy\}$, ktorá rotuje voči statoru uhlovou rýchlosťou statorového poľa ω_1 . Jednotlivé parametre v rovnici (1) vieme určiť z parametrov motora prepočítaním podľa nasledujúcich vzťahov:

$$K_{11} = \frac{3}{2} \left(L_{s1} + \frac{L_{s2}L_h}{L_{s2} + L_h} \right)^{-1} \quad (3)$$

$$K_{12} = -\frac{3}{2} \left(L_{s1} + L_{s2} + \frac{L_{s1}L_{s2}}{L_h} \right)^{-1} \quad (4)$$

$$\omega_0 = K_{11} \left[R_1 + \left(\frac{M}{L_2} \right)^2 R_2 \right] \quad (5)$$

$$M = \frac{2}{3} L_h \quad (6)$$

$$\omega_g = \frac{R_2}{L_2} \quad (7)$$

$$L_2 = \frac{2}{3} (L_{s2} + L_h) \quad (8)$$

Význam a hodnoty jednotlivých parametrov AM sú uvedené v Tabuľke č.1.

Tabuľka 1 Parametre asynchrónneho motora

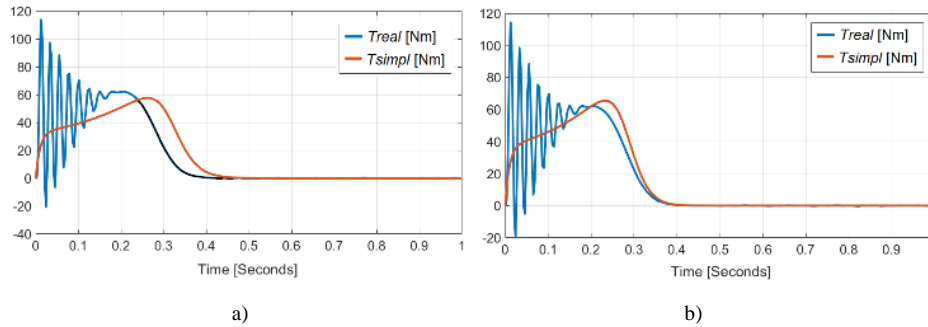
Symbol	Veličina	Hodnota/rozměr
P_N	nominálny výkon	3 kW
U_{IN}	nominálne napätie	220 V
I_{IN}	nominálny prúd	6.9 A
T_N	nominálny moment	20 Nm
n_N	nominálne otáčky	1430 rev/min
J	moment zotrvačnosti	0,1 kgm ²
R_1	odpor fázy statoru	1.8 Ω
R_2	odpor fázy rotora	1.85 Ω
L_h	hlavná indukčnosť	0.202 H
L_2	indukčnosť definovaná rovnicou (8)	0.14 H
$L_{s1}=L_{s2}$	indukčnosť	0.0086 H
M	vzájomná indukčnosť definovaná rovnicou (6)	0.13 H
n_p	počet párov pólov	2
K_{11}	parameter definovaný rovnicou (3)	59.35 H ⁻¹
K_{12}	parameter definovaný rovnicou (4)	-56.93 H ⁻¹
ω_0	parameter definovaný rovnicou (5)	207.9 s ⁻¹
ω_g	časová konštanta rotorového vinutia – rovnica (7)	13.18 s ⁻¹
ω_m	mechanická uhlová rýchlosť	[rad/s]
ω_1	uhlová rýchlosť vektora statorového napätia	[rad/s]
ω_2	frekvencia sklzu ($\omega_1 - \omega_m$)	[rad/s]
T_L	moment záťaže	[Nm]
m_1	počet fáz statora	3

Bloková schéma asynchrónneho motora, zodpovedajúca rovniciam (1-2) s parametrami upravenými podľa uvedených vzťahov (3-8) je na Obr. 1.

Presnosť vyššie uvedenej aproximácie AM je daná tromi parametrami jeho zjednodušeného modelu: maximálnym momentom $T_{max}(\omega_1)$, maximálnym sklzom $s_{max}(\omega_1)$ a náhradnou časovou konštantou elektromagnetických prechodových dejov $t_r(\omega_1)$. Ak zvolíme za kvantitatívne kritérium náhrady integrál strednej kvadratickej odchýlky medzi zjednodušeným a úplným modelom AM, tak pre teoreticky vypočítané parametre motora podľa rovníc (10) až (12) dostaneme pre $\omega_1=314$ rad/s hodnoty parametrov zjednodušeného modelu: $T_{max}=58$ Nm, $s_{max}=0.3185$ a $t_r=0.01$ s a hodnotu kritériálnej funkcie $J=214$ podľa:

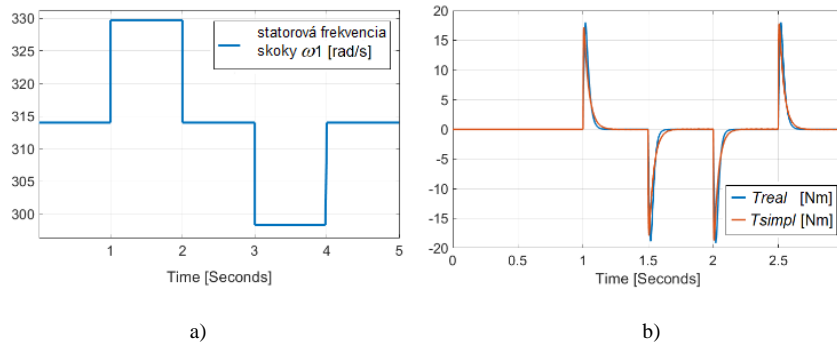
$$J = \int (T_{real} - T_{simple})^2 dt \quad (15)$$

Pri optimálnych parametroch momentom $T_{max}=58$ Nm, $s_{max}=0.32$ a $t_r=0.002$ s bude hodnota kritériálnej funkcie $J=144$. Obrázok 4a a Obr.4b ukazujú priebeh skutočného a aproximovaného momentu AM pre uvedené prípady.



Obr.4 Optimálna aproximácia parametrov AM a) vypočítané parametre b) optimalizované parametre

Porovnanie momentových priebehov úplného a zjednodušeného modelu AM pri skokoch statorovej frekvencie ω_1 v okolí nominálnej rýchlosti motora je ukázané na Obr.5. Z obrázka je zrejmé, že Klossov vzťah a zjednodušená bloková schéma z Obr.3 aproximuje AM veľmi dobre v okolí pracovného bodu ω_1 , pre ktorý boli parametre T_{max} , s_{max} , t_r optimalizované.



Obr. 5 a) Skoky ω_1 b) Porovnanie momentovej dynamiky úplného a zjednodušeného modelu AM

III. LINEARIZÁCIA NELINEÁRNEJ MOMENTOVEJ CHARAKTERISTIKY ASYNCHRÓNNEHO MOTORA

Nelinearita momentovej charakteristiky je približne daná Klossovým vzťahom (9), čo je ukázané na Obr.3. Linearizovať nejakú statickú funkciu vo všeobecnosti je možné jej súčinnom s jej inverznou funkciou (Obr.6). Moment motora je podľa tohto vzťahu funkciou rozdielu uhlovej rýchlosti statorového rotujúceho magnetického poľa a mechanickej uhlovej rýchlosti rotora ω_2 , preto platí:

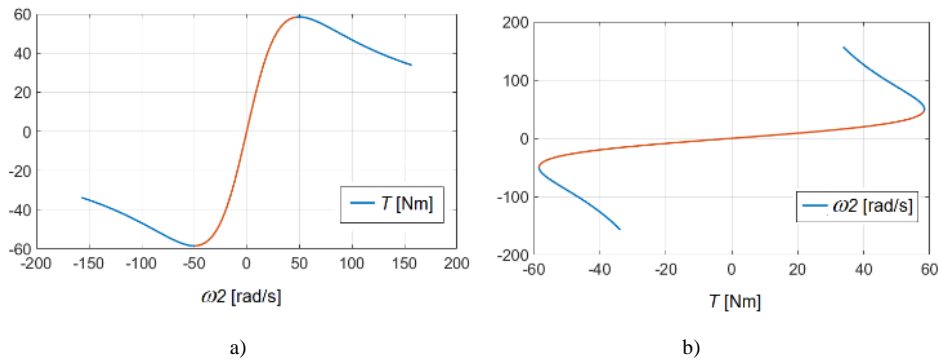
$$\omega_2 = \omega_1 - \omega_m = s \omega_1 \quad (16)$$

$$\omega_{2max} = s_{max} \omega_1 \quad (17)$$

$$T = \frac{2M_{max} s_{max} \omega_1 \omega_2}{s_{max}^2 \omega_1^2 + \omega_2^2} \quad (18)$$

$$T = \frac{2M_{max} \omega_{2max} \omega_2}{\omega_{2max}^2 + \omega_2^2} \quad (19)$$

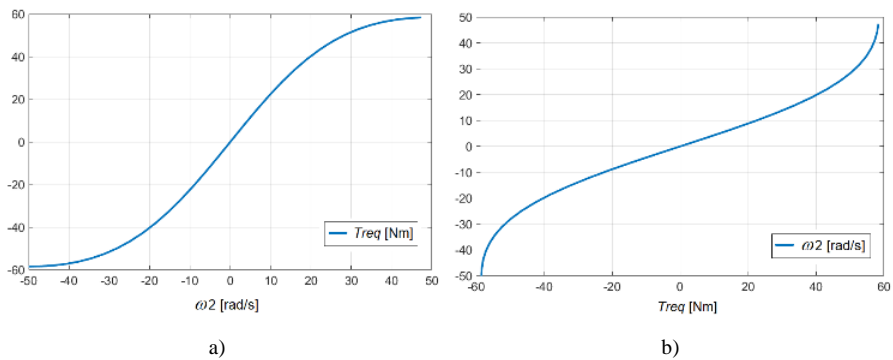
Graficky je táto nelineárna funkcia a jej inverzná funkcia znázornená na Obr.6.



Obr. 6 Statická momentová charakteristika AM
 a) pôvodná funkcia b) jej inverzia

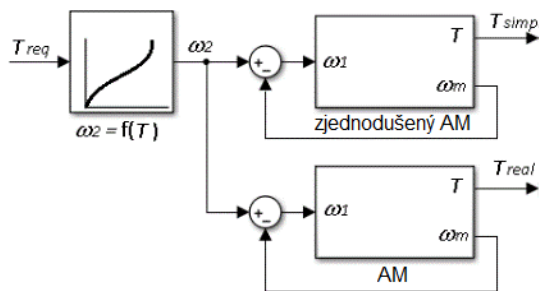
Z Obr.6b vidíme, že inverzia tejto konkrétnej nelinearity nie je jednoznačná, napr. pre moment motora $T=50$ Nm môže byť sklzová uhlová rýchlosť $\omega_2=48$ rad/s alebo aj $\omega_2=234$ rad/s. Preto bude pri linearizácii potrebné pracovať iba v jednoznačnej časti nelinearity. To znamená, že bude potrebné napr. obmedziť rozsah ω_2 na interval $\langle 0, \omega_{2max} \rangle$. Na Obr.6 je takto vymedzená pracovná oblasť pre ω_2 (a tým aj pre žiadaný moment motora pri danej rýchlosti ω_1) vyznačená červenou farbou.

Analytická inverzia vzťahu (19) je pomerne náročná a je preto výhodné realizovať grafickú inverziu podľa Obr.6b. Jednoznačná časť tejto inverzie je zobrazená na Obr. 7.



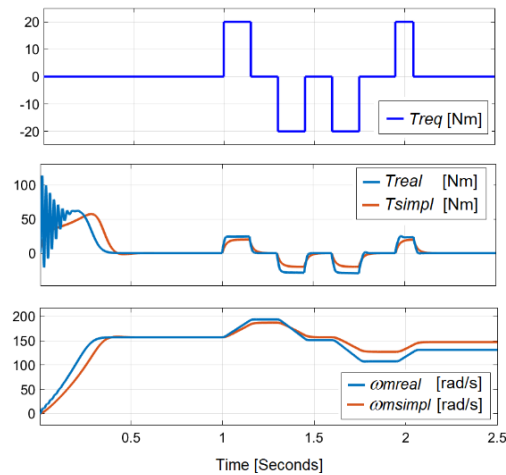
Obr. 7 a) Statická momentová charakteristika AM b) Jej inverzia v stabilnej časti Klossovho vzťahu

Na Obr.8 je uvedená linearizačná bloková schéma pre zjednodušený model AM, ktorá je porovnávaná s pripojením na úplný model AM podľa Obr.1.



Obr. 8 Linearizovaná momentová regulačná slučka AM

Obrázok (9) ukazuje, že pre zjednodušený model AM, použitý pri linearizácii, bol dosiahnutý prakticky ideálny lineárny priebeh momentovej regulačnej slučky s dynamikou zotrvačnosti 1.rádu.



Obr. 9 Porovnanie momentovej dynamiky úplného a zjednodušeného modelu AM pri skokoch T_{req}

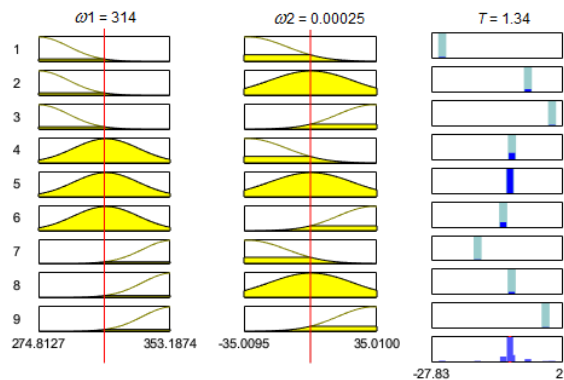
IV. FUZZY LINEARIZÁCIA MOMENTOVEJ CHARAKTERISTIKY AM

Ako bolo uvedené vyššie, momentovú regulačnú slučku AM je možné linearizovať pomocou inverzného fuzzy systému jeho statickej charakteristiky. Táto linearizácia má niektoré nevýhody:

- Problém je vyjadriť analyticky inverznú charakteristiku AM (inverzný Klossov vzťah).
- Inverzný Klossov vzťah pre konkrétny pohon je závislý od parametrov AM, ktoré poznáme len približne.
- Základné parametre Klossovho vzťahu (T_{max} , S_{max}) sú závislé od okamžitého pracovného bodu AM (konkrétne od ω_1) podľa rovníc (10), (12).

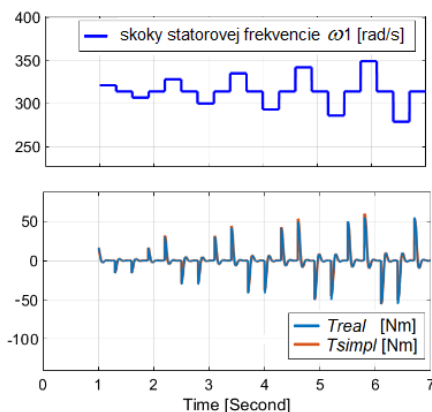
Aby sme sa vyhli uvedeným problémom, je vhodné použiť postup, pri ktorom dokážeme získať optimálnu inverznú charakteristiku AM priamo z meraní na konkrétnom pohone s AM, na čo sú vhodné metódy používané vo fuzzy systémoch, ktoré sú dnes považované za vhodný nástroj pre aproximáciu nelineárnych funkcií. V Kap.III. bol uvedený postup linearizácie nelineárnej charakteristiky AM pre vybraný pracovný bod (rozbeh AM a prevádzka v okolí nominálnej uhlovej rýchlosti a konštantné hodnoty T_{max} , S_{max}). Ak nameriame odozvu konkrétneho AM v rámci nejakého okolia vybraného pracovného bodu, potom je možné z takejto databázy získať fuzzy systém platného Klossovho vzťahu.

Získanú databázu je možné použiť na vyhľadávanie fuzzy inferenčnej štruktúry (FIS) AM popisujúcej nameraný vzťah medzi $[\omega_1, \omega_2] \rightarrow T$. Pomocou nameranej databázy je možné navrhnúť konkrétny fuzzy model štandardne známymi postupmi zhlukovej analýzy. Pre náš účel bola použitá metóda “subtractive clustering” (nástroj ANFIS v programe MATLAB), čo je rýchla a robustná metóda analýzy dát s nasledujúcimi parametrami: Range of influence = 0.4, Squash factor = 1.25, Accept ratio = 0.4, Reject ratio = 0.01. Výsledkom je fuzzy systém typu Sugeno s deviatimi pravidlami, ako je znázornené na Obr. 10.



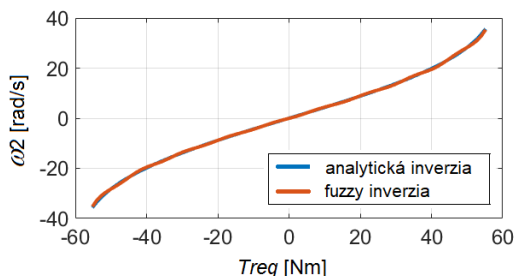
Obr. 10 Fuzzy systém typu SUGENO pre aproximáciu Klossovho vzťahu

Porovnanie ideálneho momentu AM vypočítaného Klossovým vzťahom s jeho fuzzy aproximáciou získanou z nameranej databázy je na Obr. 11. Priebehy na Obr. 11 potvrdzujú kvalitu fuzzy linearizácie momentovej charakteristiky AM, kde sa priebehy porovnaných momentov takmer úplne zhodujú.



Obr. 11 Porovnanie fuzzy aproximácie momentu AM podľa Klossovho vzťahu s ideálnym momentom.

Obrázok 12 ukazuje, že získaná fuzzy inverzia nelineárneho obrazu AM je takmer úplne identická s inverziou získanou z analytického Klossovho vzťahu.



Obr. 12 Porovnanie analytickej a fuzzy inverzie nelinearity AM

V. ZÁVER

Pre kvalitné riadenie pohonov s AM sa dnes používajú regulačné štruktúry na báze vektorového riadenia, ktoré kvôli jeho nelinearitám v princípe vyžadujú transformácie motora do zvolenej rotujúcej sústavy. Od presnosti týchto transformácií závisí kvalita navrhutej regulácie, pričom presnosť týchto transformácií závisí od znalosti aktuálnych parametrov pohonu. Takto sa riadenie pohonu transformuje na riadenie podobné jednosmerným pohonom s cudzím budením.

V predloženom príspevku je ukázaný prístup k linearizácii pohonu s AM pomocou inverzie jeho základnej nelinearity vyjadrenej analyticky Klossovým vzťahom. Využitie fuzzy aproximácie tejto nelinearity umožňuje získať jej inverziu z experimentálne nameraných údajov na konkrétnom AM. Uvedený postup odstraňuje potrebu analytickej znalosti nelinearity AM a tiež znalosti jej analytickej inverzie. Táto metóda linearizácie momentovej slučky pohonu s AM nevyžaduje principiálne žiadne transformácie a umožňuje návrh nadradených regulačných štruktúr (otáčkovej, príp. polohovej regulácie) na báze známych štandardných metód lineárneho riadenia. Použitie inverzného fuzzy „obrazu“ konkrétneho AM umožňuje navyše jeho jednoduchú aplikáciu aj na viacparametrové vstupné hodnoty a prispôbovanie nelinearity pohonu pre rôzne prevádzkové stavy a premenlivé parametre motora.

POĎAKOVANIE

Článok bol publikovaný s podporou projektu APVV-16-0206 a APVV-19-0210.

LITERATÚRA

- [1] N. K. Kumar and V. I. Gandhi, “Implementation of fuzzy logic controller in power system applications,” *J. of Intelligent & Fuzzy Systems*, vol. 36, no. 5, pp. 4115-4126, 2019. DOI: 10.3233/JIFS-169971.
- [2] M. Gaurkar and Ch. Gowder, “Application of fuzzy logic for power system,” *IJARIE*, vol. 3, no. 3, 2017. ISSN(O)-2395-43.
- [3] J. Bačík J., F. Ďurovský, P. Fedor, and D. Perduková, “Autonomous flying with quadcopter using fuzzy control and ArUco markers,” *Intelligent Service Robotics*, vol. 10, no. 3, 2017, pp. 185-194. ISSN 1861-2776. DOI: 10.1007/s11370-017-0219-8.

- [4] P. Fedor and D. Perduková, "Model based fuzzy control applied to a real nonlinear mechanical system," *Iranian Journal of Science and Technology, Trans. of Mechanical Eng.*, vol. 40, no. 2, pp. 113-124, 2016. ISSN 2228-6187. DOI 10.1007/S40997-016-0005-9.
- [5] Ch. H. Chen, Ch. Ch. Wang, Y. T. Wang, and P. T. Wang, "Fuzzy logic controller design for intelligent robots," *Mathematical problems in Eng.*, vol. 2017, Article ID 8984713, 2017. <https://doi.org/10.1155/2017/8984713>.
- [6] V. Ivanov, "A Review of fuzzy methods in automotive engineering applications," *European Transportation Research Review*, pp. 7-29, 2015.
- [7] E. Uzunsoy, "A brief review on fuzzy logic used in vehicle dynamics control," *Journal of Innovative Science and Eng. (JISE)*, vol. 2, no. 1, 2018, pp. 1-7.
- [8] I. J. Balaguer, L. Qin, Y. Shuitao, U. Supatti and Z.P. Fang, "Control for grid-connected and intentional islanding operations of distributed power generation," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 1, pp. 147-157, 2011, ISSN: 0278-0046.
- [9] H. R. Chamorro and G. A. Ramos, "Fuzzy Control in Power Electronics Converters for Smart Power Systems," in book: *Fuzzy Logic - Controls, Concepts, Theories and Applications*. pp. 157-184. 2012. DOI: 10.5772/36311.
- [10] P. Brandstetter and M. Kuchar, "Sensorless control of variable speed induction motor drive using RBF neural network," *Journal of Applied Logic*, vol. 24, pp. 97-108, 2017. DOI: 10.1016/j.jal.2016.11.

Modifikácia informačného systému DARMA pre experiment ALICE CERN

¹Martin KOPECKÝ, ²Tomáš TKÁČIK, ³Ján JADLOVSKÝ

^{1, 2, 3}Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,
Technická univerzita v Košiciach, Slovenská republika

¹martin.kopecky@student.tuke.sk, ²tomas.tkacik@tuke.sk, ³jan.jadlovsky@tuke.sk

Abstrakt — Cieľom príspevku je prezentovať modifikovaný informačný systém DARMA 3, ktorý je momentálne nasadený v rámci experimentu ALICE CERN. V článku je vykonaná analýza súčasného stavu informačného systému DARMA 2 a sú vyšpecifikované možnosti jeho modifikácie. Na základe analýzy je navrhnutý nový koncept modulárnej architektúry založenej na aplikačných knižniciach a sú vykonané príslušné modifikácie funkcionality systému. Modifikovaný informačný systém DARMA je experimentálne overený v laboratórnych podmienkach Centra moderných metód riadenia a priemyselnej informatiky na KKUI FEI TU s využitím kópie dát z experimentu ALICE CERN.

Kľúčové slová — Databázové systémy, experiment ALICE CERN, informačný systém DARMA, návrhové vzory

I. ÚVOD

Príspevok prezentuje výsledky dosiahnuté v rámci riešenia úloh diplomovej práce [1], ktorá sa zaoberá analýzou a realizáciou modifikácií informačného systému (IS) DARMA pre experiment ALICE CERN v rámci prebiehajúceho projektu *Experiment ALICE na LHC v CERN*. IS DARMA, ktorý vznikol významným vylepšením predchádzajúceho informačného systému AMANDA 3 [2], bol súčasťou experimentu ALICE CERN v rokoch 2016-2022. Hoci IS DARMA 2 spĺňal požiadavky používateľov experimentu ALICE CERN, jeho implementácia je spojená s niekoľkými výzvami, najmä v kontexte spracovania veľkých objemov dát a potenciálnej rozšíriteľnosti.

V článku sa zameriame na navrhnutie a implementáciu modifikovanej verzie tohto systému, ktorá by odstránila existujúce obmedzenia a zároveň zabezpečila splnenie nových požiadaviek ACC DCS skupiny v CERNe. V tomto kontexte je dôležitá kompatibilita s existujúcou infraštruktúrou a dátovými zdrojmi používanými v rámci experimentu ALICE CERN.

V druhej časti článku analyzujeme vývoj a súčasný stav IS DARMA, a identifikujeme jeho obmedzenia z pohľadu používateľov experimentu ALICE CERN. V tretej časti článku predstavujeme návrh a implementáciu novej verzie informačného systému DARMA s cieľom eliminovať identifikované obmedzenia. Táto časť zahŕňa prehľad inovácií v architektúre a modularite IS, ako aj pridanie nových obrazoviek. Následne predstavíme experimentálne overenie našich riešení v laboratórnych podmienkach Centra moderných metód riadenia a priemyselnej informatiky (CMMRaPI) pomocou sady jednoduchých aj zložitejších experimentov. V závere článku sumarizujeme hlavné zistenia a prínosy. Poukazujeme na výhody modifikovaného IS DARMA 3, a uvádzame námety pre ďalší rozvoj a modifikácie IS DARMA.

II. VÝVOJ IS DARMA

Informačný systém AMANDA 3 bol vyvinutý členmi výskumnej skupiny CMMRaPI na KKUI FEI TU, ktorí sú súčasne aj riešiteľmi projektu *Experiment ALICE na LHC v CERN*. Hlavným cieľom IS AMANDA 3 bolo zlepšiť efektivitu spracovania údajov z detektorov v rámci experimentu ALICE CERN. Jeho architektúra bola postavená na izolovaných decentralizovaných službách v jazyku C++, komunikujúcich cez technológiu WCF, a pracovala s dvoma typmi databáz - lokálnou systémovou databázou Microsoft SQL a vzdialenou archivačnou databázou Oracle [2].

IS DARMA 1, ktorý nasledoval po IS AMANDA 3, bol vytvorený v rokoch 2016-2018

výskumnou skupinou CMMRaPI [3]. Jeho vývoj zahŕňal analýzu IS AMANDA 3, vytvorenie nového systému pomocou pokročilých metód a nových technológií, pričom zachovával všetky pozitívne aspekty IS AMANDA 3 a pridával nové funkcionality. Na aplikačnej strane využíval jazyk C#, .NET aplikačný rámec a na strane klienta technológie Razor a JS. Celý systém bol postavený na MVC architektúre a komunikoval s Oracle databázami [4].

IS DARMA 2 vznikol v rokoch 2018-2020, kedy došlo k implementovaniu nových návrhov na modifikáciu zo strany vedenia ACC DCS skupiny v CERNe [5] [6]. Z IS DARMA 2 sa v rokoch 2020-2021 vyvinul IS DARMA 2+, ktorého cieľom bola integrácia s ďalšími vizualizačnými nástrojmi, ako napríklad grafickým rozhraním GRAFANA. V roku 2022 došlo k oddeleniu informačných systémov DARMA 2 a Grafana [7]. Tento krok umožnil lepšie zameranie sa na potreby jednotlivých systémov a ich používanie.

A. Navrhované modifikácie IS DARMA 2

Analýzou IS DARMA 2 a po konzultáciách členov výskumnej skupiny CMMRaPI s vedením ACC DCS skupiny v CERNe sme dospeli k možným modifikáciám, ktoré sú znázornené Obr. 1.



Obr. 1 Potenciálne modifikácie na inováciu IS DARMA 2.

Potenciálne modifikácie zobrazené na Obr. 1 sme implementovali nasledovne:

- **Refaktoring kódu** - refaktoring sa zameriava na odstránenie nepoužívaných funkcií, aby sa zlepšila čitateľnosť a údržba IS.
- **Rozšíriteľnosť a modularita** - cieľom je zlepšenie modularity a rozšíriteľnosti IS, čo umožní jednoduchšie pridávanie nových a zmenu existujúcich funkcií.
- **Integrácia viacerých databáz** - optimalizovaním IS pre efektívnu integráciu viacerých databázových systémov zlepšime správu databázových pripojení a prístup k údajom.
- **Vytvoriť rozhrania na alerty** - nové rozhranie pre alerty umožní lepšiu komunikáciu medzi systémom a členmi ACC DCS skupiny v CERNe.
- **Detekcia nesprávne nakonfigurovanej archivácie datapointov** - nové webové rozhranie umožní detekciu problémov s archiváciou a zabezpečí optimálnu prevádzku systému.
- **Vizuálna reprezentácia** - zahrnutie novšieho vizuálneho dizajnu IS pre lepšiu používateľskú skúsenosť.
- **Stabilizácia IS:** - vykonanie dôkladnej analýzy systému s cieľom identifikovať a odstrániť chyby, optimalizovať kód a zlepšiť výkon aplikácie.
- **Preverenie limitov systému** - implementácia metód na identifikáciu a monitorovanie limitov systému, a automatické informovanie používateľa pri ich prekročení.
- **Nasadenie na Linuxovej platforme** - modifikácia IS pre zabezpečenie kompatibility s Linuxovým prostredím, čo zjednoduší proces nasadenia a údržby.

III. NÁVRH A IMPLEMENTÁCIA NOVEJ VERZIE INFORMAČNÉHO SYSTÉMU DARMA 3

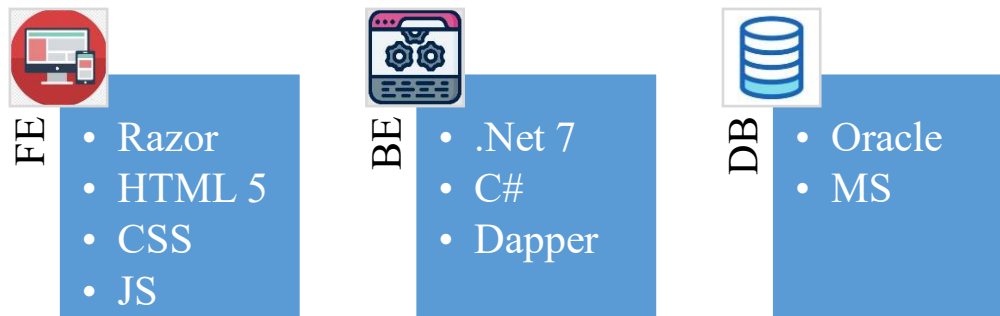
V tejto časti podrobne popíšeme implementáciu modifikácií uvedených na Obr. 1. Zmeny, ktoré navrhujeme, môžeme rozdeliť do dvoch sekcií, a to zmeny pre úpravu štruktúry kódu existujúceho systému a implementovanie nových obrazoviek. Následne sa táto časť venuje experimentom zameraným na overenie funkčnosti IS DARMA 3 v laboratórnych podmienkach CMMRaPI na KKUI FEI TU (<http://kyb.fei.tuke.sk/laboratoria/modely/cern.php>).

A. Refaktoring kódu, modifikácia architektúry a modularita

V tejto časti si popíšeme modifikácie, ktoré navrhujeme s cieľom zjednodušenia modifikácie pôvodného systému bez zmeny jeho funkcionality.

Vývoj IS DARMA 3 sme realizovali s využitím najmodernejších technológií a prístupov [8].

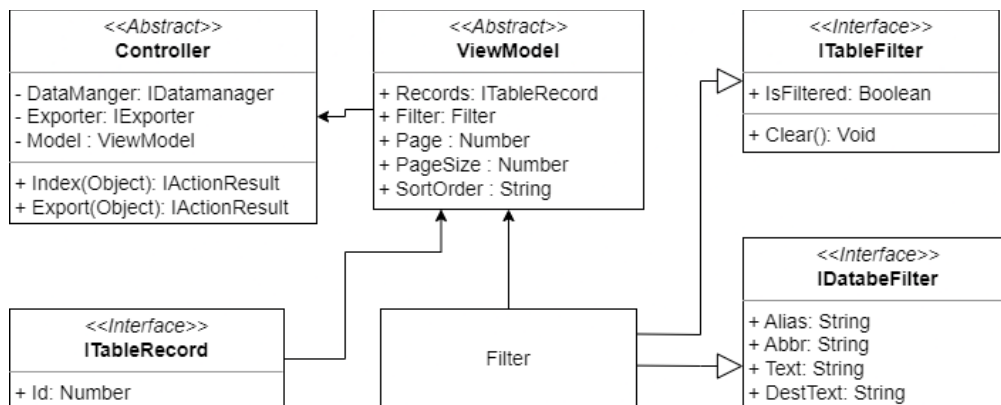
Na prácu s používateľským rozhraním využívame technológie JS a Razor stránky pre ochranu pred útokmi zo strany klienta. Na aplikačnú logiku využívame .NET 7 a ORM Dapper pre flexibilnejšiu manipuláciu s databázou. Komunikácia s DB je asynchrónna, čo zlepšuje rýchlosť a škálovateľnosť IS, a ORM poskytuje ochranu pred útokmi smerujúcimi na databázu. Využitie technológie a ich zaradenie sú vizualizované na Obr. 2. Keďže IS DARMA 2 využíval architektúru MVC, rozhodli sme sa tento prístup ponechať aj v IS DARMA 3.



Obr. 2 Použité technológie na vývoj IS DARMA 3.

Za účelom refaktorovania kódu sme sa zamerali na elimináciu nadbytočných funkcií a zlepšenie čitateľnosti a údržby informačného systému. Použitím návrhových princípov DRY a SOLID sme dosiahli značné vylepšenia. Použitie statických tried pre opakujúce sa metódy, ako napríklad ValidationHelper a DatasourceHelper, nám umožnilo zlepšiť znovu použiteľnosť kódu a dodržanie princípu DRY. Vylepšená je aj štruktúra modulov na prácu so službami, kde využívame návrhový vzor Stratégia [8] tak, že každý ViewModel teraz obsahuje IDataManager špecifický pre daný Controller.

V oblasti rozšíriteľnosti a modularity využívame prístup založený na SOLID princípoch, konkrétne Open/Closed Principle (OCP) a Liskov Substitution Principle (LSP), pre dosiahnutie lepšej flexibility a separácie komponentov. Tieto zmeny umožňujú jednoduchšie pridávanie nových funkcií a úpravu existujúcich, čo znižuje časovú náročnosť údržby IS. Príkladom použitia tohto prístupu je využitie kompozitného vzoru na vytvorenie hierarchie objektov v triede LogsViewModel [8]. Výsledná štruktúra aplikačnej vrstvy je zobrazená na Obr. 3.



Obr. 3 Vytvorená štruktúra aplikačnej vrstvy pomocou kompozitného vzoru a princípov OCP a LSP.

Zároveň sme implementovali komunikáciu s viacerými databázami prostredníctvom knižnice "DataAccess". Táto knižnica je navrhnutá podľa návrhového vzoru Fasáda [8] a na komunikáciu s databázami využíva ORM Dapper. Práca s knižnicou je navrhnutá tak, aby sa v prípade potreby vedela knižnica nezávisle modifikovať, resp. vymeniť za inú. Tieto nástroje nám umožnili zjednodušiť prácu s databázami a zabezpečiť efektívnu interakciu medzi aplikačnou logikou a databázovým systémom. Taktiež sme implementovali asynchrónnu komunikáciu s databázami pre zlepšenie rýchlosti a škálovateľnosti systému.

B. Nové komponenty v IS DARMA 3

Táto časť sa venuje implementácii nových obrazoviek podľa požiadaviek skupiny ACC DCS v CERNe v štruktúre uvedenej na Obr. 2.

Na prácu s ALERT-mi boli vytvorené dve obrazovky - AlertHistory a AlertManagement -

ktoré slúžia na sledovanie a správu alertov. Pre tieto obrazovky je navrhnutý a implementovaný filtračný nástroj, ktorý umožňuje používateľom filtrovať dáta podľa názvu detektora, času vzniku, obsahu popisu a závažnosti alertu. Vizualizácia dát je realizovaná v tabuľkovej forme, ktorá umožňuje používateľom zoradovať údaje podľa rôznych kritérií. Tabuľka je navrhnutá tak, aby poskytovala možnosť nastaviť počet zobrazených údajov, zoradovať ich podľa ľubovoľného stĺpca a stránkovanie pre jednoduchší prehľad veľkého množstva záznamov. Na Obr. 4 je zobrazená obrazovka Alerts Management s vyfiltrovanými dátami a závažnosť alertu je vizualizovaná zmenou farby stĺpca „Alert Type“.

	Alias	Ts	Text	Description	Direction	Alert Type
1	PT212	28. 1. 2021 12:07:05	HIGH PRESSURE	HIGH PRESSURE	1	WARNING
2	PT212	11. 2. 2021 13:52:38	HIGH PRESSURE	OK	0	WARNING
3	PT212	4. 2. 2021 12:35:08	ERROR	HIGH PRESSURE	0	ERROR
4	PT212	4. 2. 2021 12:35:08	HIGH PRESSURE	OK	0	WARNING
5	PT212	4. 2. 2021 12:35:08	LOW PRESSURE	LOW PRESSURE	1	WARNING
6	PT212	4. 2. 2021 12:35:08	ERROR	ERROR	1	ERROR
7	PT212	4. 2. 2021 12:35:10	ERROR	LOW PRESSURE	0	ERROR
8	PT212	4. 2. 2021 12:35:10	LOW PRESSURE	OK	0	WARNING
9	PT212	4. 2. 2021 12:35:10	HIGH PRESSURE	HIGH PRESSURE	1	WARNING
10	PT212	4. 2. 2021 12:35:10	ERROR	ERROR	1	ERROR
11	PT212	4. 2. 2021 12:35:49	ERROR	HIGH PRESSURE	0	ERROR
12	PT212	4. 2. 2021 12:35:49	HIGH PRESSURE	OK	0	WARNING
13	PT212	4. 2. 2021 12:35:49	LOW PRESSURE	LOW PRESSURE	1	WARNING
14	PT212	4. 2. 2021 12:35:49	ERROR	ERROR	1	ERROR
15	PT212	4. 2. 2021 12:36:00	ERROR	LOW PRESSURE	0	ERROR

Obr. 4 Obrazovka Alerts Management - IS DARMA 3.

Na detekciu správnej, resp. zlej konfigurácie datapointov bola vytvorená samostatná obrazovka. Kľúčovou časťou je vytvorenie algoritmu na detekciu zlej konfigurácie datapointov, kde sa deteguje neadekvátne zapisovanie údajov do databázy. Zamiera sa na detekciu časových období, v ktorých je počet zapísaných údajov odlišný od zvyčajného. Na tejto obrazovke bol tiež navrhnutý a implementovaný filtračný nástroj, ktorý umožňuje používateľom filtrovať dáta na základe databázy a schémy, názvu detektora, roka a či chcú sledovať údaje za mesiac alebo týždeň. Vizualizácia dát je opäť realizovaná v tabuľkovej forme, ktorá poskytuje rovnaké možnosti pre prácu s údajmi ako na obrazovkách v sekcii Alerts.

C. Experimentálne overenie IS DARMA 3 v laboratórnych podmienkach CMMRaPI

Táto časť sa venuje experimentálnemu overeniu modifikácií informačného systému DARMA 3 v laboratórnych podmienkach CMMRaPI na KKUI FEI TU. Testy boli vykonávané na testovacej databáze, ktorá je kópiou jednej z archivačných databáz používaných v CERN-e. Overovali sa najmä novo implementované obrazovky a moduly. Testovanie bolo realizované v dvoch etapách. Testy v prvej etape boli zamerané na overenie funkcionality IS DARMA 3, kde sa testovali nové obrazovky a moduly. Detaily týchto testov sú zaznamenané v Tab. 1.

Tab. 1 Overenie funkcionality nových obrazoviek v systéme DARMA 3.

Obrazovka	Akcia	Počet experimentov
AlertsManagement	Vyhľadávanie dát podľa filtra	2
AlertsHistory	Vyhľadávanie dát podľa filtra	2
DatapointsManagement	Vyhľadávanie dát podľa filtra	4

Druhá etapa testov sa zaoberala porovnávaním trvania požiadaviek medzi IS DARMA 2 a IS DARMA 3. Detaily týchto testov sú zaznamenané v Tab. 2.

Tab. 2 Porovnanie výkonnosti medzi informačnými systémami DARMA 2 a DARMA 3.

Akcia	IS DARMA 2	IS DARMA 3
Časová náročnosť sťahovania	2 hodiny a 18 minút	1 hodina a 47 minút (-22 %)
Súbežné požiadavky - Alias	2 minúty a 30 sekúnd	1 minúta a 56 sekúnd (-23 %)
Súbežné požiadavky - Element	5 minút a 2 sekundy	5 minút a 43 sekúnd (+14 %)

Z výsledkov testov možno zhodnotiť, že novo implementované obrazovky a moduly v IS DARMA 3 fungujú správne. Pri porovnávaní výkonnosti IS DARMA 2 a IS DARMA 3 sa ukázalo, že nová verzia systému ponúka podobný výkon v čase spracovania požiadaviek a rýchlosti sťahovania dát ako staršia verzia. Všeobecne možno konštatovať, že implementácia nových funkcionalít a modifikácií IS DARMA 3 spĺňa požiadavky používateľov experimentu ALICE CERN, čo potvrdili experimenty vykonané v laboratórnych podmienkach v CMMRaPI na KKUI FEI TU.

IV. ZÁVER

V príspevku prezentujeme úspešne modifikovaný informačný systém DARMA pre experiment ALICE CERN. Vykonali sme analýzu IS DARMA 2, navrhli a implementovali modifikácie, vrátane integrácie archivačných databáz, selekcie alertov a detekcie nesprávne nakonfigurovaných datapointov. Použitím technológií .NET 7 a ORM Dapper a moderných návrhových vzorov sme dosiahli robustné riešenie. Funkcionalita informačného systému DARMA 3 bola overená v laboratórnych podmienkach CMMRaPI na KKUI FEI TU. Informačný systém je momentálne nasadený v CERN-e na dedikovanom serveri a beží v obmedzenej prevádzke na otestovanie našej implementácie a na odhalenie potenciálnych chýb, ktoré neboli zaznamenané počas fázy vývoja a testovania. Súčasťou informačného systému je aj používateľská príručka pre CERN v anglickom jazyku.

V budúcnosti možno rozšíriť IS DARMA 3 napríklad o optimalizáciu sťahovania dát z archivačných databáz prostredníctvom paralelného asynchrónneho sťahovania. Dosiahnuté výsledky zabezpečujú základ pre budúce inovácie informačného systému DARMA 3 členmi CMMRaPI, ktorí sú riešitelia projektu *Experiment ALICE na LHC v CERN*.

POĎAKOVANIE

Táto práca bola podporovaná projektom Experiment ALICE na LHC v CERN: Štúdium silno interagujúcej hmoty v extrémnych podmienkach (ALICE KE FEI TU 0410 / 2022).

REFERENCIE

- [1] KOPECKÝ, M.: Realizácia informačných systémov na úrovni databázových systémov s väzbou na SCADA a vnorené systémy - ALICE CERN : Diplomová práca. Košice: FEI TU, 2023. 68 s. (školiťel: doc. Ing. Ján Jadlovský, CSc., konzultant: Ing. Tomáš Tkáčik)
- [2] JADLOVSKÝ, J., JADLOVSKÁ, A., SARNOVSKÝ, J., JAJČIŠIN, Š., ČOPÍK, M., JADLOVSKÁ, S., PAPAN, P., BIELEK, R., ČERKALA, J., KOPČÍK, M., CHOCHULA, P., AUGUSTINUS, A.: External Access to ALICE Controls Conditions Data. In: Journal of Physics: Conference Series (JPCS), Vol. 513, s. 1-5, 2014, ISSN 1742-6588.
- [3] JADLOVSKÝ, J., JADLOVSKÁ, S., ČABALA, J., JADLOVSKÁ, A., ČERKALA, J., KOPČÍK, M., ORAVEC, M., VOŠČEK, D., TKÁČIK, M., HANC, E., CHOCHULA, P., BOND, P.: Information system for ALICE experiment data access. In: ICALEPS 2017. - Barcelona : Mondial and Cititravel Congressos, 2017, 1-4 s. - ISBN978-3-95450-193-90.
- [4] PÁNIK, M.: Realizácia aplikácií na úrovni databázových systémov s väzbou na SCADA systémy a Embedded systémy : Bakalárska práca. Košice: FEI TU, 2018. 62 s. (školiťel: doc. Ing. Ján Jadlovský, CSc.)
- [5] PÁNIK, M.: Realizácia informačného systému pre prístup k dátam z LHC v rámci experimentu ALICE-CERN : Diplomová práca. Košice: FEI TU, 2020. 84 s. (školiťel: doc. Ing. Ján Jadlovský, CSc.)
- [6] JADLOVSKÝ, J., TKÁČIK, M., JADLOVSKÁ, S., PÁNIK, M.: DARMA - Information System for Accesing LHC Data within the ALICE Experiment at CERN, 2020. In: Acta Electrotechnica et Informatica. - Košice, FEI Roč. 20, č. 4 (2020), s. 19-26 ISSN 1335-8243.
- [7] JADLOVSKÝ, J., JADLOVSKÁ, A., JADLOVSKÁ, S., TKÁČIK, M., TKÁČIK, T., PUGELOVÁ, Z., PAZDIČ, F., LÍŠKOVÁ, D., KOPECKÝ, M. [2022]. Správa o použití finančných prostriedkov na spolufinancovanie projektu výskumu a vývoja ALICE CERN za rok 2022 (ALICE KE FEI TUKE 0410/2022), Technical report, Technická Univerzita v Košiciach.
- [8] OSBORNE, M.: Design patterns. [Online] Dostupné na: <<https://refactoring.guru/design-patterns>> [cit. 2023-04-19]

Návrh a programová implementácia sociálnej siete

¹Marek Belej, ¹Branislav Madoš, ¹Anton Baláž

¹ Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

marek.belej@student.tuke.sk, branislav.mados@tuke.sk, anton.balaz@tuke.sk

Abstract — Článok sa zaoberá návrhom a programovou implementáciou sociálnej siete určenej pre firemné prostredie s cieľom umožniť spájanie profesionálov z rôznych profesných oblastí podľa predmetov ich záujmu. Táto práca sa venuje analýze problematiky súčasného stavu sociálnych sietí, kde opisuje definíciu, históriu a súčasné najpoužívanejšie sociálne siete. Tieto analýzy vedú v ďalšej časti článku k popisu návrhu vlastnej sociálnej siete a popisu jej architektúry a zhrnutiu výsledkov jej testovania.

Keywords — sociálna sieť, sociálne médiá, sociálna platforma, webová aplikácia

I. ÚVOD

V dnešnej dobe sú sociálne siete nedeliteľnou súčasťou nášho bežného života. Je to miesto, kde si ľudia vymieňajú informácie, komunikujú a budujú vzťahy. V profesionálnom a firemnom prostredí sú sociálne siete ešte dôležitejšie, pretože umožňujú ľuďom v danom odvetví prepojiť sa a zdieľať odborné znalosti a skúsenosti. Cieľom práce je navrhnúť a programovo implementovať sociálnu sieť určenú pre firemné prostredie s cieľom umožniť spájanie profesionálov z príslušných profesijných oblastí podľa predmetov ich záujmu.

V prvom kroku táto práca definuje pojem sociálna sieť, bude sa venovať histórii sociálnych sietí a charakterizuje hlavné znaky a funkcie sociálnych sietí. Následne sa bude venovať súčasnému stavu, kde budú sociálne siete a ich hlavné funkcionality jednotlivo popísané.

Na základe analýzy problematiky sociálnych sietí a profesionálnych sociálnych sietí je vypracovaný návrh vlastnej sociálnej siete.

V poslednej časti práce je implementované riešenie popísané na základe predošlých požiadaviek, navrhnutých spôsobov riešenia a technológií.

II. SOCIÁLNE SIETE

Sociálna sieť je súbor sociálne relevantných uzlov a jednotiek, ktorými sú členovia spojení jedným alebo viacerými vzťahmi. Týmito jednotkami sú najčastejšie osoby alebo organizácie, ale v zásade možno ako uzly nazývať akékoľvek jednotky, ktoré možno pripojiť k iným jednotkám [1]. Sociálne siete predstavujú súbor online nástrojov, ktoré podporujú sociálnu interakciu medzi používateľmi. Tento termín sa často používa v kontraste s tradičnejšími médiami, akými sú televízia a knihy, ktoré poskytujú obsah masovej populácii, ale neľahčujú vytváranie alebo zdieľanie obsahu používateľmi. Bednář [2] charakterizuje sociálne siete ako sociálny softvér, ktorý umožňuje používateľom vytvárať a udržiavať spojenia s inými používateľmi prostredníctvom vzájomne prepojených kontaktov. Tieto aplikácie tiež poskytujú funkcie, ktoré umožňujú publikovanie a zdieľanie rôznych informácií v rámci virtuálnej komunity sociálnej siete.

A. História sociálnych sietí

V polovici 90-tych rokov sa začali objavovať prvotné sociálne siete, ktoré sa sústredili na spájanie ľudí v internetových diskusiách a na zdieľanie osobných informácií prostredníctvom osobných webových stránok. Boli to sociálne siete, ako napríklad theGlobe.com, GeoCities a Tripod.com. V tom čase bolo cieľom sociálnych sietí vzájomne spájať ľudí s rovnakými záujmami, ale stále to bolo veľmi obmedzené. Koncom 90-tych rokov došlo k zmene a ústredným prvkom sociálnych sietí sa stala používateľská profilová stránka. Tento prvok umožnil používateľom zostaviť vlastný

zoznam priateľov a nájsť nových, rovnako zmysľajúcich ľudí. Prvý portál, ktorý tento prvok využil bol Friendster v roku 2002. Nasledovali LinkedIn a MySpace v roku 2003. V roku 2004 vznikol Facebook, ktorý je dnes najväčšou a najpopulárnejšou sociálnou sieťou na svete s viac ako 2,96 miliardami aktívnych používateľov. V súčasnosti existuje viac ako 200 internetových stránok, ktoré sú založené na modeli sociálnych sietí a ich popularita neustále rastie [3].

Internetová stránka Classmates.com je uvádzaná ako prvá sociálna sieť. Bola založená v roku 1995 Conradsom Randym. Sociálna sieť Classmates.com bola pôvodne určená na vyhľadávanie spolužiakov z materských škôl, zo základných, stredných a vysokých škôl.

V roku 2002 Jonathan Abrams a Ross McKinnon založili Friendster, službu pre zoznamovanie. Cieľom bolo vytvoriť bezpečné a účinné prostredie na stretávanie sa ľudí, prezeranie profilov a nadviazanie kontaktov. Na začiatku sa Friendster stal populárnym medzi blogermi, účastníkmi Burning Man Art festivalu.

Sociálnu sieť a web stránku MySpace založilo v roku 2003 niekoľko zamestnancov americkej internetovej marketingovej spoločnosti eUniverse. Tím eUniverse videl potenciál stránok ako Friendster a rozhodol sa založiť vlastnú sociálnu sieť. Sprístupnenie prvej verzie stránky online trvalo iba desať dní. MySpace ťažil z dvadsiatich miliónov používateľov a e-mailových predplatiteľov eUniverse, ktorí boli nasmerovaní na novú sociálnu sieť. Myspace sa rýchlo stal populárnym medzi dospelými a mladým dospelým publikom. Známky úpadku Myspace sa objavili 19. apríla 2008, keď ho Facebook predbehol v rebríčku Alexa. Odvtedy členstvo v Myspace naďalej klesalo a stránka už nikdy nedokázala znovu získať takú úroveň návštevnosti, akú mala počas vrcholných rokov. Myspace vyskúšal niekoľko stratégií, aby znovu získal svoje publikum s obmedzeným úspechom. Zahŕňali rozsiahle redizajny stránok a posun v zameraní na hudbu a zábavu, a nie na sociálne siete. Tieto pokusy však zväčša zlyhali. Myspace dnes zostáva online, ale jeho používateľská základňa sa výrazne zmenšila. Do roku 2019 klesol mesačný počet návštevníkov Myspace na sedem miliónov. V tom roku technická chyba spôsobila, že všetok obsah nahraný pred rokom 2015 zmizol, čo ešte viac poškodilo túto sociálnu sieť [4].

LinkedIn, obchodne orientovaná webová stránka sociálnych sietí založená v roku 2002 so sídlom v Mountain View v Kalifornii. Na rozdiel od iných sociálnych sietí, ako sú Facebook a Myspace, ktoré sú často čisto rekreačné, LinkedIn kladie dôraz na profesionálne kontakty používateľa. Používatelia vytvárajú profilové stránky, ktoré majú štruktúru podobnú životopisu v tom, že používatelia môžu zhrnúť svoju kariéru, inzerovať svoje konkrétne zručnosti a uviesť svoje vzdelanie a históriu zamestnania. Webová stránka bola spustená v roku 2003. Rast bol spočiatku pomalý. V roku 2005 LinkedIn predstavil služby, ktoré umožnili spoločnostiam zverejňovať pracovné ponuky a vyhľadávať v sieti potenciálnych zamestnancov. Okrem týchto profesionálnych služieb LinkedIn umožňuje spoločnostiam inzerovať na stránke. LinkedIn sa konečne stal ziskovým v roku 2007. V tom roku mal LinkedIn viac ako 15 miliónov členov a do roku 2011 mal LinkedIn viac ako 100 miliónov členov po celom svete. Počítačová verejná ponuka LinkedIn v tom istom roku priniesla 353 miliónov dolárov. V roku 2016 LinkedIn získal Microsoft za približne 26 miliárd dolárov. Nasledujúci rok hlásila viac ako 500 miliónov členov v približne 200 krajinách [5].

Facebook založili v roku 2004 Mark Zuckerberg, Eduardo Saverin, Dustin Moskovitz a Chris Hughes, ktorí boli študentmi Harvardskej univerzity. Facebook sa stal najväčšou sociálnou sieťou na svete. Spoločnosť má komplikovanú rannú históriu, ktorá začala na Harvardskej univerzite v roku 2003 ako Facemash, online služba pre študentov na posúdenie atraktívnosti svojich spolužiakov. TheFacebook.com bola spustená vo februári 2004 a umožnila študentom Harvardskej univerzity zverejniť svoje fotografie a osobné informácie, ako napríklad ich rozvrh hodín a príslušnosť k rôznym klubom. Do konca roka 2004 dosiahol TheFacebook milión aktívnych používateľov. V treťom kvartáli roka 2008 mal Facebook 100 miliónov používateľov a v poslednom kvartáli roka 2022 to bolo až 2,96 miliardy.

B. Klasifikácia a hlavné znaky sociálnych sietí

Sociálne siete majú mnoho rôznych znakov a funkcií, ktoré im umožňujú plniť ich úlohu a poskytovať pridanú hodnotu pre používateľov. Medzi tieto znaky patria napríklad hodnotenie, zdieľanie obsahu, odlúčenie od času a priestoru, dostupnosť a mnoho ďalších. Tieto znaky je možné považovať aj za ciele sociálnej siete [5].

Dostupnosť Kľúčom k dostupnosti sociálnych sietí je internetové pripojenie, pretože sociálne siete sú webové stránky alebo mobilné aplikácie, ktoré sú umiestnené na internete a ktoré vyžadujú internetové pripojenie pre prístup k nim.

Tvorba obsahu Používatelia môžu tvoriť a meniť obsah webu. Sociálne médiá sú sociálne predovšetkým preto, že "čitatelia oznamujú svoj vlastný názor na diskutovanú tému a tým

ovplyvňujú názory ostatných" [6]. Jednou z hlavných charakteristík sociálnych sietí je možnosť používateľov tvoriť a meniť obsah, čím sa stávajú aktívnymi účastníkmi a tvorcami obsahu.

Multimedialita Sociálne siete sa často charakterizujú svojou schopnosťou pracovať s rôznymi formami multimédií, ktoré v súčasnosti umožňujú používateľom nahráť rôzne typy súborov a formátov, ako sú fotografie, videá, animácie, zvukové záznamy a mnoho ďalších. Tieto prvky umožňujú používateľom zdieľať svoje skúsenosti a emócie prostredníctvom vizuálnych a zvukových prvkov, čo umocňuje ich interakciu a komunikáciu.

Hodnotenie Je jedným zo znakov sociálnej siete, pretože používateľom umožňuje vyjadriť svoj názor a poskytnúť spätnú väzbu na produkt, službu, značku alebo iného používateľa. Týmto spôsobom sa vytvárajú virtuálne komunity, ktoré diskutujú o rôznych témach a navzájom sa ovplyvňujú.

Zdieľanie Používatelia na sociálnych sieťach môžu zdieľať rôzne druhy obsahu, ako sú fotografie, videá, články alebo iné príspevky. Týmto spôsobom sa umožňuje šírenie informácií a komunikácia medzi používateľmi, ktorí môžu byť navzájom od seba oddelení geograficky alebo inými faktormi. Zdieľanie obsahu tiež umožňuje používateľom prezentovať svoje názory a záujmy a nadviazať nové kontakty.

Odlúčenie od času a priestoru Je umožnené používateľom komunikovať a interagovať navzájom bez ohľadu na to, kde sa nachádzajú alebo v akom čase. Týmto spôsobom sa umožňuje vytváranie virtuálnych komunít a vzťahov medzi používateľmi, ktorí by sa inak nemuseli nikdy stretnúť.

Komunikácia a interakcia Jednými z kľúčových znakov sociálnych sietí sú práve komunikácia a interakcia. Sociálne siete umožňujú používateľom komunikovať a interagovať s ostatnými ľuďmi bez ohľadu na miesto a čas. Cieľom komunikácie a interakcie na sociálnych sieťach je výmena informácií a názorov medzi používateľmi. Používatelia môžu zdieľať svoje myšlienky, názory, zážitky a emócie prostredníctvom rôznych formátov, ako sú textové príspevky, fotografie, videá, a získavať spätnú väzbu od ostatných používateľov.

Sociálne siete možno klasifikovať podľa rôznych kritérií, ako sú napríklad:

- Podľa orientácie na zisk
 - ziskové
 - neziskové
- Podľa počtu používateľov
 - malé, ktoré majú len pár tisícok používateľov
 - veľké, ktoré majú milióny používateľov
- Podľa geografickej polohy
 - celosvetové
 - lokálne
- Podľa účelu
 - špecifické
 - sociálne siete bez zámeru
- Podľa zamerania
 - používateľské (Facebook)
 - mikroblovacie (Twitter)
 - obsahovo orientované (Youtube)

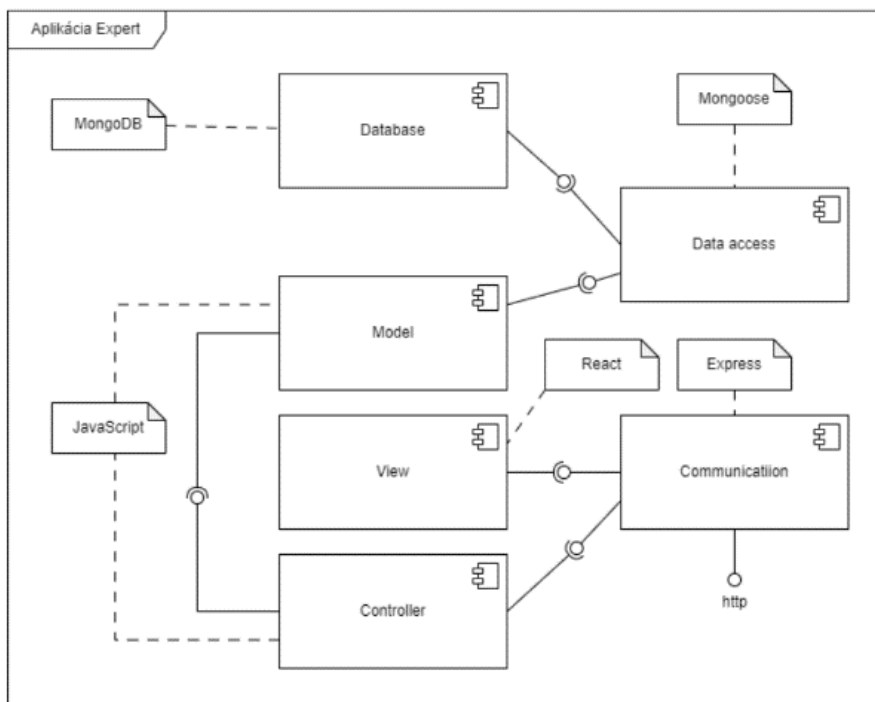
Osobitnou kategóriou, ktorá je zaujímavá aj z hľadiska práce popisovanej týmto článkom, sú sociálne siete, ktoré sú určené pre ľudí s profesionálnymi záujmami, ako sú napr. LinkedIn, Doximity, LawLink, Canvas, ResearchGate, Behance, Github a ďalšie [7]. Problematika sociálnych sietí je súčasne rozoberaná z hľadiska rôznych, čoraz partikulárnejších pohľadov [8][9].

III. NÁVRH A IMPLEMENTÁCIA VLASTNEJ SOCIÁLNEJ SIETE

Návrh vlastnej sociálnej siete určenej pre profesionálov je založený na myšlienke, že bude vytvorená sociálna sieť so známou, bežnou funkcionalitou, ako je profil používateľa, vytváranie kontaktov, komunikácia, vyhľadávanie, tvorba príspevkov a interakcie s nimi. To, čo túto sociálnu sieť výrazne odlišuje od ostatných, je vytváranie komunít do ktorých používateľ môže patriť, od všeobecnej komunity jeho priateľov cez profesnú komunitu, komunitu na základe jeho zamestnaneckého pomeru, voľnočasových aktivít, ako je šport apod., pričom celý obsah súvisiaci s danou komunitou bude zoskupený do jednej záložky. Bude tak možné oddeliť kontakty, príspevky, konverzácie a sponzorský obsah do jednotlivých záložiek.

Základná funkcionálna sociálnej siete bude zahŕňať najmä

- Profil používateľa - Vytvorenie profilu používateľa, ktorý bude zahŕňať
 - meno a priezvisko,
 - povolanie,
 - možnosť prídania fotografie,
 - vlastnú nástenku s jeho príspevkami,
 - zoznam kontaktov,
 - komunity v ktorých je členom.
- Vytvorenie kontaktov: Používateľ tejto siete by mal mať možnosť pridávania ostatných používateľov pre vytváranie siete kontaktov. Toto pridávanie je implementované formou žiadostí, aby sa predišlo nechceným kontaktom. Taktiež je k dispozícii možnosť odporúčania pre pridávanie do kontaktov.
- Komunikácia: Priama komunikácia medzi používateľmi patrí ku kľúčovým funkciám sociálnej siete pre zlepšovanie vzťahov medzi používateľmi.
- Vyhľadávanie: Možnosť vyhľadávania ostatných používateľov alebo komunit.
- Tvorba príspevkov: Budovanie obsahu sociálnej siete používateľov.
- Interakcia s príspevkami: Možnosť diskusie ku príspevku v rámci komentárov a ohodnotenie príspevkov.
- Komunity: Vytvorenie skupiny používateľov, ktorá bude zahŕňať:
 - názov komunity,
 - kategóriu,
 - popis,
 - správu komunity autorom alebo iným povereným používateľom,
 - komunitnú nástenku s príspevkami,
 - komunitnú komunikáciu,
 - zoznam používateľov v tejto komunite.



Obr. 1 MVC architektúra navrhnujej sociálnej siete.

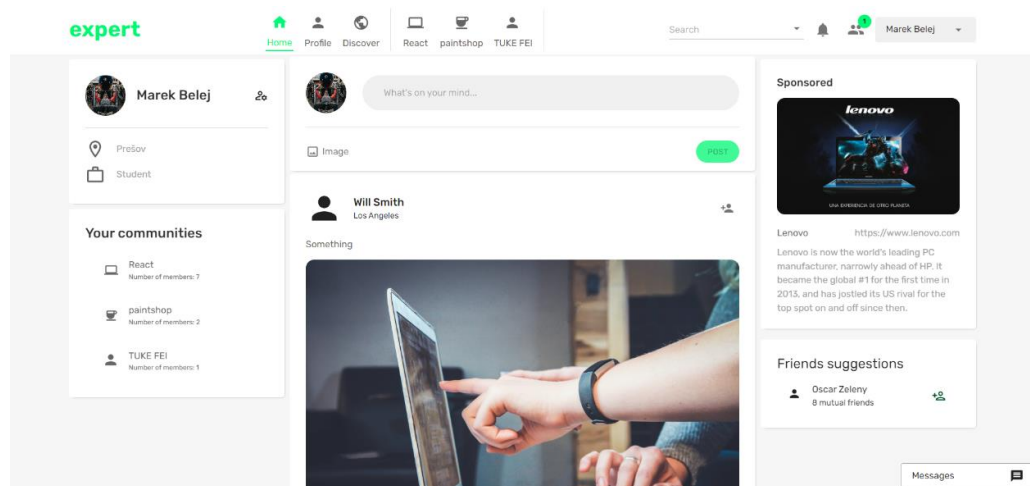
Aplikácia je podľa návrhu vykresľovaná pomocou CSR, aby sa zabezpečil plynulý chod aplikácie. Nie je teda potrebné aby sa pri každom volaní na server vykresľovala celá stránka nanovo. Architektúra webovej aplikácie s využitím MERN stacku sa skladá z dvoch hlavných častí: frontend a backend, ktoré boli vyvíjané samostatne, ale zároveň paralelne.

Frontend pozostáva z Reactu, ktorý je používaný pre vytvorenie používateľského rozhrania a súčasne zabezpečuje logiku aplikácie. Backend sa skladá z Node.js a Expressu, ktoré sú používané pre vytvorenie REST API pre komunikáciu s frontendom a databázou MongoDB. Architektúra tejto webovej aplikácie pozostáva z viacerých vrstiev.

Najdôležitejšie vrstvy sú:

- Prezentačná vrstva: Táto vrstva je tvorená Reactom a zabezpečuje používateľské rozhranie aplikácie. Túto časť používateľa vidia a interagujú s ňou.
- Aplikačná vrstva: Táto vrstva sa nachádza na serveri a zabezpečuje logiku aplikácie. Používa sa na spracovanie dát a na spracovanie požiadaviek od klientov. Používané technológie pre implementáciu tejto vrstvy sú Node.js a Express.
- Údajová vrstva: Táto vrstva zahŕňa databázu MongoDB a API pre komunikáciu s ňou. MongoDB poskytuje úložisko pre všetky dáta aplikácie. API zabezpečuje prístup k týmto dátam.

Podrobnejšie dokumentuje architektúru navrhutej webovej aplikácie predstavujúcej sociálnu sieť Obr.1. Na Obr. 2 potom možno vidieť používateľské rozhranie sociálnej siete, zobrazujúce jej domovskú stránku.



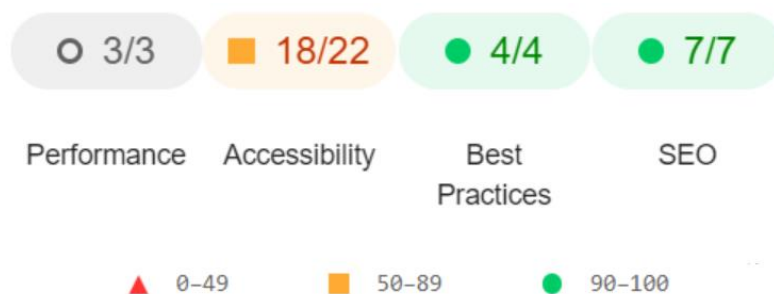
Obr. 2 Domovská stránka navrhutej sociálnej siete.

IV. TESTOVANIE DOSIAHNUTÉHO RIEŠENIA

Výsledky testovania sú kľúčovou súčasťou každého testovacieho procesu a poskytujú podrobný prehľad o tom, ako webová aplikácia funguje. V prípade testovania klientskej a serverovej časti aplikácie je dôležité sledovať rôzne parametre, ktoré ovplyvňujú výkon a výkonnosť aplikácie. Výsledky testovania ukázali, že aplikácia je schopná zvládnuť požadovanú úroveň záťaže a výkon klientskej časti je v súlade s odporúčaniami a najlepšimi postupmi. Testovanie poukázalo na to, že pre škálovanie počtu používateľov aplikácie je nutné navýšiť funkcie MongoDB, ktoré už ale nie sú bezplatné. Celkovo možno zhodnotiť, že aplikácia splnila požiadavky a očakávania, čím je možné prehlásiť testovanie za úspešné.

Prvá fáza testovania klienta prebiehala už počas vývoja aplikácie, kedy bolo potrebné vždy kontrolovať funkčnosť a určiť chyby, ktoré majú negatívny vplyv na výkon alebo použiteľnosť aplikácie. Pri vývoji boli využité tri prehliadače, Google Chrome, Firefox a Microsoft Edge. Hlavným prehliadačom bol Google Chrome, ktorý poskytuje aj nástroje pre testovanie webových stránok Lighthouse. Je to open-source projekt, ktorý umožňuje vývojárom a testerom vyhodnocovať kvalitu webových stránok pomocou rôznych parametrov. Lighthouse poskytuje detailné analýzy webových stránok vrátane výkonnosti, prístupnosti alebo najlepších postupov SEO. Výsledok testovania pomocou tohto nástroja je zachytený na Obr.3.

Test prebehol na stránke vyhľadávania komúní, kde výkon splňal všetky potrebné parametre. Dostupnosť je vyznačená oranžovou farbou. Tieto kontroly zvyčajne zdôrazňujú príležitosti na zlepšenie dostupnosti webovej aplikácie. Najlepšie praktiky sú ukazovateľom toho, či daná webová aplikácia využíva moderné spôsoby vývoja webových aplikácií. V poslednom bode je poukázané, že aplikácia je optimalizovaná pre vyhľadávače.



Obr. 3 Výsledky testovania nástrojom Lighthouse.

V. ZÁVER

V závere tejto práce možno konštatovať, že boli úspešne naplnené jej ciele. Bola vytvorená plne funkčná sociálna sieť pre profesionálov s jednoduchým a intuitívnym používateľským rozhraním, ktorá umožňuje používateľom diskutovať, zdieľať informácie a budovať nové kontakty. Aplikácia bola testovaná a overená a môže byť používaná v reálnom profesionálnom prostredí. Tento projekt môže prispieť k rozvoju profesionálneho prostredia a zlepšeniu spolupráce medzi rôznymi odborníkmi.

POĎAKOVANIE

Táto publikácia bola publikovaná s podporou Operačného programu integrovaná infraštruktúra: Research in the SANET Network and Possibilities of Its Further Use and Development (ITMS code: 313011W988), spolufinancovaná ERDF a s podporou projektu KEGA 002TUKE-4/2021 Implementation of Modern Methods and Education Forms in the Area of Cybersecurity towards Requirements of Labour Market.

REFERENCIE

- [1] J. Scott, P.J. Carrington, "The SAGE Handbook of Social Network Analysis", SAGE Publications, 2011. ISBN 9781446250112.
- [2] V. Bednář, "Marketing na sociálních sítích", Computer Press, Brno, Česká republika, 2011. ISBN 978-80-251-3320-0.
- [3] M. Bočák, J. Rusnák, "Médiá a text 3, Mediálny text: variácie mediálneho diskurzu - popkultúra", Prešovská univerzita v Prešove, Prešov, 2012. ISBN 978-80-555-0667-8.
- [4] A. Gale, "Myspace, the Rise and Fall of an Internet Empire", 2023, Dostupné tiež z: <https://greekreporter.com/2023/01/03/myspace-fall/>. [cit. 19.03.2023].
- [5] M. McLuhan, "Výbor z díla. Člověk, média a elektronická kultura", Jota, 2000. ISBN 8072171283.
- [6] T. Bouda, BOUDA, "Sociální média", Inflow: information journal. 2010, roč.2, č.8. ISSN 1802-9736.
- [7] L. Stanescu a R. Stanescu, Social Medical Network, International Conference on System Theory Control and Computing, Sinaia, Romania, Oct. 19-21, 2022, pp. 152-156., DOI: <https://doi.org/10.1109/ICSTCC55426.2022.9931896>
- [8] MZ. Li, X. Li, Y. Zhang, JG. Pan a R. Yang, Implementation of a Social Network Information Dissemination Model Incorporating Negative Relationships, Computing and Informatics, Volume 41, Issue 6, pp. 1510-1540, 2022, ISSN: 1335-9150, https://doi.org/10.31577/cai_2022_6_1510
- [9] W. He, Y. Han, XF. Hu, W. Liu, BY. Yang a HZ. Chen, From idea endorsement to idea implementation: A multilevel social network approach toward managerial voice implementation, Sage Publications, London, England, Nov. 2020, Volume 73, Issue 11, pp. 1563-1582, ISSN: 0018-7267, eISSN: 1741-282X, <https://doi.org/10.1177/0018726719882999>

Modelovanie Superkapacitorov pre Aplikácie Hybridných Zásobníkov Energie

¹Dávid BODNÁR, ²František ĎUROVSKÝ

¹ Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

² Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹david.bodnar@tuke.sk, ²frantisek.durovsky@tuke.sk

Abstrakt — Superkapacitorové technológie (SC) sa začínajú objavovať v elektrických vozidlách (EV) a stacionárnych aplikáciách na ukladanie energie. Často sú prepojené s inými technológiami skladovania energie, ktoré spolu vytvárajú hybridný systém skladovania energie (HESS). Modelovanie difúzných procesov v SC však naďalej zostáva náročné. V tomto článku bol urobený stručný prehľad možných modelov superkapacitorov. Použitý bol model ekvivalentného elektrického obvodu a bol vytvorený pomocou Simscape Electrical v prostredí Matlab/Simulink. Identifikácia parametrov bola vykonaná pre 300 F SC článok a výstup simulácie bol porovnaný s nameranými údajmi. Výsledky ukázali vysokú presnosť SC modelu. Vytvorený model je možné použiť v BMS systémoch SC modulov, ale aj na účely dimenzovania hybridných zásobníkov energie.

Kľúčové slová — Model ekvivalentného elektrického obvodu, samovybíjanie, Simscape electrical, superkapacitor

I. ÚVOD

V posledných rokoch získavajú superkapacity väčšiu pozornosť pre svoju vysokú hustotu výkonu, spoľahlivú prevádzku, dlhú životnosť a nízky vnútorný odpor [1]. Existujú tri možné typy SC: elektrický dvojvrstvový kondenzátor (EDLC), hybridné kondenzátory a pseudo kondenzátory [2]. V tomto článku sa EDLC bude označovať ako SC.

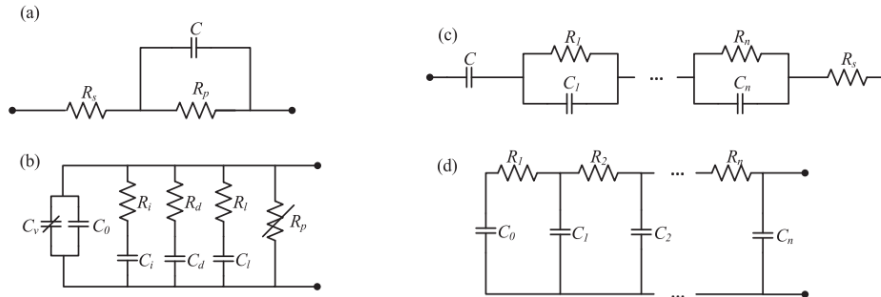
V SC sú obe elektródy vyrobené z aktivovaného uhlíka. Ich vybíjacia charakteristika je takmer lineárna [3]. Na modelovanie elektrického správania sa SC sa najčastejšie používajú elektrochemické modely, modely ekvivalentných elektrických obvodov (ECM) alebo modely neceločíselného rádu [4]. Môžu sa použiť aj modely umelej inteligencie, ako sú napríklad neurónové siete [5]. SC modelovanie je v porovnaní s modelovaním batérie o niečo jednoduchšie, keďže v SC neprebiehajú žiadne chemické reakcie a napätie SC je približne priamo úmerné stavu nabitia SC. Modelovanie SC má však stále svoje výzvy, ako je napríklad modelovanie difúzných procesov, samovybíjania alebo tepelné modelovanie.

II. ECM SUPERKAPACITORA

ECM modely sú v SC modelovaní pomerne rozšírené vďaka jednoduchej implementácii. Najpoužívanejšie z nich sú znázornené na obr. 1. Najjednoduchší ECM model SC je rezistor zapojený do série s kondenzátorom a jedným paralelným rezistorom reprezentujúcim samovybíjanie (obr. 1 a)). Ďalším spôsobom je kondenzátor pripojený paralelne k niekoľkým paralelným RC vetvám, ktoré predstavujú difúzne procesy v SC článku (obr. 1 b)). RC páry môžu byť zapojené aj do série s kondenzátorom (obr. 1 c)). Rovnaký model ECM je teda možné použiť pre SC aj batérie. Posledná metóda ECM je znázornená na obr. 1 d), ktorá bola vytvorená na simuláciu kapacity pórovitej elektródy a odporu elektrolytu. Táto metóda (obr. 1 d) je však dosť náročná na identifikáciu parametrov.

Samovybíjanie je spôsobené dvoma rôznymi mechanizmami, a to difúziou iónov a zvodovým prúdom [4]. Difúzia iónov je vlastne redistribúcia iónov v rámci elektródy kvôli pórovitosti elektródy SC. Počas nabíjania alebo vybíjania dochádza k nerovnomernému rozloženiu iónov v rámci elektródy, preto sa v rôznych častiach elektródy vyskytujú rôzne úrovne potenciálu. Redistribúciu iónov možno pozorovať počas obdobia odpočinku, kedy SC článok nie je zaťažený, ako pokles alebo zvýšenie napätia. Po určitom čase odpočinku sa ióny rovnomerne prerozdedia, rozdiely potenciálu v rôznych častiach elektródy sa minimalizujú a napätie SC zostáva relatívne stabilné. V prípade nabíjania článku dochádza k nárastu napätia, ale po ukončení nabíjania, kedy

je už článok v pokoji, dochádza k poklesu napätia v čase. Naopak, ak sa článok s rovnomerne rozloženými iónmi začne vybíjať, tak po ukončení vybíjania je pozorovaný nárast napätia SC článku [6]. Toto zvýšenie napätia sa tiež označuje ako samonabíjanie [7]. Tieto difúzne procesy sú zodpovedné za najvýznamnejšie zmeny napätia SC počas obdobia odpočinku a možno ich pozorovať hlavne počas prvých niekoľkých minút. Trvá však hodiny, kým SC opäť dosiahne rovnovážny stav. Difúzne procesy sú modelované pomocou RC párov, ako je znázornené na obr. 1 b) a c). Tie môžu mať veľmi veľký vplyv na celkové napätie článku, najmä pri SC s menšou kapacitou.

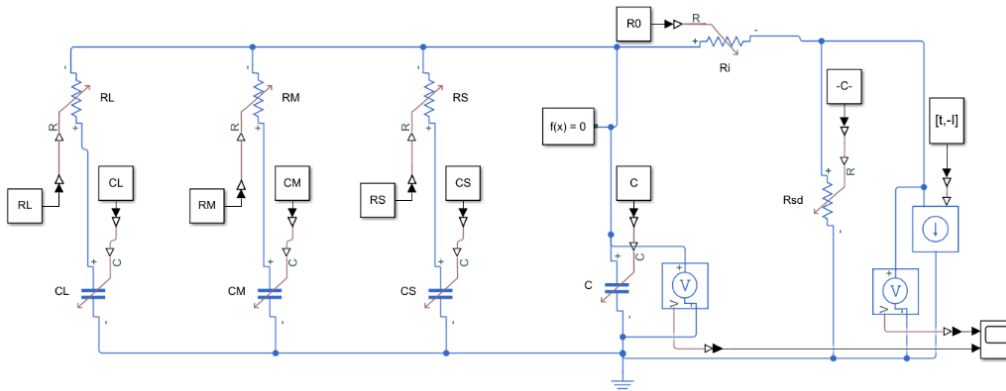


Obr. 1 ECM modely SC [4]

Samovybíjanie tiež pozostáva zo zvodového prúdu, pretože elektrolyt nemá nekonečne veľký odpor. Tento prúd je veľmi malý, ale pri dlhodobom skladovaní SC môže tiež výrazne znížiť napätie, respektíve stav nabitia SC. Zvodový prúd sa zvyčajne modeluje ako paralelne zapojený rezistor s relatívne vysokou hodnotou.

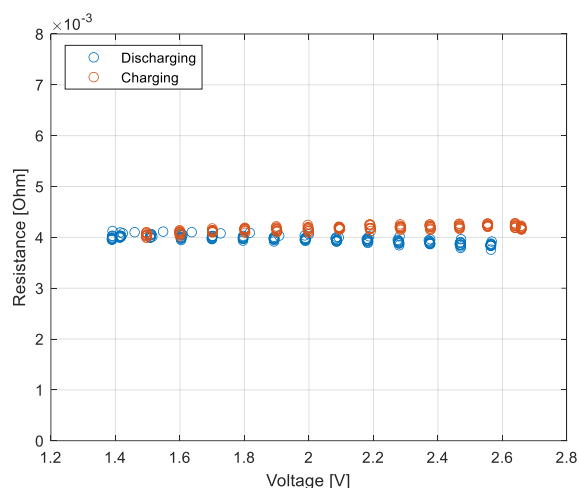
III. MODEL SC V PROSTREDÍ MATLAB/SIMULINK

Na vytvorenie modelu SC bol na identifikáciu parametrov vybraný typ SC článku SCCY62B307SSB. SC článok má menovité napätie 2,7 V a kapacitu 300 F. Všetky testy boli vykonané na testeri batériových modulov Arbin, model LBT22043. Na SC modelovanie sme použili Simscape Electrical v Matlabe (obr. 2). Zvolili sme ECM zložený z troch paralelne zapojených RC vetiev (obr. 1 b)), aby bolo možné modelovať aj difúzne procesy.



Obr. 2 ECM model SC v prostredí Matlab/Simulink

Identifikácia parametrov bola vykonaná pomocou Curve Fitting Toolbox v MATLAB a bol použitý model ECM tretieho rádu. Nameraná veľkosť sériového vnútorného odporu je znázornená na obr. 3 v závislosti od napätia, respektíve stavu nabitia (SoC). Vnútorný odpor sa vzhľadom od SoC mení len minimálne a pohybuje sa okolo hodnoty 4 mΩ. Ostatné parametre boli identifikované podľa napäťovej odozvy počas doby odpočinku, ktorá trvala jednu hodinu (obr. 4). Článok sa striedavo nabíjal na 2,7 V, a po odpočinku sa vybíjal na 1,35 V.



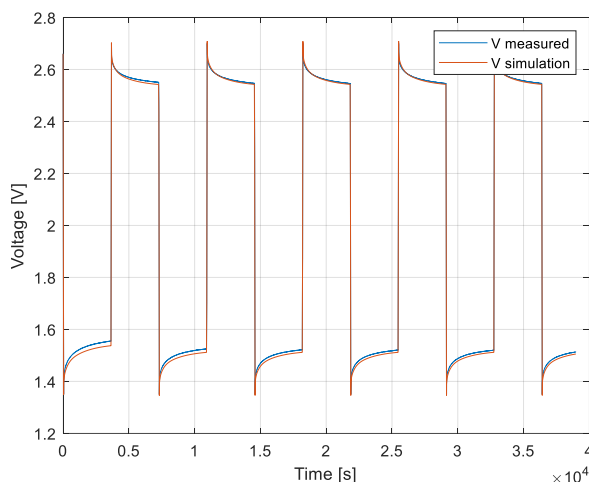
Obr. 3 Vnútny sériový odpor SC článku s kapacitou 300 F v závislosti od napätia SC

Všetky identifikované parametre sú zhrnuté v tabuľke 1, kde C predstavuje nameranú kapacitu SC článku, R_0 vnútorný sériový odpor, a ostatné parametre C_x a R_x predstavujú parametre paralelných vetiev, pričom indexom S sú označené parametre najkratšej časovej konštanty, indexom M parametre stredne veľkej časovej konštanty a indexom L dlhej časovej konštanty.

Takto vzniknutý model a jeho výsledky boli porovnané s nameraným priebehom napätia SC pod vyššie spomínaným priebehom záťažného prúdu. Na obr. 4 sú namerané priebehy porovnané s výsledkom simulácie a boli zaznamenané len nepatrné odchýlky.

Tabuľka 1 Identifikované parametre SC modelu

Parameter	C [F]	R_0 [Ω]	C_S [F]	R_S [Ω]	C_M [F]	R_M [Ω]	C_L [F]	R_L [Ω]
Hodnota	315	0,004	9,64	3,5	14,31	19,17	25,12	69,85



Obr. 4 Napät'ová odozva 300 F SC na vybíjacie a nabíjacie impulzy a porovnanie s výsledkami simulácie SC modelu

IV. ZÁVER

V tomto článku bol navrhnutý SC model v prostredí Matlab/Simulink. Výsledky simulácie ukazujú vysokú presnosť modelovania difúzných procesov. Vytvorený model je možné použiť v BMS systémoch SC a pri projektovaní HESS. Samovybíjanie SC spôsobené zvodovým prúdom nebolo v tomto článku skúmané, čo si vyžaduje ďalší výskum. Navyše, na posúdenie vplyvu teploty na samovybíjanie a difúziu je potrebné ďalšie odborné skúmanie. To si vyžaduje ďalšie experimenty s použitím teplotnej komory, ktorá by udržiavala konštantnú teplotu SC článku počas celej doby experimentu. Rozšírením tejto práce by mohlo byť aj modelovanie celého HESS vrátane ďalších zásobníkov energie ako je batéria a modelovanie ďalších potrebných komponentov HESS ako je DC/DC menič. Účinnosť všetkých použitých komponentov hrá kľúčovú úlohu v celkovej účinnosti HESS, preto posúdenie vplyvu SC na celkový systém zostáva predmetom ďalšieho štúdia.

POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe zmluvy č. APVV-15-0750 a zmluvy č. APVV-18-0436.

Tento výskum bol zároveň umožnený vďaka podpore DCE&S group na Technickej univerzite v Delfte, Holandsko.

REFERENCES

- [1] A. Khaligh and Zhihao Li, "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 Trans Veh Technol*, vol. 59, no. 6, pp. 2806–2814, Jul. 2010, doi: 10.1109/TVT.2010.2047877.
- [2] A. González, E. Goikolea, J. A. Barrena, and R. Mysyk, "Review on supercapacitors: Technologies and materials," *Renewable and Sustainable Energy Reviews*, vol. 58, pp. 1189–1206, May 2016, doi: 10.1016/j.rser.2015.12.249.
- [3] D. Slafstein, F. M. Ibanez, and K. Siwek, "Supercapacitor Modeling: A System Identification Approach," *IEEE Transactions on Energy Conversion*, pp. 1–11, 2022, doi: 10.1109/TEC.2022.3212617.
- [4] L. Zhang, X. Hu, Z. Wang, F. Sun, and D. G. Dorrell, "A review of supercapacitor modeling, estimation, and applications: A control/management perspective," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1868–1878, Jan. 2018, doi: 10.1016/j.rser.2017.05.283.
- [5] G. J. Adekoya, O. C. Adekoya, U. K. Ugo, E. R. Sadiku, Y. Hamam, and S. S. Ray, "A mini-review of artificial intelligence techniques for predicting the performance of supercapacitors," *Mater Today Proc*, vol. 62, pp. S184–S188, 2022, doi: 10.1016/j.matpr.2022.05.079.
- [6] J. Kowal *et al.*, "Detailed analysis of the self-discharge of supercapacitors," *J Power Sources*, vol. 196, no. 1, pp. 573–579, Jan. 2011, doi: 10.1016/j.jpowsour.2009.12.028.
- [7] J. J. Quintana, A. Ramos, M. Diaz, and I. Nuez, "Energy efficiency analysis as a function of the working voltages in supercapacitors," *Energy*, vol. 230, Sep. 2021, doi: 10.1016/j.energy.2021.120689.

Návrh metódy pre meranie zdravia stromov

¹Peter GIROVSKÝ, ²Daniel KORMANSKÝ

^{1,2} Katadra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach

¹peter.girovsky@tuke.sk, ²daniel.kormansky@student.tuke.sk

Abstrakt — V tomto príspevku predstavujeme možnosti využitia snímačov na detekovanie poškodenia kmeňa stromu. V úvodnej časti príspevku sú uvedené základné metódy pre určovanie zdravia stromov. Na základe prevedenej analýzy merania zdravia stromov, boli navrhnuté dva prístupy, ktoré boli následne aj otestované. V práci sú uvedené výsledky navrhovaných experimentálnych metód merania zdravia stromov pomocou termovízie a merania odporu, ktoré sa zameriavajú na detekciu dutín v kmeni stromu.

Kľúčové slová — snímače, meranie, diagnostika, termovízia, elektrický odpor

I. ÚVOD

Dnešná doba nám stále viac a viac ukazuje, aká je environmentalistika pre ľudí dôležitá. Zeleň v mestách, obytných oblastiach a rôznych mestských parkoch je neodlúčiteľnou súčasťou nášho života, nielen po estetickú, ale aj praktickú stránku. Stromy už roky slúžia ako prírodná klimatizácia, ktorá počas globálneho otepľovania ešte viac naberá na dôležitosti a s tým aj starostlivosť o stromy samotné. Čas vplýva na stromy rovnako, ako na nás, po určitej dobe sa môžu na stromoch objaviť rôzne poškodenia. Stromy môžu pri určitej úrovni poškodenia predstavovať veľké riziko. Najväčším problémom sú poškodenia, ktoré nie sú voľným okom viditeľné. Na zistenie a následnú diagnostiku skrytých poškodení sa využívajú viaceré metódy, technologické postupy ako aj rôzne typy snímacích zariadení.

Pri štandardnom diagnostikovaní a analyzovaní zdravia stromu sú definované určité parametre a postupy, ktorými sa hodnotí kondícia stromov rastúcich mimo lesa. Táto činnosť sa realizuje za účelom ochrany zdravia človeka, ochrany majetku a ochrany samotnej prírody. Hodnotenie zdravotného stavu stromu zohľadňuje široké spektrum faktorov vplývajúcich na stromy. Posudzovanie stavu stromov sa realizuje systémovým prístupom, ktorý pozostáva z niekoľkých krokov:

- Hodnotenie základných plôch
- Hodnotenie jednotlivých stromov
 - Súpis stromov
 - Dendrologický prieskum
 - Návrh pestovateľských opatrení
- Nadväzujúce a špecializované prieskumy

Individuálne hodnotenie jednotlivých stromov sa vykonáva v prípadoch kde je potrebné zhodnotiť samostatné stromy ktoré sú jednoznačne lokalizovateľné. Za účelom evidencie a zostavenia plánu kontroly, ale aj samotnej starostlivosti o stromy sa vypracováva hodnotenie spolu so súpisom stromov na každej evidovanej ploche zelene.

Pri súpise stromov sa v niektorých špeciálnych prípadoch môže vyžadovať aj evidencia pňov konkrétnych stromov spolu s ich parametrami (obvod a výška). Hodnotením a diagnostikou stavu stromu sa zaoberá dendrologický prieskum, ktorý sa realizuje pre množstvo rôznych účelov, či už ide o rekonštrukciu zelene, revitalizáciu prostredia, výrubové konanie a pod. Na základe dendrologického prieskumu sa následne môže meniť škála hodnotených atribútov drevín, pričom na zmenu hodnotených atribútov vplýva samotný stav stromu, nakoľko účel prieskumu je volený na základe kondície stromu. [1]

II. ANALÝZA ZDRAVIA STROMU

Nakoľko poškodenie stromu častokrát nevieme zhodnotiť voľným okom, využívajú sa na to prístrojové metódy, ktoré sa používajú prevažne na identifikáciu špecifických poškodení stromu. Tieto metódy sú zdrojom informácií pre ďalší postup analýzy. Pri používaní prístrojových metód sa berie ohľad na určité kritériá: [1]

- daná metóda nesmie poškodiť svojou montážou analyzovaný strom,
- nesmie dôjsť k poškodeniu beľového dreva stromu,
- stanovište stromu musí ostať nenarušené.

Výstupy z merania určujú úroveň poškodenia a na základe následnej analýzy sa zhodnotí riziko a pravdepodobnosť zlyhania stromu. Výber správnej prístrojovej metódy [7,8,9] pre meranie závisí od časti stromu, na ktorú má byť aplikovaná, a od atribútu poškodenia, ktorý chceme analyzovať. Za pomoci prístrojových metód sa vykonávajú štyri typy analýz: [1,2,3,4,6]

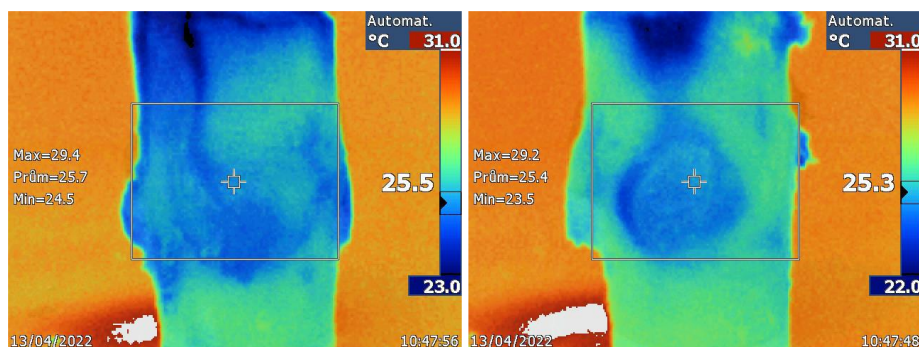
- 1) Analýza architektúry orgánov stromu
 - pôdny radar
 - akustické senzory
 - LIDAR
- 2) Analýza stavu kmeňa
 - akustické merania
 - akustická tomografia
 - ťahová skúška
- 3) Analýza stavu kmeňa a kostrových konárov
 - akustické merania
 - akustická tomografia
 - elektricko-impedančná tomografia
- 4) Analýza stability koreňového systému
 - ťahová skúška

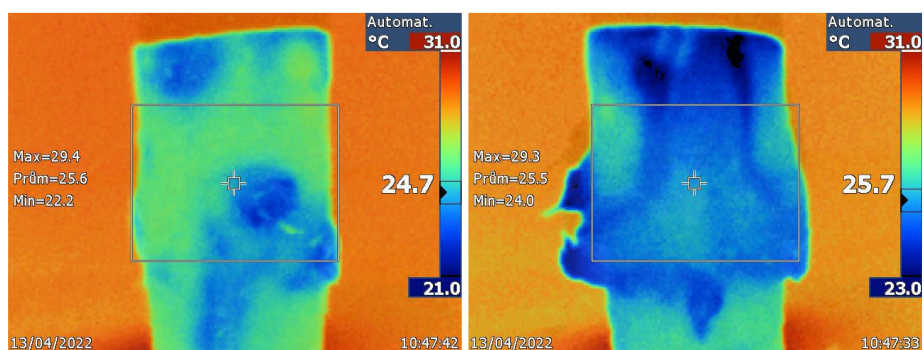
III. MERANIE POŠKODENIA KMEŇA NA ZÁKLADE VYBRANÝCH METÓD

V rámci riešenia práce sme sa zamerali na testovanie nových prístupov merania pre analýzu poškodenia kmeňa stromu. Pri meraní sme použili experimentálnu metódu pomocou snímania termovíznou kamerou a metódu, pri ktorej sme využili materiálové vlastnosti dreva ako izolantu.

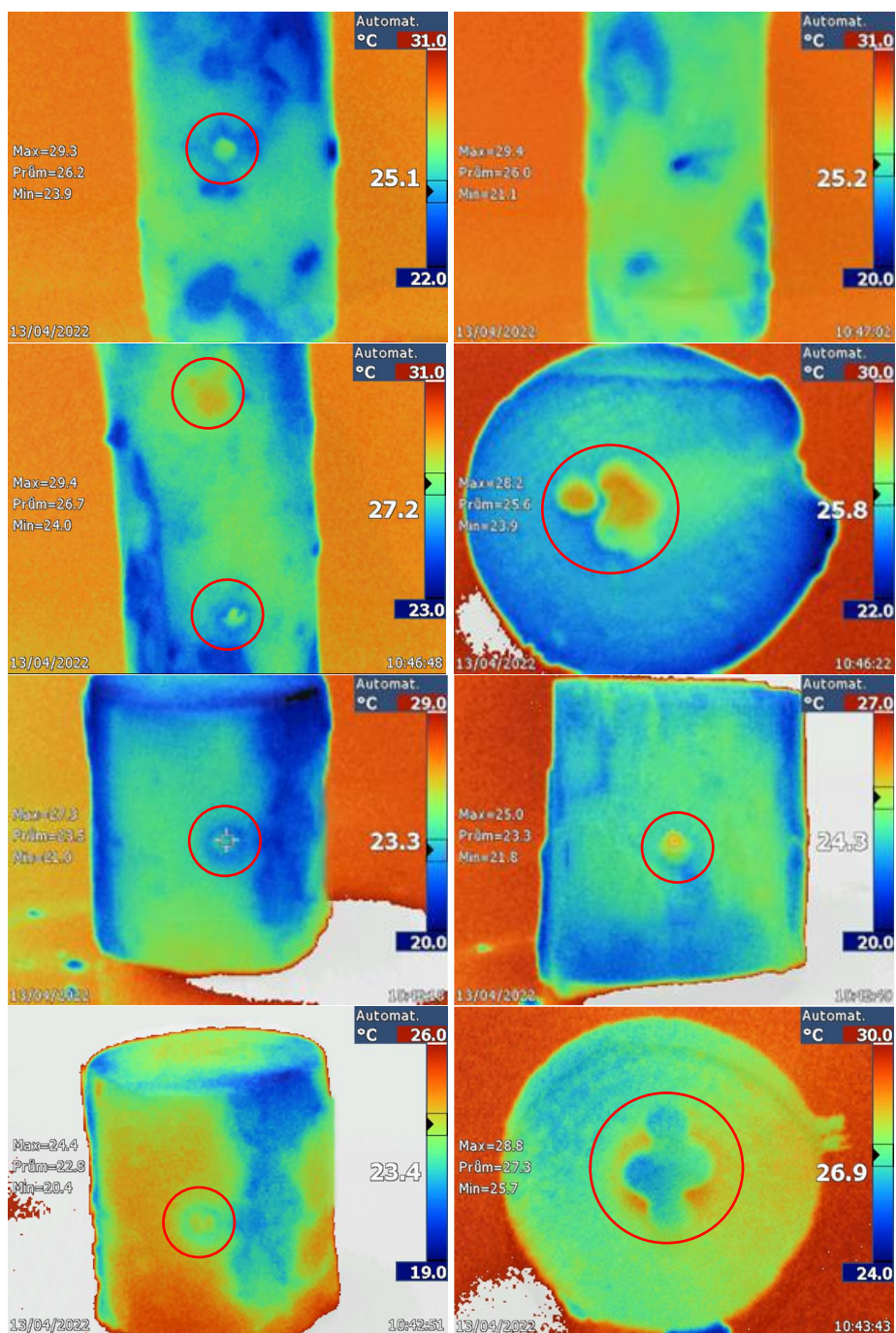
A. Analýza poškodenia kmeňa pomocou termovízie

V tomto prípade sme sa zamerali na snímanie infračerveného žiarenia, pričom jeho veľkosť závisí od povrchovej teploty predmetu a intenzity vyžarovania jeho povrchu. Následne prístroj detekuje množstvo energie vyžiarenej v infračervenom spektre predmetu a túto informáciu použije na odhad jeho teploty. Väčšina meraných materiálov, ako je lakovaný kov, drevo, voda, koža a textil, vyžarujú energiu veľmi efektívne a je ľahké získať z nich veľmi presné merania. Pri povrchoch, ktoré efektívne vyžarujú energiu (vysoká intenzita vyžarovania), sa faktor intenzity vyžarovania odhaduje na 95 % (čiže 0,95). Princíp celého merania je pomerne jednoduchý. Za pomoci termokamery FLUKE [5] sme na troch vzorkách pozorovali, či bude možné zachytiť teplotný rozdiel na miestach, kde boli úmyselne poškodené časti kmeňa tak, aby nám vznikli pomyselné dutiny. Snímali sme vzorky s poškodením obr.2, ktoré predstavovalo dutinu kmeňa s rôznym polomerom, hĺbkou a rôznym umiestnením. Následne sme tieto snímky porovnávali so snímkami urobenými na vzorke bez poškodenia, ktoré sú zobrazené na obr. 1.





Obr. 1 Termovízne snímky nepoškodeného kmeňa stromu.



Obr. 2 Termovízne snímky poškodeného kmeňa stromu.

Ako vidno z prezentovaných snímok, použitím termovízie je možno detekovať poškodenia stromu, avšak tieto poškodenia sa vo väčšine prípadov nachádzali blízko povrchu kmeňa. V prípade ak je poškodenie hlboko v kmeni stromu a nie je významné, táto metóda zrejme nebude

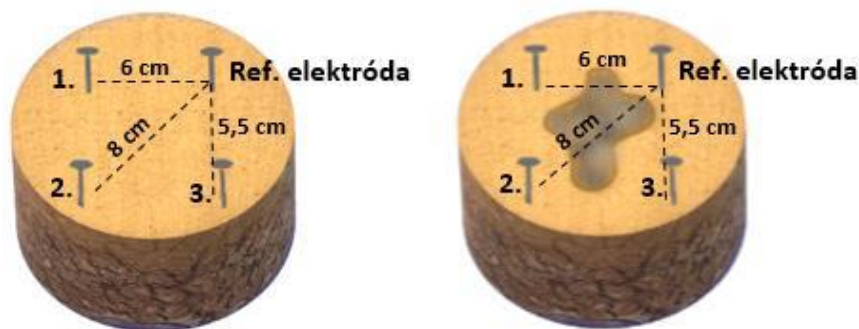
dostatočná na odhalenie takéhoto poškodenia. Výhodou tejto metódy je to že je úplne neinvazívna a teda ani najmenej nepoškodzuje kmeň stromu, ďalšou výhodou je, že na základe dostatočného počtu snímkov je možné vytvárať 3D model kmeňa stromu pre ďalšiu analýzu.

B. Analýza poškodenia kmeňa odporovým meraním

V tomto prípade sme sa zamerali na meranie izolačných vlastností dreanava, pričom sme vychádzali z predpokladu, že v oblasti poškodenia kmeňa stromu dôjde k zmene izolačných vlastností, teda k zmene elektrického odporu. Pre overenie presnosti merania boli použité priama a nepriama metóda merania elektrického odporu.

1) Priame meranie elektrického odporu

Na meranie sme použili prístroj HIOKI 3511-50 LCR HiTester, ktorý používa mostíkové zapojenie, ktoré je pripojené k zdroju striedavého napätia. Umiestnenie meracích elektród pri tomto meraní je zobrazené na obr. 3.



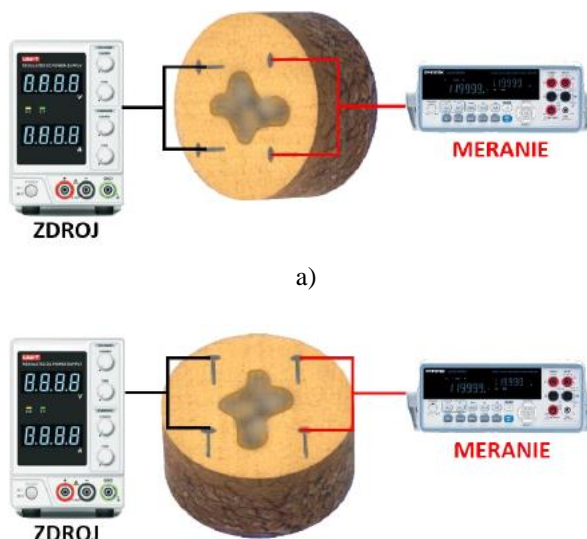
Obr. 3 Umiestnenie elektród pri priamom meraní kmeňa bez poškodenia a s poškodením.

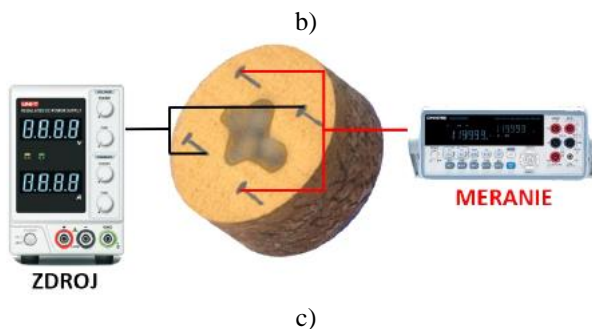
Tabuľka 1
 Namerané hodnoty elektrického odporu pri priamom meraní

Meranie na vzorke bez poškodenia		Meranie na vzorke s poškodením		Rozdiel hodnoty medzi vzorkami
Meranie medzi elektródami	Odpor	Meranie medzi elektródami	Odpor	
Ref. elektróda - 1	269,8kΩ	Ref. elektróda - 1	172,9kΩ	36,02%
Ref. elektróda - 2	7,706MΩ	Ref. elektróda - 2	269,2kΩ	96,51%
Ref. elektróda. - 3	354,7kΩ	Ref. elektróda - 3	171,7kΩ	51,76%

2) Nepriame meranie elektrického odporu

Pre porovnanie sme urobili aj meranie nepriamou metódou. V podstate išlo o podobný princíp ako pri predošlom meraní s tým rozdielom, že sme menili dvojicu napájajúcich elektród v troch rôznych zapojeniach zobrazených na obr.4. Merali sme V/A metódou pri napájacom napätí a následne bola vypočítaná hodnota elektrického odporu.





Obr. 4 Umiestnenie elektród pri nepriamom meraní odporu.

Tabuľka 2
 Namerané hodnoty odporu pri nepriamom meraní

	Meranie na vzorke bez poškodenia		Meranie na vzorke s poškodením		Rozdiel hodnoty medzi vzorkami
Zapojenie a)	U[V]	30	U[V]	30	39,89%
	I [mA]	0,076	I [mA]	0,1265	
	R[kΩ]	394,7	R[kΩ]	237,1	
Zapojenie b)	U[V]	30	U[V]	30	60,69%
	I [mA]	0,058	I [mA]	0,1475	
	R[kΩ]	517,2	R[kΩ]	203,3	
Zapojenie c)	U[V]	30	U[V]	30	97,87%
	I [mA]	0,002	I [mA]	0,094	
	R[kΩ]	15	R[kΩ]	319,1	

Oproti predhadzujúcej metóde merania, pri meraní elektrického odporu dochádza k čiastočnému poškodeniu kmeňa pri upevnení meracích elektród, avšak tieto poškodenia nie sú nijako významné ani trvalé. Z výsledkov merania uvedených v tabuľke 1 pre priame meranie a tabuľke 2 pre nepriame meranie, môžeme vidieť, že zmena odporu pri vnútornom poškodení stromu je významná a teda týmto spôsobom je možné detekovať hlavne poškodenie vo vnútri kmeňa stromu.

IV. ZÁVER

V rámci tejto práce sme sa zamerali na metódy merania zdravia stromov, ktoré sa využívajú v praxi. Na základe prevedenej analýzy meracích metód boli navrhnuté dva princípy merania, ktoré sme následne aj prakticky overili. Vzhľadom na komplexnosť analýzy zdravia stromu sme sa rozhodli venovať metódam, ktoré sa zameriavajú iba na meranie poškodení v konkrétnej časti stromu a to na detekciu dutín v kmeni stromu. Pre zisťovanie dutín v kmeni stromu sme zvolili metódu využívajúcu termovíziu a meranie elektrického odporu. Výsledky týchto meraní boli prezentované v tomto článku. V rámci prvej testovanej metódy sme využívali snímanie infračerveného žiarenia a na základe dosiahnutých výsledkov, ktoré boli prezentované v tomto článku môžeme konštatovať, že sme boli schopní detekovať poškodenia stromu pomocou tejto experimentálnej metódy. Pri využití tejto metódy by však bolo potrebné urobiť podrobnejšie merania a nastavenia termovíznej techniky pre vylepšenie tejto metódy a prípadné vytváranie komplexných 3D modelov, z ktorých by sme vedeli podrobnejšie identifikovať jednotlivé poškodenia. Avšak dosiahnuté výsledky pomocou tejto metódy nás privádzajú k záveru, že daná metóda má potenciál pre ďalšie využitie. Druhá navrhnutá metóda využívala meranie elektrického odporu. Aj pri tejto metóde sme dosiahli povzbudivé výsledky, nakoľko zmena elektrického odporu pri poškodenej a nepoškodenej vzorke bola výrazná a dokázali sme detekovať poškodenia vo vnútri kmeňa. Aj keď táto metóda nebola oproti predchádzajúcej neinvazívna, pri jej využití nedochádza k trvalému poškodeniu kmeňa stromu. V rámci tejto meracej metódy predpokladáme, že podrobnejšími meraniami na obvode stromu a následnou analýzou by bolo možné identifikovať vnútorné poškodenia.

POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-19-0210.

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy,

výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23.

LITERATÚRA

- [1] V. Paganová „Arboristický štandard : Hodnotenie stavu stromov“ : Slovenská poľnohospodárska univerzita v Nitre, 2019. 54 s. ISBN 978-80-552-2013-0
- [2] I. Virgala, E. Miková, M. Sukop, M. Vagaš, T. Lipták, M. Kelemen „Snímače a prevodníky“ : Technická univerzita v Košiciach, 2017. 183 s. ISBN 978-80-553-2879-9
- [3] J. Šturcel „Snímače a Prevodníky“ : STU v Bratislave, 2002. 262 s. ISBN 80-227-1712-6
- [4] M. Holubčík, L. Holubčíková, N. Kantová „Termovízia v praxi“ : Žilinská univerzita, Katedra energetickej techniky, CEIT a.s. 2015. 72 s. ISBN 978-80-971684-3-8
- [5] Fluke Corporation, 2007-2010. „Ti25 Thermal Imager Users Manual“ : August 2007, Rev.2, 8/10, Printed in USA. PN2803044, 26 s. dostupný na internete : <https://www.fluke-direct.com/pdfs/cache/www.fluke-direct.com/ti9-9hz/manual/ti9-9hz-manual.pdf>
- [6] X. Wang, J. Wiedenbeck, S. Liang. „Acoustic Tomography for Decay Detection in Black Cherry Trees“ : PKP/PS 2009. 137 s. dostupné na internete : <https://www.fs.usda.gov/treesearch/pubs/33263>
- [7] Fakopp Enterprise Bt. User Manual, 2020. „Manual for the ArborSonic3D acoustic tomograph“ . Pdf : 64 s. dostupný na internete : <https://www.fakopp.com/en/product/arborsonic/>
- [8] Fakopp Enterprise Bt. User Manual, 2018. „Manual for ArborElectro Impedance Tomograph“ . Pdf : 37 s. dostupný na internete : <https://www.fakopp.com/en/product/arborelectro/>
- [9] Argus electronic gmbh. User Manual, 2012. „ Manual for PiCUS: TreeTronic Electrical Resistance Tomograph for trees “ Pdf : 36 s. dostupný na internete : <https://www.argus-electronic.de/en/tree-inspection/support/download/manuals/picus-treetronic-r-version-1>

Modulárne robotické rameno

¹Peter GIROVSKÝ, ²Jaroslava ŽILKOVÁ, ³Peter NEMERGUT

^{1,2,3} Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,
Technická univerzita v Košiciach

¹peter.girovsky@tuke.sk, ²jaroslava.zilkova@tuke.sk, ³peter.nemergut@student.tuke.sk

Abstrakt — V tomto príspevku predstavíme návrh a konštrukciu modulárneho robotického ramena. Robotické rameno bolo nahnuté s ohľadom na jeho modulárnosť aby v prípade potreby bolo možné vytvoriť rameno s rôznym počtom stupňov voľnosti. V rámci návrhu je urobená aj jednoduchá statická analýza vzhľadom na pôsobiace sily v ramene. Diely robotického ramena boli vytlačené pomocou 3D tlačiarne. Riadenie robotického ramena zabezpečovala riadiaca jednotka určená pre pohony Dynamixel.

Kľúčové slová — robotické rameno, riadenie, konštrukcia, 3D tlač

I. ÚVOD

Zavádzanie automatizácie a robotizovaných pracovísk do priemyselnej výroby a technologických procesov sa stáva čoraz väčšou neoddeliteľnou súčasťou automobilových, strojárskych, elektrotechnických, medicínskych či potravinárskych priemyslov. Automatizácia so sebou prináša zefektívnenie výroby, vyššiu kvalitu, stabilitu a najmä zníženie manuálnej práce. Neustály vývoj automatizácie smeruje od robotizovaných pracovísk k vytváraniu priemyselných sieťových zoskupení čo vedie k vytváraniu plne automatizovaných inteligentných tovární.

Z uvedených dôvodov je dôležité sa danou problematikou čoraz viac zaoberať a venovať jej patričnú pozornosť. Už pri štúdiu na stredných či vysokých školách by sa študenti mali mať možnosť oboznámiť s problematikou automatizácie, základnými pojmami robotiky a pochopiť princípy robotizácie. Problémom je fakt, že sa to častokrát deje iba na teoretickej rovine a málokedy majú študenti možnosť stretnúť sa s konkrétnou technológiou. Je to spôsobené hlavne tým, že si školy nemôžu dovoliť takúto technológiu zadovážiť.

Je preto dôležité nachádzať iné spôsoby alebo alternatívy, ako študentom pomôcť pri štúdiu tejto problematiky a nachádzať možnosti, ako sa môžu aspoň čiastočne priblížiť k reálnym technológiám, ktoré by si mohli aj fyzicky vyskúšať, čím sa daná technológia pre nich stáva viac prístupná a dopomáha k celkovo lepšiemu pochopeniu danej témy.

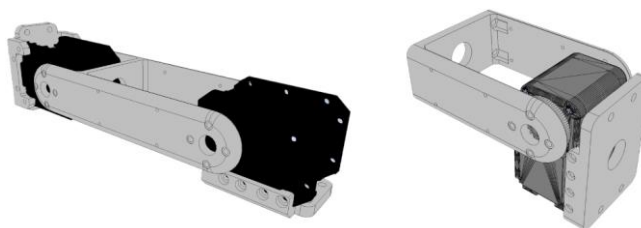
II. NÁVRH ROBOTICKÉHO RAMENA

Pri návrhu konštrukcie sme definovali konštrukciu robotického ramena so 6 stupňami voľnosti a kinematickou štruktúrou s reťazcom R-R-R-R-R-R, nakoľko sa v praxi najviac používa. Uloženie servopohonov Dynamixel, ktoré boli použité ako aktuátory, bude priamo v kĺboch z dôvodu zjednodušenia konštrukcie a modulárnosti. Parametre použitých servopohonov dynamixel sú uvedené v tabuľke 1.

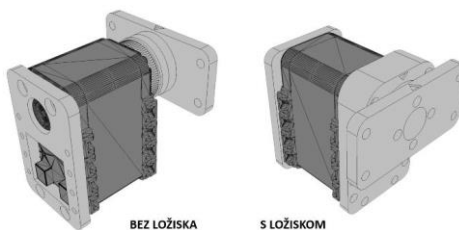
Tabuľka 1
Parametre použitých pohonov Dynamixel

	Servopohony Dynamixel		
	AX-12A	AX-18A	MX-28AT
Hmotnosť	54,6 g	55,9 g	77 g
Rozsah natočenia	0-300°	0-300°	0-360°
Rozlíšenie	0-1023	0-1023	0-4095
Blokovací moment	1,5 Nm (pri 12 V)	1,8 Nm (pri 12 V)	2,5 Nm (pri 12 V)
Uhlová rýchlosť naprázdno	59 otáčok/minúta (pri 12 V)	97 otáčok/minúta (pri 12 V)	55 otáčok/minúta (pri 12 V)
Prevod	254:1	254:1	193:1

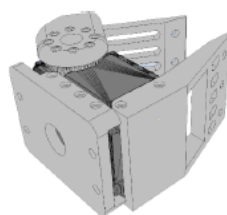
Uloženie zároveň bolo realizované tak, aby bolo čo najbližšie k osi rotácie, aby sme dosiahli čo najvyšší užitočný krútiaci moment. Modularita robotického ramena spočívala vo vyhotovení modulov, ktoré budú môcť byť medzi sebou spájané pomocou univerzálnej príruby v požadovanej konfigurácii. Jednotlivé časti pre vyhotovenie modulov boli vytlačené pomocou technológie 3D tlače. Navrhnuté moduly sú zobrazené na obr. 1-3.



Obr. 1 Modul s dlhým a krátkym ramenom



Obr. 2 Modul pre kolmú rotačnú os

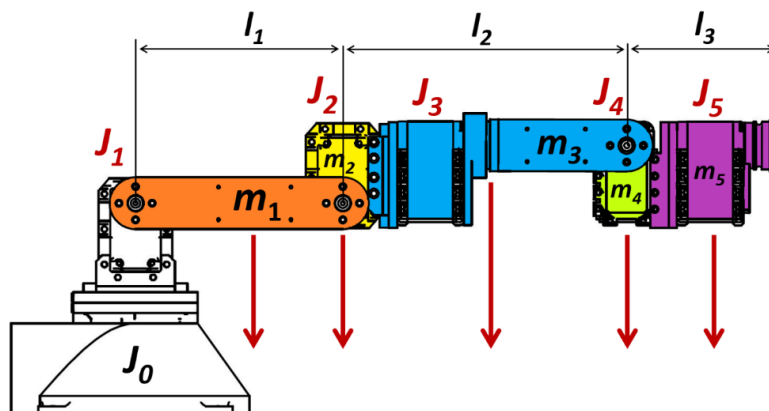


Obr. 3 Koncový efektor

Pre zvolenú testovaciu konštrukciu robotického ramena bola následne urobená statická analýza.

A. Statická analýza ramena

Pomocou statickej analýzy sme zisťovali, aký krútiaci moment pôsobí na servopohon, ktorý je upevnený v kĺbe J_1 . Bol uvažovaný najkritickejší stav t.j. ak budú všetky ramená vo vodorovnej polohe. Pri tejto analýze uvažujeme hmotnosti a dĺžky jednotlivých dielov tak, ako sú vyobrazené a farebne odlišené na obr. 4. Pre zjednodušenie modelu bol uvažovaný účinok pôsobenia sily sústredený v strede farebne vyznačenej časti ramena resp. servopohonu.[12, 13]



Obr. 4 Robotické rameno bez koncového modulu

Tabulka 2
 Parametre pre statickú analýzu

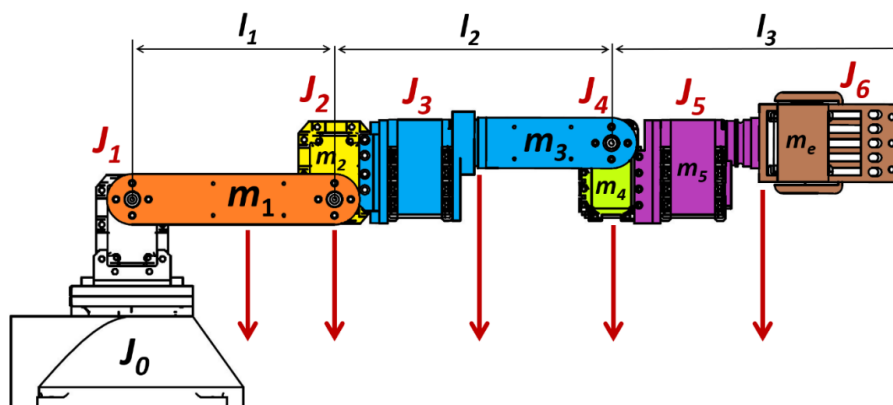
Parameter	Hodnota	Popis
m_1	36 g	rameno s hmotnosťou m_1
m_2	82 g	servopohon v kĺbe J_2 + skrutky
m_3	128 g	servopohon v kĺbe J_3 + ložisko + s hmotnosťou m_3
m_4	59 g	servopohon v kĺbe J_4 + skrutky
m_5	86 g	servopohon v kĺbe J_5 + prírubou + skrutky
m_e	89 g	koncový efektór
l_1	10 cm	vzdialenosť medzi kĺbmi J_1 a J_2
l_2	13,7 cm	vzdialenosť medzi kĺbmi J_2 a J_4
l_3	7,3 cm / 14,45 cm	vzdialenosť medzi kĺbmi J_4 a prírubou J_5 / vzdialenosť medzi kĺbmi J_4 a J_6

V Tab. 2 sú uvedené vstupné parametre, kde m je hmotnosť vyznačených častí odlišených farbou a l dĺžka.

Výpočet krútiaceho momentu, ktorý pôsobí na servopohon v kĺbe J_1 :

$$M_1 = \frac{l_1}{2} \cdot m_1 + l_1 \cdot m_2 + \left(l_1 + \frac{l_2}{2}\right) \cdot m_3 + (l_1 + l_2) \cdot m_4 + \left(l_1 + l_2 + \frac{l_3}{2}\right) \cdot m_5 \quad (1)$$

Výpočtom následne dostávame hodnotu krútiaceho momentu pre robotické rameno bez koncového efektora $M_1=0,678Nm$.



Obr. 5 Robotické rameno s koncovým modulom

Doplnením koncového modulu, ktorý sa môže v závislosti od aplikácie meniť, prepočítame krútiaci moment pôsobiaci na servopohon v kĺbe J_1 :

$$M_1 = \frac{l_1}{2} \cdot m_1 + l_1 \cdot m_2 + \left(l_1 + \frac{l_2}{2}\right) \cdot m_3 + (l_1 + l_2) \cdot m_4 + \left(l_1 + l_2 + \frac{l_3}{2}\right) \cdot (m_5 + m_e) \quad (2)$$

V tomto prípade dostávame hodnotu krútiaceho momentu $M_1=0,978Nm$. Obdobne môžeme vyjadriť aj momenty, ktoré pôsobia na servopohony v kĺboch J_2 a J_4 :

$$M_2 = \frac{l_2}{2} \cdot m_3 + l_2 \cdot m_4 + \left(l_2 + \frac{l_3}{2}\right) \cdot (m_5 + m_e) \quad (3)$$

$$M_4 = \frac{l_3}{2} \cdot (m_5 + m_e) \quad (4)$$

Z výpočtov dostávame potrebné hodnoty momentov v jednotlivých kĺboch robotického ramena bez záťaže $M_2=0,525Nm$ a $M_4=0,124Nm$. Keďže ako servopohon v kĺbe J_1 bol zvolený servomotor MX-28AT, ktorého maximálny krútiaci moment je $M_m=2,5Nm$, dostaneme spätným výpočtom hodnotu maximálnej záťaže pre dané robotické rameno.

$$M_m = \frac{l_1}{2} \cdot m_1 + l_1 \cdot m_2 + \left(l_1 + \frac{l_2}{2}\right) \cdot m_3 + (l_1 + l_2) \cdot m_4 + \left(l_1 + l_2 + \frac{l_3}{2}\right) \cdot (m_5 + m_e) + (l_1 + l_2 + l_3) \cdot m_t \quad (5)$$

Z rovnice (5) následne dostaneme hodnotu maximálnej záťaže $m_t=407g$.

III. PRAKTICKÉ OVERENIE

Pri praktickom overení funkčnosti robotického ramena so 6 stupňami boli jednotlivé moduly radené tak, aby sa výsledná konštrukcia čo najviac podobala skutočnému konštrukčnému usporiadaniu robotického ramena používaného v priemyselnej výrobe. Konštrukcia robotického ramena je zobrazená na obr.6.



Obr. 6 Robotické rameno so šiestimi stupňami voľnosti

Na riadenie a ovládanie ramena bola použitá riadiaca jednotka CM-530. Táto riadiaca jednotka obsahuje konektory pre priame pripojenie servopohonov a ich napájanie, 6 tlačidiel a vypínač. Na riadiacej jednotke sa nachádza konektor pre komunikáciu s počítačom (mini USB). Táto riadiaca jednotka môže byť napájaná z externého zdroja (v našom prípade s parametrami DC 12V, 5A) alebo z batérie s typizovaným konektorom. Riadiaci program pre testovanie robotického ramena bol vytvorený v prostredí RobotPlus .

IV. ZÁVER

V tomto článku sme sa venovali návrhu a zostrojeniu konštrukcie modulárneho robotického ramena, ktoré je možné použiť vo výučbe. Pri návrhu sme vychádzali z poznatkov o aktuálne používaných kinematických štruktúrach, spôsobov ovládania a programovania, ktoré sú využívané v priemyselnej praxi. Zároveň sme sa snažili vyhotoviť modulárne robotické rameno tak, aby bolo cenovo dostupné a bolo možné ho kedykoľvek opätovne vyhotoviť.

Modulárne robotické rameno bolo vyhotovené pomocou technológie 3D tlače. Môžeme konštatovať, že vyskladané robotické rameno má požadovanú tuhosť a dostatočnú presnosť pre potreby výučby. Zároveň možnosť riadenia servopohonov Dynamixel nie len pomocou uvedenej riadiacej jednotky, ale pomocou počítača, prípadne mikropočítačov umožňuje navrhovať a testovať vlastné riadiace algoritmy pre zostavené robotické rameno.

POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-19-0210.

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23.

LITERATÚRA

- [1] KOVÁČ, Jozef, Rudy Vladimír a Kováč Juraj. Automatizácia výroby. 1. vyd. Košice: TU, 2016. 304 s. ISBN 978-80-553-2311-4 (brož.)
- [2] ŠTEFANÁK, Peter. Konštrukcia priemyselných robotov a manipulatorov. 1. vyd. Bratislava: Alfa, 1985. 266 s. ISBN (brož.)
- [3] PALKO, Anton a Smrček Juraj. Robotika: koncové efekty pre priemyselné a servisné roboty: navrhovanie, konštrukcia, riešenia. Košice: TU, 2004. 272 s. ISBN 80-8073-218-3 (viaz.)
- [4] VITRALAB, Leonardo da Vinci Programme, Automatizačná a robotická technika [online]. Košice, 09.2011. Verzia dokumentu LLP/LDV/TOI/2009/SK/93100530 [cit. 11.05.2022]. Dostupné na internete: http://www.sjf.tuke.sk/vitralab/files/vystupy/prirucka_sk_final.pdf
- [5] Robotické manipulatory – priemyselné roboty [online]. Spišská Nová Ves: Stredná priemyselná škola technická. [cit. 11.05.2022]. Dostupné na internete: https://cloud2q.edupage.org/cloud/3-Roboticke_manipulatory_-_priemyselne_roboty.pdf?z%3AAAE6h2I%2FoZJhJTST2bRCJpTFwyZRyChUTHXlseRbEoNM2h5nfgZcl6zAYY05980xD

- [6] BELOVEŽČÍK, Andrej. Základné kinematické štruktúry [online]. Košice: Katedra robotiky, Strojnícka fakulta TU, 2018. [cit. 11.05.2022]. Dostupné na internete: <https://kvtar.abell.sk/cms/index.php/sk/start/zakladne-kinematicke-dvojice>
- [7] JURÍŠICA, Ladislav, Hubinský, Peter a Kardoš Ján. Robotika [online]. Bratislava: Fakulta elektrotechniky a informatiky, Slovenská technická univerzita, 2005. [cit. 11.05.2022]. Dostupné na internete: https://www.researchgate.net/publication/267783300_Robotika
- [8] Články. Definícia, štruktúra a funkcie priemyselného robota [online]. Bardejov: Stredná škola Juraja Henischa. Posledná zmena 08.03.2021.[cit. 12.05.2022].Dostupné na internete: https://ssjh.sk/dokuwiki/doku.php/mnk/definicia_struktura_a_funkcie_priemyselneho_robota
- [9] Články. Robotshop.com: Basics: How Do I Choose a Robotic Arm? [online]. Tutorials. Posledná zmena 26.07.2021. [cit. 11.05.2022]. Dostupné na internete: <https://www.robotshop.com/community/tutorials/show/basics-how-do-i-choose-a-robot-arm>
- [10] WANG, Xiaoyu a Baron, Luc. Topology and Geometry of Serial and Parallel Manipulators [online]. Canada: Polytechnique of Montreal, 04.2008. [cit. 11.05.2022]. Dostupné na internete:https://www.researchgate.net/publication/221786847_Topology_and_Geometry_of_Serial_and_Parallel_Manipulators/link/02e7e52c442c4be9cf000000/download
- [11] DAILY AUTOMATION. DELTA Roboty [online]. [cit. 11.05.2022]. Dostupné na internete: <https://www.dailyautomation.sk/delta-roboty/>
- [12] MEDEIROS, Alex. Robotic Arms, Theory Forces [online]. Web Project about Robotic arms, 2014. [cit. 11.05.2022]. Dostupné na internete: http://ffden-2.phys.uaf.edu/webproj/211_fall_2014/Alexander_Medeiros/page2.html
- [13] ADEBOLA, O. Simeon. A locally sourced robotic arm [online]. Nashville: Department of Teaching and Learning, Vanderbilt University, 16.12.2012. [cit. 11.05.2022]. Dostupné na internete: <https://roboarmblog.com/2012/09/16/a-locally-sourced-robotic-arm-2/>
- [14] DAILY AUTOMATION. Pohyby robota [online]. [cit. 11.05.2022]. Dostupné na internete: <https://www.dailyautomation.sk/06-programovanie-robotov-pohyby-robota/>
- [15] OWEN-HILL, Alex. What Are the Different Programming Methods for Robots? [online]. Článok. Posledná zmena 27.07.2021. [cit. 11.05.2022]. Dostupné na internete: <https://blog.robotiq.com/what-are-the-different-programming-methods-for-robots>
- [16] DAILY AUTOMATION. Online programovanie robotov [online]. [cit. 11.05.2022]. Dostupné na internete: <https://www.dailyautomation.sk/online-programovanie-robotov/>
- [17] <https://www.automate.org/news/reducing-the-variables-getting-accurate-motion-from-inaccurate-robots>
- [18] PRIMATICS. Torque Conversion [online]. [cit. 12.05.2022]. Dostupné na internete: https://www.primatics.com/wp-content/uploads/2020/01/primatics_unit_conversions.pdf
- [19] ISO 8373:2021. Robotics – Vocabulary. Third edition. Geneva: ISO copyright office, 2021. [online]. Dostupné na internete: <https://www.iso.org/obp/ui/#iso:std:iso:8373:ed-3:v1:en>
- [20] ACKERMAN, Evan. iRobot Brings Robot Modules to Modular Robots. [online]. Článok. Posledná zmena 28.03.2011. [cit. 23.05.2022]. Dostupné na internete: <https://spectrum.ieee.org/imobot-brings-robot-modules-to-modular-robots>
- [21] CANONICAL ROBOTS. D module 14 [online]. [cit. 23.05.2022]. Dostupné na internete: <https://www.canonicalrobots.com/en/replacement-parts/2023/module-dual-axis-14-hans-robot-detail>
- [22] VACHEK, Jaroslav. Fyzika: Prehľad učiva základnej školy. 1. vyd. Bratislava: Slovenské pedagogické nakladateľstvo, 1979. 328 s. ISBN (viaz.)
- [23] ROBOTIS. Dynamixel [online]. [cit. 23.05.2022]. Dostupné na internete: <https://www.robotis.us/dynamixel/>
- [24] ROBOTIS. Dynamixel-P [online]. [cit. 23.05.2022]. Dostupné na internete: <https://www.robotis.us/dynamixel-p/>
- [25] ROBOTIS. Dynamixel-X [online]. [cit. 23.05.2022]. Dostupné na internete: <https://www.robotis.us/x-series/>
- [26] ROBOTIS e-Manual. MX-28T/R/AT/AR [online]. [cit. 23.05.2022]. Dostupné na internete: <https://manual.robotis.com/docs/en/dxl/mx/mx-28/>
- [27] ROBOTIS e-Manual. AX-12A [online]. [cit. 23.05.2022]. Dostupné na internete: <https://manual.robotis.com/docs/en/dxl/ax/ax-12a/>
- [28] ROBOTIS e-Manual. AX-18A [online]. [cit. 23.05.2022]. Dostupné na internete: <https://manual.robotis.com/docs/en/dxl/ax/ax-18a/>
- [29] Články. 3dskladom.sk: Správny výber filamentu je základ úspechu [online]. Posledná zmena 28.02.2021. [cit. 23.05.2022]. Dostupné na internete: <https://www.3dskladom.sk/vyber-druh-filamentu>
- [30] ROBOTIS e-Manual. Integrated Sensor (AX-S1) [online]. [cit. 23.05.2022]. Dostupné na internete: <https://manual.robotis.com/docs/en/parts/sensor/ax-s1/>
- [31] ROBOTIS e-Manual. CM-530 [online]. [cit. 23.05.2022]. Dostupné na internete: <https://manual.robotis.com/docs/en/parts/controller/cm-530/>
- [32] ROBOTIS shop. Software [online]. [cit. 23.05.2022]. Dostupné na internete: http://en.robotis.com/service/downloadpage.php?ca_id=10
- [33] ROBOTIS e-Manual. R+ Manager 1.0 [online]. [cit. 23.05.2022]. Dostupné na internete: <https://manual.robotis.com/docs/en/software/rplus1/manager/>
- [34] ROBOTIS e-Manual. R+ Motion 1.0 [online]. [cit. 23.05.2022]. Dostupné na internete: <https://manual.robotis.com/docs/en/software/rplus1/motion/>
- [35] ROBOTIS e-Manual. R+ Task 1.0 [online]. [cit. 23.05.2022]. Dostupné na internete: https://manual.robotis.com/docs/en/software/rplus1/task/getting_started/

Robotické rameno pre triedenie objektov

¹Peter GIROVSKÝ, ²Jaroslava ŽILKOVÁ, ³Milan JANO

^{1,2,3} Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach

¹peter.girovsky@tuke.sk, ²jaroslava.zilkova@tuke.sk, ³milan.jano@student.tuke.sk

Abstrakt — Cieľom tohto príspevku je predstaviť návrh a implementáciu algoritmov pre rozpoznávanie základných geometrických tvarov (kruh, štvorec, trojuholník) a farby (červená, zelená, modrá), ktoré sú využité pre automatické triedenie objektov pomocou robotického ramena. Pre návrh algoritmu na rozpoznávanie objektov ako aj pre návrh riadiaceho algoritmu robotického ramena bol použitý MATLAB. Robotické rameno bolo zostavené zo servopohonov Dynamixel a pre nastavenie a ovládanie programu pre klasifikáciu a triedenie objektov bolo v prostredí MATLAB vytvorené grafické užívateľské rozhranie.

Kľúčové slová — robotické rameno, rozpoznávanie objektov, servopohon, strojové videnie

I. ÚVOD

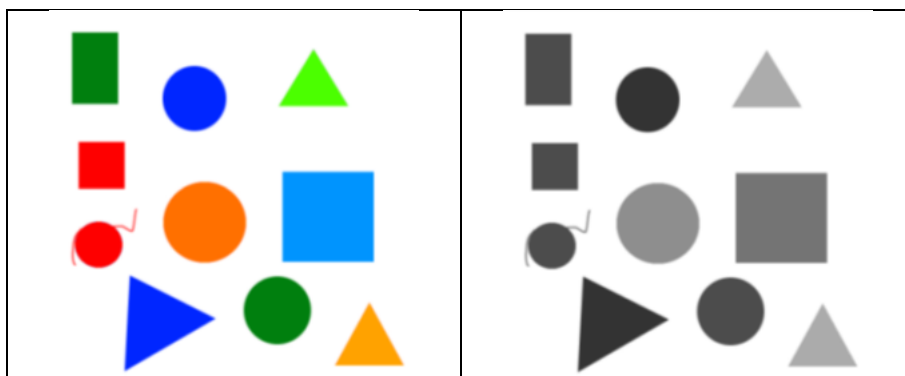
Automatizácia výroby je trendom, ktorý bude pokračovať aj naďalej a zdokonaľovaním jednotlivých techník a metód využívaných pri automatizovaných procesoch, bude možné v budúcnosti nahrádzať ľudskú prácu aj pri zložitých operáciách, ktoré sa doposiaľ nepodarilo nahradiť technikou. Jednou takouto oblasťou je oblasť počítačového videnia, resp. rozpoznávania objektov, tvarov a pod. Aj keď kvalitu ľudského oka je ťažko napodobniť, Priemyselné kamerové systémy takéto úlohy zvládajú nepretržite, pričom zároveň odbremeňujú človeka a chránia jeho zrak. Vlastnosti akými sú presnosť, rýchlosť, robustnosť a opakovateľnosť zaradujeme medzi hlavné výhody využitia priemyselných kamerových systémov. Príkladom môžeme byť napríklad kontrola kvality súčiastok na dopravnom páse. Táto kontrola môže byť veľmi rýchla, rádovo sa jedná o milisekundy, vykonávať sa môže nepretržite a s veľmi malou chybovosťou. Vo výsledku to znamená možnosť zrýchlenia kontroly kvality a teda možnosť zrýchlenia výrobného procesu a zvýšenia kvality produktov.

II. ROZPOZNÁVANIE OBJEKTOV

Rozpoznávanie objektov pomocou počítačovej techniky, resp. počítačové videnie pozostáva z niekoľkých krokov. V prvom kroku je potrebné získať obraz, ktorý sa bude následne digitálne spracovávať a vyhodnocovať. Ako už bolo spomenuté vyššie, na riešenie tejto úlohy bol využívaný program MATLAB a na spracovanie obrazu z webkamery bol použitý Image Processing Toolbox.

Pri digitálnom spracovaní obrazu je odstránenie šumu jednou z techník predbežného spracovania získaného obrazu. Obrazový šum sa prejavuje ako náhodná zmena informácie o jase alebo farbe snímaného objektu, pričom existujú rôzne typy obrazového šumu. Jedným z nich je napr. náhodný výskyt čiernych a bielych pixelov na odstránenie ktorých sa využíva mediánové filtrovanie. Mediánové filtrovanie predstavuje nelineárnu operáciu, ktorá sa často používa pri spracovaní obrazu na zníženie takéhoto šumu, pričom je efektívnejšie ako konvolúcia, v prípade ak je cieľom súčasne znížiť šum a zachovať okraje.

Ďalším krokom v oblasti spracovania obrazu je jeho segmentácia. Segmentácia obrazu je proces, pri ktorom sa obraz rozdelí do viacerých častí, nazývaných segmenty. Pri segmentácii obrazu sa jednotlivým pixelom, ktoré spolu súvisia (pixely ktoré v obraze reprezentujú jeden samostatný objekt), priraduje rovnaké označenie. Pixely s rovnakým označením, potom vytvárajú jeden segment nazývaný tiež „*superpixel*“. Segmentáciou obrazu získame súbor segmentov ktoré pokrývajú celý obraz a pomocou tohto súboru segmentov môžeme jednoduchšie identifikovať jednotlivé časti obrazu.



Obr. 1 Výsledok spracovania obrazu pomocou Image Processing Toolboxu.

V rámci spracovania obrazu potrebujeme pre automatické triedenie poznať základné vlastnosti jednotlivých snímaných objektov.

- *BoundingBox* - určuje polohu a veľkosť najmenšieho rámčeka obsahujúceho oblasť, ktorú vytvára objekt. Pomocou tejto funkcie získame informáciu o polohe ľavého horného rohu rámčeka v súradniciach (x, y) a o veľkosti rámčeka v pixeloch (šírka výška). V našom prípade na základe tejto informácie budeme určovať veľkosť a pozíciu objektu pre nastavenie polohy a koncového efektora robotického ramena.
- *Centroid* – určuje polohu ťažiska objektu v súradniciach, x,y pri 2D obraze.
- *Perimeter* – vypočíta vzdialenosti okolo hraníc objektu.
- *Regionprops* - vypočíta obvod na základe vzdialenosti medzi každým susedným párom pixelov okolo okraja oblasti.
- *FilledArea* - Výstupom tejto funkcie je skalárna hodnota, ktorá udáva počet pixelov, ktoré tvoria daný objekt. Tento údaj je pre nás veľmi dôležitý, keďže bol potrebný pre určenie tvaru objektu.

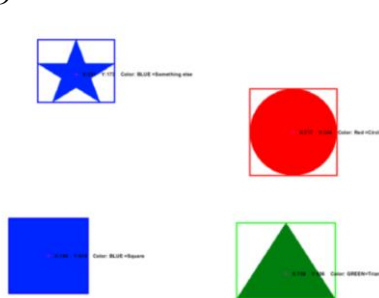
A. Návrh algoritmu pre rozpoznávanie objektov

V rámci riešenia problému bolo potrebné vytvoriť algoritmus na rozpoznávanie základných farieb (červená, zelená, modrá) a rozpoznávanie tvarov (štvorec, kruh, trojuholník). Základom bolo získanie obrazu z kamery. Výsledkom získania farebného obrazu z kamery je trojrozmerná matica. Táto matica obsahuje informácie o veľkosti a farbe objektu. Následne potrebujeme daný obraz spracovať, t.j. konvertovať RGB obraz do odtieňov šedej, pomocou 2D mediánového filtra odstrániť šum z obrazu a po odstránení šumu je potrebné previesť obraz na binárny (čierna-pozadie, biela-objekt), identifikovať vlastnosti objektov, uvedených vyššie, ktoré sú zapísané do tabuľky Na základe získaných vlastností bola na základe jednoduchého vzťahu zypočítaný parameter kruhovitosti objektu.

$$Kruhovitost' = 4\pi \cdot \frac{Obsah}{Obvod^2} \quad (1)$$

Táto definícia kruhovitosti zodpovedá definícii kruhovitosti, ktorá bola popísaná E. P. Coxom v roku 1927.[1] Podľa tohto parametra sme následne určili tvar objektu.

- Kruh - 0.8 až 1.1
- Štvorec - 1.273
- Trojuholník - 1.6539



Obr. 2 Rozpoznávanie tvaru objektu.

Pre otestovanie algoritmu, programu bol použitý obr.2 so štvorcem, kruhom, trojuholníkom a hviezdou. Algoritmus dokázal rozoznať tvary a farby, pričom hviezdou definoval ako neznámy objekt.

B. Algoritmus riadenia robotického ramena

Tabuľka údajov o jednotlivých objektoch bola využitá pre robotické rameno zobrazené na obr.3, na triedenie objektov podľa tvaru a farby. V rámci navrhnutého algoritmu boli na a výber dve druhy triedenia.

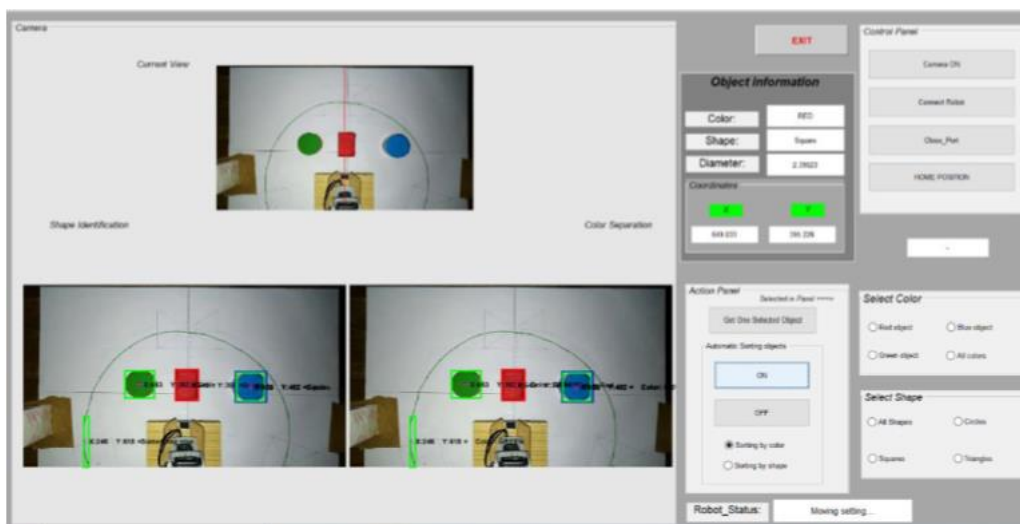
- Prvé triedenie bolo založené na princípe zvolenia farby a tvaru, kde robot odkladal iba objekty ktoré vyhovovali zvolenému tvaru a farbe.
- Druhé triedenie bolo založené na princípe triedenia všetkých objektov a ich odkladanie na príslušné miesto, pričom tu sme tu mali na výber odkladanie na základe farby alebo na základe tvaru.



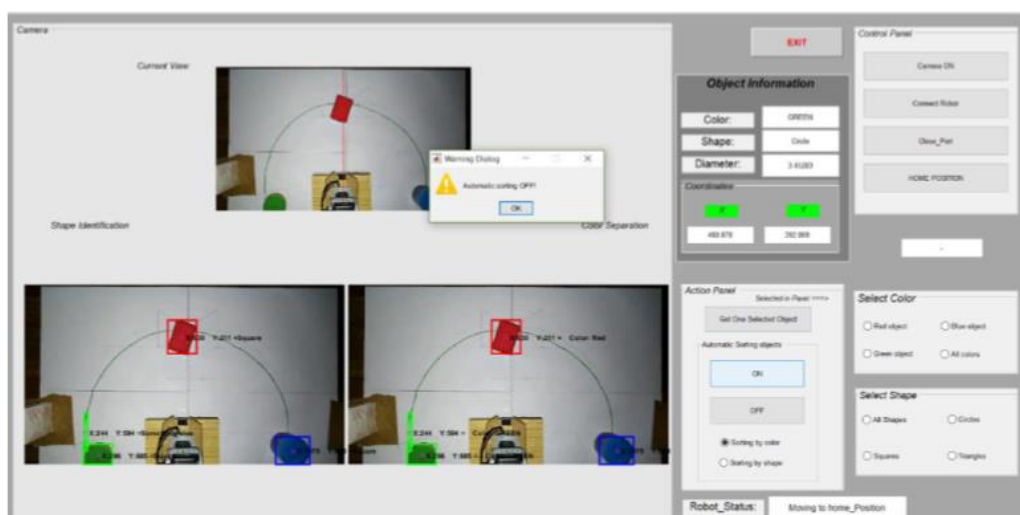
Obr. 3 Robotické rameno

Informácie o objektoch uložené v tabuľke slúžili pre určenie súradníc jednotlivých objektov a na základe zvoleného typu triedenia, boli premiestňované jednotlivé objekty na definované koncové polohy (obr.4 a obr.5).

Pre výpočet in inverznej kinematiky robota bola použitá funkcia $[FiBx]=inverse_kinematic(X, Y, Z, d)$, ktorej vstupnými hodnotami sú súradnice objektu X, Y, Z v centimetroch a vzdialenosť objektu od základne robota d . Výstupom funkcie je opäť riadkový vektor, ktorý obsahoval informácie o uhloch, o ktoré sa musia jednotlivé pohony natočiť. Vektor obsahuje informácie iba o prvých štyroch pohonoch, informácia pre piaty pohon, ktorý tvorí efektor ramena, sa počítala v samostatnej funkcii. Vstupným parametrom pre výpočet bola šírka objektu, ktorá sa následne prepočítala na centimetre a využitím trigonometrie sa vypočítal potrebný uhol pre pohon na uchopenie objektu. Po uchopení objektu, sa robot presunul aj s objektom na definovanú pozíciu a podľa použitej metódy triedenia, uložil objekt na správne miesto. Po uložení objektu na správne miesto sa robot vrátil späť do domovskej pozície. Po vrátení do domovskej pozície sa vytvorila nová fotka miesta s objektami. Následne znova prebehla klasifikácia objektov, vyberal sa jeden objekt počítala sa inverzná kinematika.



Obr. 4 Začiatok triedenia

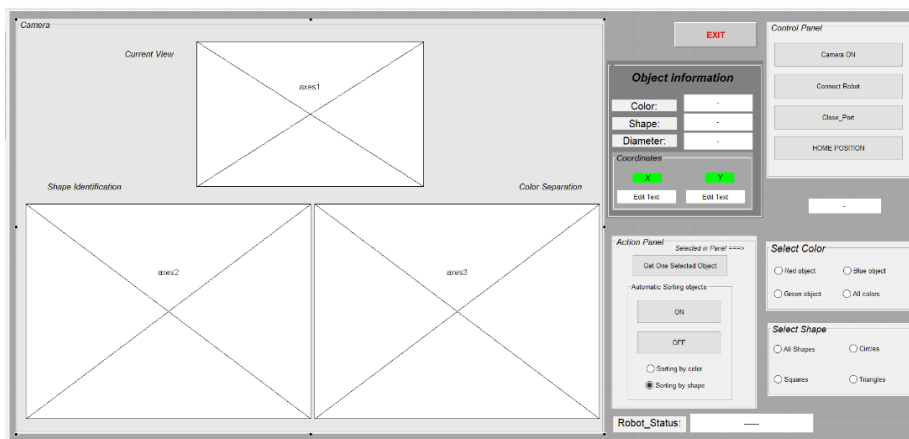


Obr. 5 Koniec triedenia

V rámci programu bolo v MATLABe vytvorené užívateľské rozhranie (obr.6), ktoré slúžilo pre nastavenie a ovládanie parametrov triedenia.

- **Control Panel** slúži na zapnutie kamery a prípravu robotického ramena.
- **Home Position** slúži pre vrátenie robota do domovskej pozície.
- **Action Panel** slúži na výber druhu triedenia, kde na výber sú dva druhy triedenia.
 - Get One Selected Object slúži na zozbieranie jedného konkrétneho objektu s definovaným tvarom a farbou. Definovaný tvar a farbu si užívateľ vyberie v paneloch napravo Select Color a Select Shape.
 - Automatic Sorting Objects, táto funkcia triedy všetky objekty podľa farby alebo tvaru.
- **Object information**, obsahuje základné informácie o objekte, ktorý sa chystá robot zdvihnúť a premiestniť. Informácie o farbe, tvare, šírke a súradniciach stredu v pixeloch.

Na ľavej strane obrazovky je aktuálny pohľad kamery a dve samostatné okná pre zobrazenie snímky, s ktorou sa priešenieovalo vrátane ohraničenie každého jedného objektu, ktorý algoritmus rozpoznal spolu so súradnicami a popisom.



Obr. 6 GUI pre nastavenie triedenia

III. ZÁVER

V tomto príspevku sme v krátkosti prezentovali problematiku počítačového videnia, ktorú sme využívali pre rozpoznávanie objektov a ich následné automatické triedenie pomocou robotického ramena. Ako základný nástroj pre návrh algoritmu rozpoznávania objektov, ako aj návrh algoritmov riadenia robotického ramena bol využívaný program MATLAB spolu s jeho toolboxami. Pre ovládanie robotického ramena, slúžiaceho na automatické triedenie, bolo taktiež v programe MATLAB vytvorené grafické užívateľské rozhranie. Aj keď testovaním navrhnutých algoritmov bola overená ich funkčnosť, ďalšie vylepšenia, hlavne čo sa týka rozpoznávania obrazu by mohlo priniesť využitie umelých neurónových sietí.

POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-19-0210.

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23.

LITERATÚRA

- [1] Cox, E. P. (1927). A Method of Assigning Numerical and Percentage Values to the Degree of Roundness of Sand Grains. *Journal of Paleontology*, 1(3), 179–183. <http://www.jstor.org/stable/1298056>
- [2] Zhao, Y., Guo, S., Wang, Y. a kol. Prototypová metóda neštruktúrovaného vnímania chirurgického stavu a navigácie pre robota endovaskulárnej chirurgie založená na CNN. *Med Biol Eng Comput* 57, 1875-1887 (2019). <https://doi.org/10.1007/s11517-019-02002-0> <https://doi.org/10.1007/s11517-019-02002-0>
- [3] Takashimizu, Y., Iiyoshi, M. New parameter of roundness R: circularity corrected by aspect ratio. *Prog. in Earth and Planet. Sci.* 3, 2 (2016). <https://doi.org/10.1186/s40645-015-0078-x>
- [4] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.
- [5] CAMACHO, Cezanne. 2018. Machine and deep learning: Convolutional Neural Networks URL: < https://cezanne.github.io/Convolutional_Neural_Networks/ >
- [6] Ružický E., Ferko A.: Počítačová grafika a spracovanie obrazu, SAPIENTIA, Bratislava, 1995
- [7] Šikudová E., Černeková Z., Benešová W., Haladová Z., Kučerová J.: Počítačové videnie, detekcia a rozpoznávanie objektov, WIKINA, Praha, 2014
- [8] Nilsson, Fredrik & Murhed, Anders: Select the best technology for your vision application, dostupné na internete URL <https://cdn.sick.com/media/docs/2/22/322/whitepaper_sick_ag_whitepaper_select_the_best_technology_for_your_vision_application_en_im0063322.pdf>
- [9] Zhu, L., Spachos, P., Pensini, E., & Plataniotis, K.N. (2021). Deep learning and machine vision for food processing: A survey. *Current Research in Food Science*, 4, 233 - 249.
- [10] Sinčák P., Andrejková G.: Neurónové siete Inžiniersky prístup(1. diel), URL: < <https://fmfi-uk.hq.sk/Informatika/Neuronove%20Siete/NS1/all.html> >
- [11] Rosenblatt, F. (1958) The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review*, 65, 386. <https://doi.org/10.1037/h0042519>
- [12] AgBot II Robotic Site-specific Crop and Weed Management Tool, URL:< <https://research.qut.edu.au/future-farming/projects/robot-platform-design-agbot-ii-a-new-generation-tool-for-robotic-site-specific-crop-and-weed-management/>>
- [13] Image Recognition with Deep Neural Networks and its Use Cases(dec 11, 2019), URL:< <https://www.altexsoft.com/blog/image-recognition-neural-networks-use-cases/>>
- [14] McCool, Chris, Perez, Tristan, & Upcroft, Ben (2017) Mixtures of Lightweight Deep Convolutional Neural Networks: Applied to Agricultural Robotics. *IEEE Robotics and Automation Letters*, 2(3), Article number: 78491671344-1351. URL:< <https://eprints.qut.edu.au/209663/>>
- [15] Khuyen Le(Mar 8, 2021) LeNet and MNIST handwritten digit recognition URL:<<https://medium.com/mllearning-ai/lenet-and-mnist-handwritten-digit-classification-354f5646c590>>

Model turbíny malej vodnej elektrárne s fuzzy modelom účinnosti

¹Pavol FEDOR, ²Peter Radváni

¹Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,
Technická univerzita v Košiciach, Slovenská republika

²EneroControl s.r.o., Pri plynárni 2, 040 01 Košice, Slovenská republika

pavol.fedor@student.tuke.sk, radvani@energocontrol.sk

Abstract — V článku je uvedený návrh modelu hydraulického turbíny ako najdôležitejšieho podsystemu malej vodnej elektrárne, ktorého výpočet účinnosti je analyticky náročný a závislý od parametrov, ktoré sa často získavajú len teoretickým odhadom. Na vytvorenie modelu účinnosti bol v článku použitý fuzzy systém na základe nameraných prevádzkových údajov, a to prietoku turbíny a jej mechanickej uhlovej rýchlosti bez potreby presnej znalosti jej parametrov.

Keywords — malé vodné elektrárne, hydraulická turbína, fuzzy systém, účinnosť turbíny

I. ÚVOD

Využívanie hydroenergetického potenciálu vodných tokov na výrobu elektrickej energie je vysoko pozitívnym globálnym, dlhodobým opatrením na znižovanie významných negatívnych vplyvov na životné prostredie a zdravie, pretože predstavuje environmentálne najpriateľnejší spôsob získavania elektrickej energie [1-3]. Výroba elektrickej energie vo vodných elektrárnach zaťažuje životné prostredie neporovnateľne menej, než výroba na báze tradičných fosílnych palív [4-6]. V súčasnosti existuje viacero spôsobov modelovania vodných elektrární pozostávajúcich z jednotlivých podsystemov [7-9]. Štruktúra modelov malých vodných elektrární (MVE) je v porovnaní s nimi jednoduchšia, pretože vzhľadom na ich konštrukčné usporiadanie (ide o systémy menších výkonov) môžeme niektoré vplyvy zanedbať (napríklad vplyv časovej konštanty vln vody, javu vodného kladiva, vplyv takzvaných nárazových komôr a pod. [10-12].

Predložený článok popisuje model hydraulického turbíny, ktorá patrí k najdôležitejším podsystemom MVE. Na rozdiel od doteraz známych modelov hydraulických turbín, kde je ich účinnosť vypočítavaná analyticky, je pre určenie jej účinnosti v článku navrhnutý fuzzy systém.

Návrh fuzzy systému vychádza len z nameraných prevádzkových údajov ako sú prietok turbíny a jej mechanickej uhlovej rýchlosti, pričom nie je potrebná presná znalosť jej parametrov, ktoré sa počas prevádzky menia a je problém ich presne určiť.

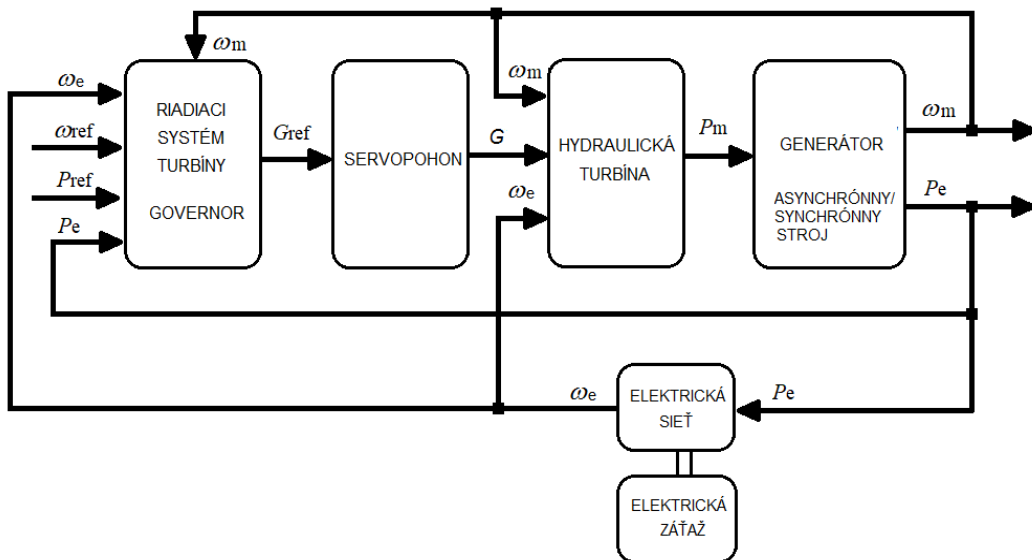
II. MODEL MALEJ VODNEJ ELEKTRÁRNE

Model malej vodnej elektrárne je pomerne zložitý nelineárny dynamický systém, ktorý sa vo všeobecnosti skladá zo štyroch základných podsystemov [9] (Obr.1):

- Governora – predstavujúceho systém riadenia turbíny;
- Servopohonu, ktorý slúži ako pohon na reguláciu prietoku vody cez hydraulickú turbínu pomocou ovládania ventilov alebo rozvádzačov lopatiek podľa výstupného signálu Governora;
- Hydraulického turbíny – predstavujúceho proces premeny energie;
- Elektrického generátora a elektrickej siete.

Hydraulická turbína je kvôli jej dôležitej funkcii v celom procese výroby elektrickej energie považovaná za srdce vodnej elektrárne. Turbína premieňa kinetickú alebo potenciálnu energiu vody na mechanickej rotáciu, ktorá pohybuje rotorom generátora cez spoločný hriadeľ. Na dosiahnutie realistických a relevantných výsledkov je preto potrebné použiť jej presný model.

Význam jednotlivých veličín na Obr.1 je uvedený v Tab.1.



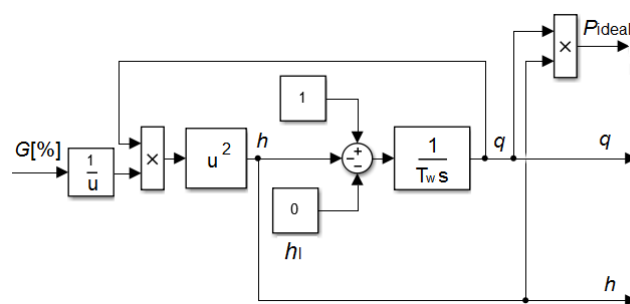
Obr. 1 Bloková schéma malej vodnej elektrárne MVE

Tabuľka 1. Popis veličín pre simulačný model MVE

Premenné	Jednotky
ω_{ref} – referenčná hodnota uhlovej rýchlosti hriadeľa je obvykle zhodná s elektrickou uhlovou rýchlosťou siete	[rads ⁻¹]
P_{ref} – referenčný výkon turbíny	[W]
P_e – elektrický výkon generátora	[W]
ω_e – elektrická uhlová rýchlosť siete	[rads ⁻¹]
G – poloha servopohonu	[%]
G_{ref} – referenčná hodnota polohy servopohonu	[%]
P_m – mechanický výkon na hriadeľi turbíny	[W]
ω_m – mechanická uhlová rýchlosť hriadeľa	[rads ⁻¹]

III. MODEL PODSYSTÉMU HYDRAULICKEJ TURBÍNY

Obrázok 2 zobrazuje blokovú schému podsyntému hydraulickéj turbíny, ktorý bol vytvorený na základe teoretického a matematického popisu turbíny uvedeného v literatúre [5].



Obr. 2 Bloková schéma modelu vodnej turbíny

Modelovanie prívodového potrubia, ako súčasť znázorneného modelu, bolo robené za predpokladu nestlačiteľnosti vody a za predpokladu, že prívodové potrubie je dostatočne pevné s dĺžkou L a plochou prierezu A . Úbytok výšky vodného stĺpca následkom vplyvu prívodového potrubia je zanedbaný. Veličina T_w je časová konštanta zmeny prietoku označovaná ako časová konštanta vody.

Rovnica 1 predstavuje základný vzorec použitý pre návrh simulačného modelu hydraulickéj turbíny. Význam jednotlivých parametrov je uvedený v Tab.2.

$$\frac{d\bar{q}}{dt} = (\bar{h}_0 - \bar{h} - \bar{h}_1)g \frac{A}{L} \quad (1)$$

Tabuľka 2. Zoznam parametrov použitých v rovnici (1)

Symbol	Popis	Jednotka
q	Prietok turbíny	[m ³ /s]
A	Plocha prierezu prívodného potrubia	[m ²]
L	Dĺžka prívodného potrubia	[m]
g	Gravitačné zrýchlenie	[m/s ²]
h_0	Výška vodného stĺpca	[m]
h	Výška vodného stĺpca na vstupe do turbíny	[m]
h_t	Strata výšky vodného stĺpca spôsobená trením v prívodnom potrubí	[m]

A. Model hydraulickej turbíny zohľadňujúci straty v potrubí

V reálnej prevádzke hydraulická turbína nepracuje so 100% účinnosťou a z toho vychádza požiadavka na zohľadnenie tejto skutočnosti aj pri jej modelovaní. Spôsobov, ako zakomponovať vplyv účinnosti turbíny do simulačného modelu je viacero [5,7].

Jedným z najjednoduchších spôsobov je použitie takzvaného prietoku pri nulovom zaťažení. Tento prietok označovaný ako q_{nl} je konštantou, charakterizujúcou hodnotu stálych výkonových strát turbíny, pričom vo vzorci je táto konštanta odčítaná od aktuálneho prietoku turbínou, a až táto výsledná hodnota prietoku je prenasobená výškou vodného stĺpca h podľa nasledujúcej rovnice

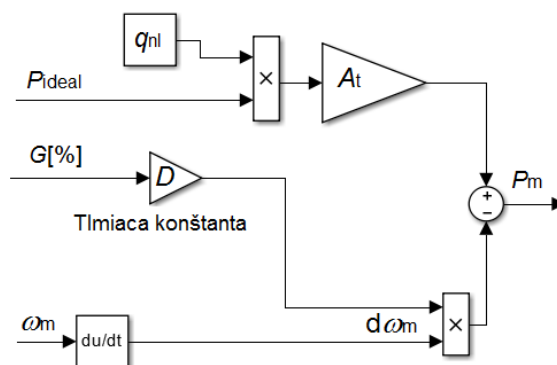
$$P_m = A_t h (q - q_{nl}) - DG \Delta \omega \tag{2}$$

Rovnako sa počas prevádzky elektrárne uplatňuje aj tlmiaci efekt turbíny, ktorý je funkciou rozdielu rýchlostí, polohy servopohonu a polohy ventilu prívodu vody. Veličina $\Delta \omega$ reprezentuje zmenu mechanickej rýchlosti hriadeľa turbíny, D je veľkosť priemeru prívodného potrubia (tiež ako tlmiaci koeficient turbíny) a G je poloha servopohonu v percentách.

Parameter A_t je považovaný za zosilnenie hydraulickej turbíny. V rôznych vedeckých publikáciách sa jeho vyčíslenie líši, napr. podľa literatúry [7] ho môžeme vypočítať nasledovne

$$A_t = \frac{\text{Výkon Turbíny [MW]}}{(\text{Výkon Generátora [MVA]} h_r (q_r - q_{nl}))} \tag{3}$$

pričom veličiny s dolným indexom r môžeme považovať za nominálne. To znamená, že q_r je prietok vody turbínou pri jej nominálnom zaťažení a h_r je výška vodného stĺpca potrebná na dosiahnutie spomínaného nominálneho prietoku q_r . Zvyčajne sú tieto hodnoty súčasťou technickej dokumentácie jednotlivých systémov vodných elektrární. Model výpočtu strát hydraulickej turbíny je zobrazený na Obr.3.



Obr. 3 Model výpočtu strát hydraulickej turbíny

B. Model zohľadňujúci účinnosť turbíny prostredníctvom výpočtu jej účinnosti

Tento variant tvorby simulačnej štruktúry hydraulickej turbíny vychádza zo základného vzťahu popisujúceho výpočet jej výstupného mechanickeho výkonu P_m

$$P_m = \eta \rho q g h \tag{4}$$

kde η predstavuje hodnotu účinnosti hydraulickej turbíny; ρ je hustota vody (zvyčajne sa používa hodnota 1000 kg/m³); q je aktuálna hodnota toku vody turbínou; g je gravitačné zrýchlenie (9,81 m/s²); a h je aktuálna výška vodného stĺpca vodnej elektrárne. V modeli sú

použité normované veličiny. V tomto prípade je na vyjadrenie účinnosti hydraulickéj turbíny použitý výpočet, ktorý je uvedený v literatúre [8-9], pre ktorý je potrebné poznať parametre turbíny ako polomer lopatiek turbíny - R_{blades} [m] a veľkosť plochy, ktorú dané lopatky „obtiajú“ - A_{blades} [m²]. Taktiež sú pre vyčíslenie nevyhnutné aj aktuálne hodnoty toku vody turbínou q , ako aj jej mechanická uhlová rýchlosť ω_m

$$\eta(\lambda, q) = \left[\frac{1}{2} \left(\frac{90}{\lambda_i} + q + 0.78 \right) \exp \left(\frac{-50}{\lambda_i} \right) \right] (3,33q) \quad (5)$$

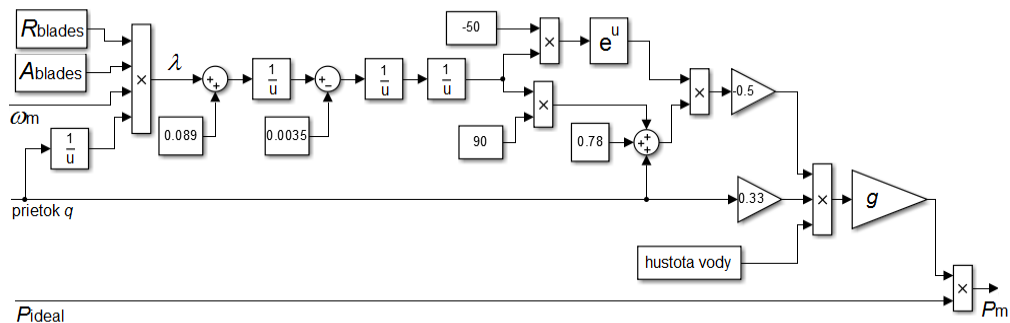
kde pre λ_i platí

$$\lambda_i = \left[\frac{1}{(\lambda + 0,089)} - 0,0035 \right]^{-1} \quad (6)$$

a λ je definovaná ako

$$\lambda = \frac{R_{blades} A_{blades} \omega_m}{q} \quad (7)$$

Vstupom do výpočtu účinnosti hydraulickéj turbíny by bol v tomto prípade aktuálny prietok vody turbínou q a mechanická uhlová rýchlosť turbíny ω_m . Matematickému popisu zodpovedá blokovo schéma na Obr.4.

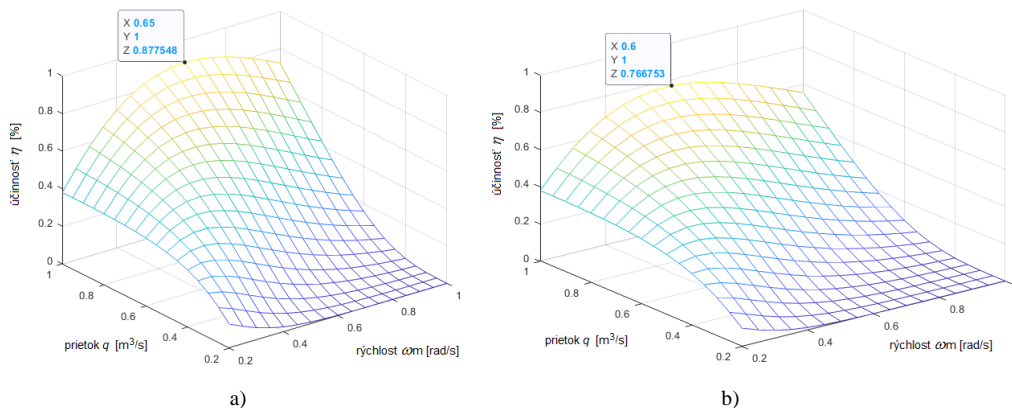


Obr. 4 Model výpočtu účinnosti hydraulickéj turbíny

IV. FUZZY MODEL ÚČINNOSTI HYDRAULICKEJ TURBÍNY

Všetky vyššie uvedené spôsoby výpočtu účinnosti turbíny sú založené na pomerne komplikovaných a nie vždy presných výpočtoch, a to najmä z dôvodu potreby detailnej znalosti jej parametrov. Napr. podľa rovnice (7) závisí účinnosť od parametrov lopatiek konkrétnej turbíny, ktoré sa počas prevádzky môžu výrazne meniť. Táto závislosť je ďalej silne nelineárna (rovnice 5-7) a závislá od viacerých približne určených parametrov. Napr. pre polomer lopatiek $R_{blades} = 1.9$ m a plochu ich obtierania $A_{blades} = 9$ m² je ukázaná účinnosť pre rôzne prietoky a mechanické uhlové rýchlosti turbíny na Obr.5a.

Pri zmene plochy obtiekania lopatiek (napr. opotrebovaním) na $A_{blades} = 8$ m² sa účinnosť turbíny zmení podľa Obr.5b. Maximálna účinnosť turbíny v tomto prípade je približne 76,6%, zatiaľ čo v predchádzajúcom prípade bola 87,7%.



Obr. 5 Účinnosť turbíny a) $R_{blades} = 1.9$ m a $A_{blades} = 9$ m², $\eta = 87,7\%$
 b) $R_{blades} = 1.9$ m a $A_{blades} = 8$ m², $\eta = 76,6\%$

Z obrázka 5 je zrejmy rozdiel v maximálnej možnej hodnote účinnosti turbíny, spôsobený len opotrebovaním lopatiek turbíny, čo má za následok zmenu vnútornej plochy lopatiek (parameter A_{blades}). Tento parameter má podľa rovnice (7) významný vplyv na výpočet účinnosti hydraulickéj turbíny, a preto je problematické zohľadňovať jeho vplyv pri tvorbe jej simulačného modelu.

Ďalšie kroky nášho výskumu v oblasti modelovania podsystemu hydraulickéj turbíny sú v zvažovaní ďalších možností z hľadiska univerzálneho, no stále presného a rýchleho výpočtu jej účinnosti.

Obraz účinnosti turbíny je možné vytvoriť aj na základe merania prietoku vody turbínou q , mechanickej uhlovej rýchlosti hriadeľa ω_m a účinnosti turbíny η v ustálených prevádzkových stavoch, tak ako je to uvedené v Tab.3.

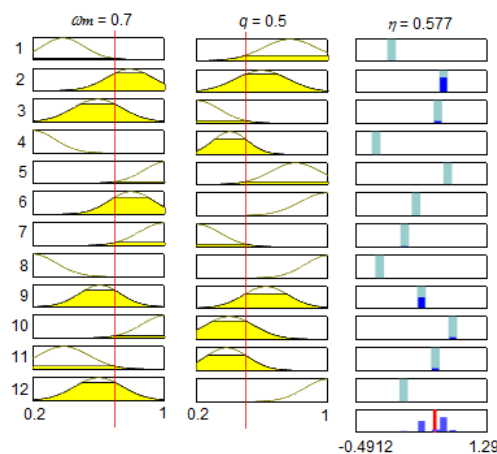
Údaje z Tab. 3 možno použiť na vytvorenie fuzzy inferenčnej štruktúry (FIS) modelovanej účinnosti popisujúcej nameraný vzťah medzi $[\omega_m, q] \rightarrow \eta$.

Na základe týchto údajov bol navrhnutý fuzzy model účinnosti pomocou štandardných nástrojov zhukovej analýzy, ktorý vedie k zlepšeniu kvality simulačného modelovania a k skráteniu času potrebného na tvorbu fuzzy modelu [13]. Základnou charakteristikou zhukovej analýzy je zníženie celkového počtu fuzzy pravidiel a ich kvalitné počiatkové nastavenie. Z veľkého množstva dostupných metód v oblasti vývoja adaptívnych fuzzy systémov bol vybraný adaptívny neuro-fuzzy interferenčný systém ANFISEDIT (súčasť softvéru programu MATLAB) [14]. Použila sa metóda „subtractive clustering“ s nasledujúcimi parametrami: *Range of influence* = 0.5, *Squash factor* = 1.1, *Accept ratio* = 0.45, and *Reject ratio* = 0.005.

Tabuľka 3 Namerané účinnosti turbíny v ustálených prevádzkových stavoch

Účinnosť η		ω_m				
		0,2	0,4	0,6	0,8	1
q	0,2	12,77%	30,36%	36,23%	37,83%	37,97%
	0,4	2,74%	26,40%	49,60%	64,41%	73,24%
	0,6	0,35%	13,64%	40,90%	66,92%	87,22%
	0,8	0,04%	5,68%	27,20%	56,26%	84,31%
	1	0,00%	2,11%	16,07%	42,03%	72,49%

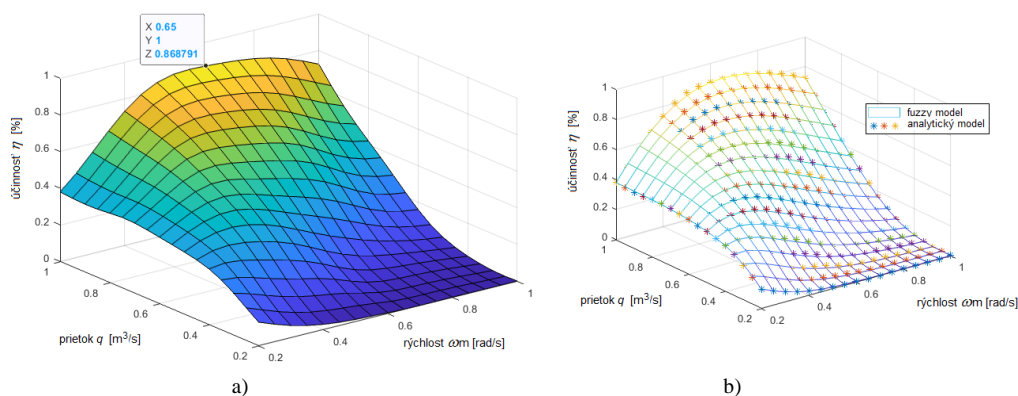
Ako výsledok sme získali statický fuzzy systém typu Sugeno s dvanástimi pravidlami, ako je to znázornené na Obr. 6.



Obr. 6 Štruktúra fuzzy modelu účinnosti hydraulickéj turbíny – Sugeno typ s 12 pravidlami

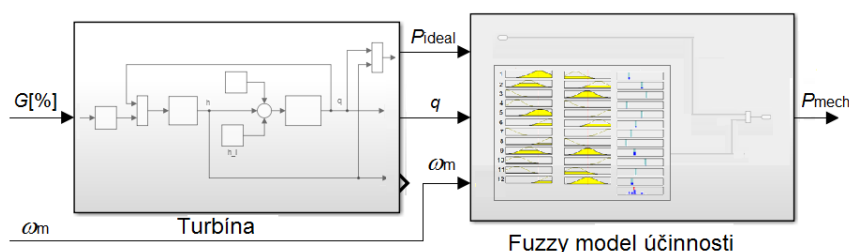
Navrhovaný fuzzy model nahrádza numerický výpočet účinnosti turbíny na základe rovníc (5)-(7). Tieto rovnice, ako už bolo spomenuté, obsahujú niektoré parametre, ktoré nevieme celkom presne určiť.

Z obrázka 7a) vidíme, že účinnosť na základe fuzzy modelu pre $\omega_m=0.65$ a $q=1$ je $\eta=87,62\%$. Na základe porovnania účinnosti získanej z fuzzy modelu a účinnosti vypočítanej analyticky (pozri Obr.5), môžeme konštatovať, že navrhnutý fuzzy model dosahuje takmer rovnakú presnosť ako analytický model. Ukazuje to aj grafické porovnanie účinnosti získaných analyticky a na základe fuzzy modelu, ktoré je zobrazené na Obr. 7b).



Obr. 7 a) Fuzzy model účinnosti hydraulického turbíny
 b) Porovnanie analytického a fuzzy modelu účinnosti turbíny

Bloková schéma modelu hydraulického turbíny s fuzzy modelom jej účinnosti je znázornená na Obr. 8.



Obr. 8 Bloková schéma modelu hydraulického turbíny so zohľadnením jej účinnosti

V. ZÁVER

Predložený článok popisuje aj stručný prehľad súčasných poznatkov a spôsobov v oblasti modelovania vodných turbín ako súčasti malých vodných elektrární. Je tu navrhnutý fuzzy model vodnej turbíny so zohľadnením jej účinnosti. Navrhnutý fuzzy model je možné získať z nameraných údajov prietoku vody turbínou a jej mechanickej uhlovej rýchlosti. Ukazuje sa, že navrhnutý fuzzy model dosiahol prakticky rovnakú presnosť ako analytické modely. Takýto fuzzy model je možné použiť v procese návrhu a vývoja optimálneho riadenia energie hydraulických turbín, ako aj v efektívnom procese určovania skutočnej miery opotrebovania turbíny.

POĎAKOVANIE

Článok bol publikovaný s podporou projektu APVV-16-0206 a APVV-19-0210.

LITERATÚRA

- [1] H.S. Sachdev, A.K. Akella and N. Kumar, „Analysis and evaluation of small hydropower plants: A bibliographical survey,“ *Renewable and Sustainable Energy Reviews*, vol. 51, 2015, pp. 1013-1022. ISSN 1364-0321. doi.10.1016/j.rser.2015.06.065.
- [2] O. Doso and S. Gao, „An overview of small hydro-power development in India,“ *AIMS Energy*, vol.8, no. 5, 2020, pp.896-917. doi 10.3934/energy.2020.5.896
- [3] S.P. Adhau, R.M. Moharil, P.G. Adhau, „Estimation of micro hydro-power plant capacity from potential sites,“ in: *Proceedings of IEEE power electronics, drives and energy systems conference – PEDES*, Bengaluru, India; September 2012, pp. 1–5. doi.10.1109/PEDES.2012.6484339
- [4] S. Overshoote, *Renewable energy: Power for sustainable future*, Oxford University Press, 2018 , p. 656. ISBN 9780198759751.
- [5] A.K.Akella, R.P.Saini, „Social, economical and environmental impacts of renewable energy systems,“ in: *Int J Renew Energy*, vol. 34, no.2., 2009, pp. 390-396. doi.10.1016/j.renene.2008.05.002
- [6] K. Mahesh, *Social, Economic, and Environmental Impacts of Renewable Energy Resources*, Chapter in: K. Eloghene Okedu, A. Tahour, and A. Ghani Aissaou, Eds., “Wind Solar Hybrid Renewable Energy System,” Feb. 2020, doi: 10.5772/intechopen.77440.
- [7] A. Acakpovi, E.B. Hagan, and F.X. Fifatin, “Review of Hydropower Plant Models,“ *International Journal of Computer Applications*, vol.108, no.18, 2014, pp. 33-38. doi.10.5120/19014-0541

- [8] N. Kishor and J. Fraile-Ardanuy, *Modeling and Dynamic Behaviour of Hydropower Plants*, London: The Institution of Engineering and Technology, 2017. p. 279. ISBN: 978-1-78561-195-7.
- [9] G. Munoz-Hernandez, S. Mansoor, D. Jones, *Modelling and Controlling Hydropower Plants*. London: Springer, 2013. ISBN 978-1-4471-6221-6.
- [10] R. Olexa, D. Perduková and P. Fedor, „Fuzzy Approach of Modeling a Hydraulic Turbine Efficiency,“ *MM Science Journal*, vol. 2020, 2020, pp. 4086-4092. ISSN 1803-1269. doi.10.17973/MMSJ.2020_11_2020033.
- [11] D. Tsuanyo, B. Amougou, A. Aziz, B.N. Nnomo, D. Fioriti and J. Kenfach, „Design models for small run-of-river hydropower plants: a review,“ *Sustainable Energy res*, vol.10, no.3, 2023. doi.10.1186/s40807-023-00072-1.
- [12] M. Sattouf, “Simulation Model of Hydro Power Plant Using Matlab/Simulink,” *International Journal of Engineering Research and Applications*, vol. 4, no.1, 2014, pp 295-301. ISSN 2248-9622.
- [13] P. Fedor, D. Perdukova, “Model Based Fuzzy Control Applied to a Real Nonlinear Mechanical System”, *Iranian Journal of Science and Technology, Transactions of Mechanical Engineering*, vol. 40, 2016, pp. 113-124. DOI 10.1007/S40997-016-0005-9.
- [14] D. Perduková, P. Fedor, M. Lacko, „DC Motor Fuzzy Model Based Optimal Controller”, *MM Science Journal*, vol. 2021, 2021, pp. 4879-4885. DOI: 10.17973/MMSJ.2021_10_2021033.

Fyzikálny model kontinuálnej linky

¹Pavol SMOLEŇ, ²František ĎUROVSKÝ

¹ Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,

² Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,
Technická univerzita v Košiciach

¹pavol.smolen@student.tuke.sk, ²frantisek.durovsky@tuke.sk

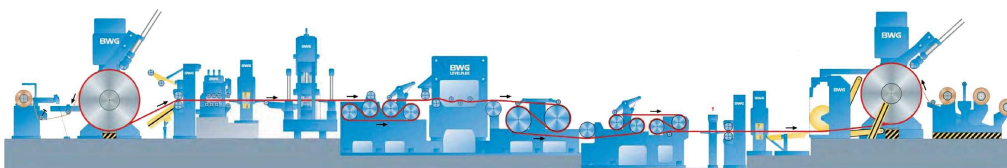
Abstrakt - Článok opisuje fyzikálny model kontinuálnej linky určený na výučbu riadenia zložitých mechatronických systémov. Mechanickú časť modelu tvorí odvíjačka, navíjačka, pracovné a ťažné valce a horizontálny zásobník materiálu. Riadenie modelu zabezpečuje technologický regulátor na báze PLC s technologickou kartou, napájanie motorov je realizované frekvenčnými meničmi. Model umožňuje nastaviť rôzne konfigurácie kontinuálnej linky. Článok je doplnený o jednoduchý príklad usporiadania linky s príslušnou riadiacou štruktúrou a nameranými priebehmi.

Kľúčové slová - riadenie kontinuálnych liniek, frekvenčný menič, programovateľný logický automat, technologický regulátor

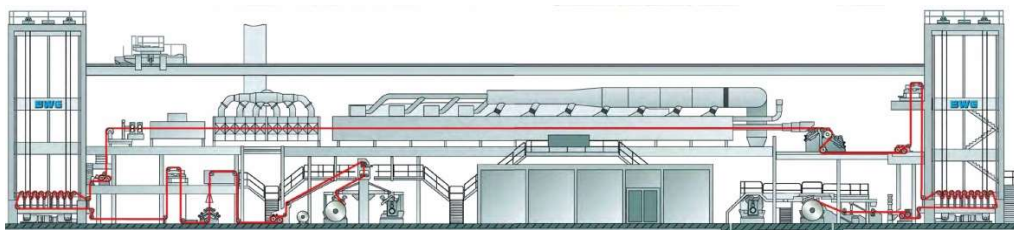
I. ÚVOD

Linky na spracovanie materiálov (Material Processing Lines - MPL) sa používajú v rôznych odvetviach priemyslu: v hutníckom, papierenskom, drevárskom, chemickom priemysle, v tlačiarňach a pod. Linky bývajú rozdelené na vstupný, technologický a výstupný úsek. Materiál je odoberaný na vstupe MPL zo zvitku na odvíjačke, postupuje cez technologický úsek a po úprave je navíjaný na výstupnej strane na navíjačku. V prípade, že sa jedná o linku s priečnym delením materiálu, kde sa pás reže alebo strihá na tabule, je na výstupe baliaca sekcia, kde sa narezaný materiál stohuje a následne balí.

V prípade jednoduchších technologických operácií sa pás zavádza do linky pri každom zvitku (Obr. 1). Pri zložitejšej technológii, kde zavádzanie materiálu do technologickej časti je komplikované, sa pás z nového zvitku na vstupe privarí na predchádzajúci pás, takže prechod pásu cez technologickú časť je plynulý a nevyžaduje opätovné zavedenie pásu (Obr. 2). V tomto prípade sa zvykne používať termín "nekonečný pás".



Obr. 1 Príklad jednoduchej výrobnéj linky [1]



Obr. 2 Príklad kontinuálnej výrobnéj linky [1]

II. KONFIGURÁCIA KONTINUÁLNYCH LINIEK

A. Mechanická časť kontinuálnej linky

Konfigurácia vstupného a výstupného úseku závisí od procesov v technologickom úseku (TÚ). Ak je zavádzanie pásu do linky jednoduché, robí sa pri každom zvitku. Vtedy stačí na

vstupe jedna odvíjačka, ťažné valce a priečne nožnice, ktoré pás vstupujúci do linky zarovnajú. V TÚ sa urobí želaná úprava materiálu (napr. valcovanie, ťahové rovanie, úprava hrúbky, pozdĺžne strihanie a pod.) a na výstupe sa upravený materiál navinie na navíjačku. Keďže pracovný cyklus nie je spojitý, aj na výstupe z linky stačí jedna navíjačka.

V prípade, ak je TÚ zložitejší, využíva sa tzv. nekonečný pás. Zavedenie pásu sa robí pri nižších rýchlostiach a s prestávkami. Materiál, ktorý sa používa na zavedenie, tak nemá požadovanú kvalitu a po prechode linkou sa musí vyradiť. Preto sa zavádzanie robí len pri spúšťaní linky a po prípadnej poruche spojenej s roztrhnutím pásu.

V hutníckom priemysle sa linky s nekonečným pásom používajú napr. pri morení oceľových pásov. Pás z teplej valcovne má na povrchu oxidy železa, ktoré je nutné pred ďalším valcovaním odstrániť. Oceľový pás preto prechádza cez vane s kyselinou chlorovodíkovou. Pobyt pásu vo vani s kyselinou je časovo obmedzený. Ak by bol pás vo vani kratší čas, neočistil by sa dostatočne. Ak by bol naopak vo vani dlhšie, príp. by sa úplne zastavil, kyselina by poškodila povrch pásu. Preto musí byť prechod pásu cez vane rovnomerný s približne konštantnou rýchlosťou.

Ďalším príkladom je žihacia linka. Žihanie sa robí po valcovaní za studena, kedy treba po valcovaní na tandemovej valcovacej trati vrátiť oceľovému pásu plasticitu. Pás prechádza komorami, kde sa zohreje na teplotu cca 700-800°C, a potom sa riadene chladí. Kombináciou ohreву a chladenia sa nastavujú mechanické parametre pásu, ako sú pevnosť a plasticita. Zavádzanie pásu cez vyhriate žihacie pece s ochrannou atmosférou je komplikované a nedá sa robiť pri pracovnej rýchlosti. Podobná situácia je pri linkách, kde pás prechádza cez vaňu s roztaveným materiálom (napr. zinkom), ktorý sa nanáša na pás, pri elektrolytickom pokovovaní (nanášanie cínu), pri lakovaní a pod.

V prípade liniek s nekonečným pásom je nutné zabezpečiť, aby materiál cez TÚ prechádzal rovnakou rýchlosťou, a to aj počas zavádzania pásu na vstupe a odbavovania navinutého materiálu na výstupe linky. Vstupný úsek je preto doplnený o druhú odvíjačku, zväračku a zásobník materiálu. Keď sa odvinie pás z prvej odvíjačky, je vstupný úsek zastavený. Pás z druhej odvíjačky sa privarí na predchádzajúci pás a vstupný úsek sa znovu rozbehne. Počas manipulácie so zvitkami sa pás do TÚ odoberá zo vstupného zásobníka. Po privarení pásu sa vstupná časť rozbehne na vyššiu rýchlosť, ako má TÚ, a opäť naplní vstupný zásobník, aby bol pripravený na ďalšiu výmenu.

Podobne je riešený výstupný úsek, kde po navínutí požadovanej dĺžky materiálu na navíjačku sa úsek zastaví, pás sa rozstrihne, zvitok sa z navíjačky stiahne a presunie na balenie. Do navíjačky sa zavedie koniec pásu a výstupný úsek sa rozbehne. Aj výstupný úsek sa rozbieha na vyššiu rýchlosť ako TÚ, aby sa vyprázdnil výstupný zásobník, kam sa ukladal materiál počas manipulácie so zvitkom.

B. Riadenie kontinuálnych liniek

Pracovné stroje kontinuálnych liniek sú poháňané zväčša elektrickými motormi. Niektoré pracovné mechanizmy využívajú pneumatické alebo hydraulické pohony, ale tým sa nebudeme venovať. Keďže stroje sú navzájom viazané cez spracovávaný materiál, je nutné riadiaci systém doplniť o technologickú reguláciu. Riadenie MPL býva preto rozdelené do viacerých úrovní [2]:

- riadenie pohonov - DC (drive control),
- technologická regulácia - PC (process control),
- nadradené riadenie a logické riadenie linky - PLC.

Najnižšiu úroveň predstavuje riadenie elektrických motorov (DC), ktoré poháňajú jednotlivé pracovné stroje. Motory sú napájané z výkonových meničov so základnými regulačnými slučkami otáčok a momentu. K tejto úrovni riadenia sú zaraďované aj snímače. Ak snímače nie sú priamo súčasťou riadenia pohonov, sú ich signály vyhodnocované v PLC.

Vyššou úrovňou riadenia je technologická regulácia (TC). Tá riadi spoluprácu jednotlivých pohonov a definuje ich vzájomné rýchlosti a ťahy. Snímače, ktoré priamo využívajú jednotlivé technologické uzly, sú pripojené k tejto úrovni riadenia. Komunikácia medzi TC a DC prebieha po niektorej z priemyselných zberníc (Profibus, Profinet, Powerlink, Ethercat a pod.).

Logické riadenie linky (obsluha signálov z ovládacích pultov, signály na blokovanie, spúšťanie a ovládanie linky, vizualizácia) a spracovanie signálov zo snímačov, ktoré sa nepodieľajú priamo na riadení technologických uzlov, je realizované prostredníctvom PLC. PLC tvorí zároveň rozhranie medzi riadenou linkou a informačným systémom podniku, z ktorého sa získavajú údaje o spracovávanom materiáli (rozmery, materiálové charakteristiky, požiadavky na ďalšie spracovanie na danej linke, napr. rozmery na delenie materiálu a pod.).

Technologická regulácia je špecifická tým, že vyžaduje, aby sa regulačné zásahy vykonávali

s definovanou periódou - v tzv. task-och. TC má viacero task-ov s rôzne veľkými periódami. Jednotlivé regulátory sa zaraďujú do príslušných task-ov podľa toho, aký rýchly zásah daný regulačný obvod vyžaduje. Tento zásah sa potom cyklicky opakuje s definovanou periódou. TC sa zvykne zaraďovať aj do úrovne nadradeného riadenia (PLC), kde sa realizuje buď na tzv. technologických PLC alebo na PLC vybavených špeciálnou technologickou kartou.

Výrobcovia ponúkajú balíky programových modulov, ktoré uľahčujú návrh riadenia pre MPL (napr. Siemens [3], BR [4]). Výhodou balíkov je, že sa jedná o modulárnu stavebnicu, z ktorej sa dá riadenie MPL poskladať. Jednotlivé moduly sú odladené, takže zabezpečujú základnú funkcionálnosť. Na druhej strane, často sa jedná o „čierne skrinky“ s definovaným rozhraním, ktorých vnútorná štruktúra je používateľovi skrytá. V prípade, ak je nutné robiť úpravu alebo zásah do takýchto blokov, býva to problém.

C. *Výučba problematiky riadenia kontinuálnych liniek*

Svoje špecifiká má aj výučba problematiky riadenia kontinuálnych liniek. Ešte cca pred desiatimi rokmi výrobcovia pri technologických kartách uvádzali aj štruktúru riadenia. Pri kreslení regulačných schém sa využívali blokové schémy, čo je v oblasti regulácie tradičný prístup. Schémy boli prehľadné a užívateľ na ich základe vedel odhadnúť, ako sa bude regulačná štruktúra správať. Tieto štruktúry boli navrhované univerzálne, takže pokrývali väčšinu aplikácií v praxi. V prípade, že užívateľ potreboval štruktúru modifikovať, a systém to umožňoval, skúsený užívateľ vedel, kde má zasiahnuť a aký výsledok zmena prinesie.

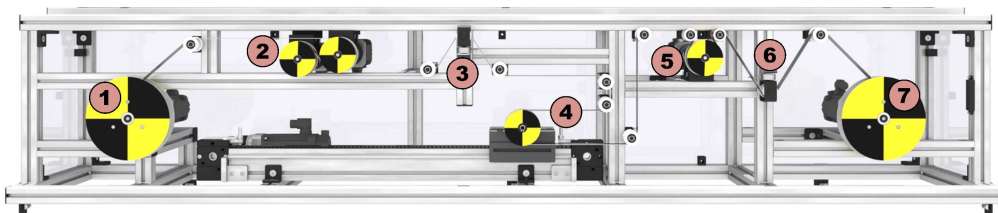
Kontinuálne výrobné linky a ich komponenty možno považovať za deterministické systémy s premenlivými parametrami. Na ich riadenie sa najčastejšie používajú klasické regulátory. V prípade, že zmeny parametrov sú významnejšie, používajú sa regulátory vybavené určitou mierou adaptivity. Uvedený prístup zabezpečuje spoľahlivú prevádzku aj pri zmene niektorých parametrov linky. Používanie blokových schém pri riadení deterministických sústav je tradičný prístup, ktorý je prehľadný a zároveň nezakrýva fyzikálnu podstatu riadenej sústavy. Práve tento prístup je z hľadiska výučby výhodný, lebo umožňuje študentom lepšie pochopiť vzájomné vzťahy tak v rámci riadenej sústavy, ako aj ich prepojenie s regulačnou štruktúrou. Bohužiaľ, používanie hotových modulov - „čiernych skriniek“, tieto súvislosti zahaľuje. Je síce výhodné pri opakovaných aplikáciách, kde šetrí čas, ale vyžaduje skúseného a vyškoleného užívateľa, ktorý už pozná fyzikálnu podstatu riadenej sústavy, čo nie je prípad študentov.

Druhým problémom pri výučbe riadenia zložitých mechatronických systémov je absencia praktických experimentov. Výpočty pre nastavenie komponentov linky a ich simulačné overenie je síce možné na urobiť počítači, ale tento prístup vyžaduje veľkú mieru technickej fantázie u študentov, aby si vedeli predstaviť, ako sa linka správa v skutočnosti. V kontinuálnej linke je viacero pohonov viazaných prostredníctvom materiálu, a tieto pohony sa navzájom ovplyvňujú. Na ich riadenie nestačí nastaviť sólo pohon, ale je nutný tréning, kde sa študent učí analyzovať priebehy z linky (zvyčajne rýchlosti a momenty susediacich pohonov), a na ich základe rozhodovať o potrebných úpravách riadiacej štruktúry. Zároveň je dobré, ak vidí, ako sa linka správa reálne. Študent tak získava fyzikálny pohľad na riadenie linky a určité zručnosti pri práci s nástrojmi na riadenie a monitorovanie linky. To bol hlavný dôvod, prečo bol navrhnutý a postavený fyzikálny model kontinuálnej linky s variabilnou štruktúrou.

III. FYZIKÁLNY MODEL KONTINUÁLNEJ LINKY

A. *Mechanická časť linky*

Usporiadanie fyzikálneho modelu kontinuálnej linky je na Obr. 3. Model je uložený v ráme z hliníkových profilov. Obsahuje šesť pohonov, ktoré reprezentujú základné komponenty kontinuálnych výrobných liniek: odvíjačku (1), ťažné S-valce (2), snímač sily (3), zásobník (4), ťažný valec (5), snímač sily (6) a navíjačku (7).

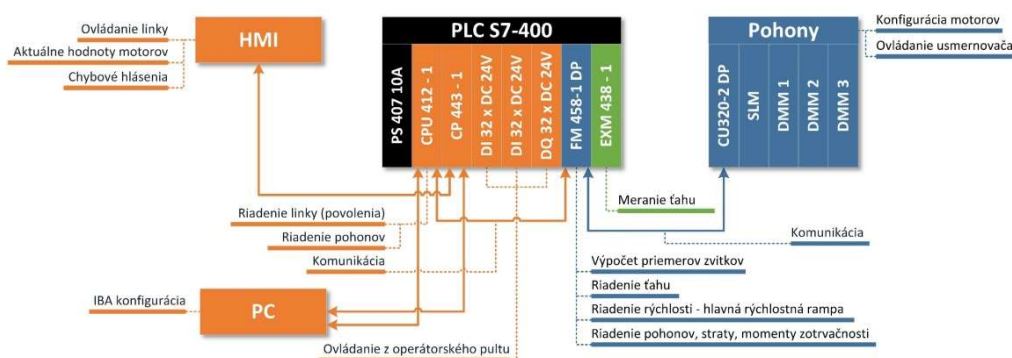


Obr. 3 Usporiadanie fyzikálneho modelu linky: 1 - Odvíjačka; 2 - S-valce 1; 3 - Snímač sily 1; 4 - Zásobník; 5 - Ťažný valec 2, 6 - Snímač sily 2; 7 - Navíjačka

a navíjačku (7). Do zostavy sú zaradené dva snímače sily (3) a (6), pomocou ktorých sa dá realizovať priama regulácia ťahu. Okrem toho je do linky zaradený ďalší pohon valca s väčším opásaním (5), pomocou ktorého sa dajú realizovať ťahové štruktúry. Štruktúru linky je možné podľa potreby meniť (viď časť D). Konštrukcia je zámerne robustná, aby vydržala aj neúmyselné, ale často nešetrné zaobchádzanie, ku ktorému môže dôjsť pri praktických cvičeniach. Motory, výkonové meniče aj riadiaci systém je postavený na komponentoch firmy Siemens, čo zaručuje ich vzájomnú kompatibilitu.

B. Riadenie a pohony

Riadiaci systém je postavený na báze programovateľného logického automatu SIMATIC S7-400 (CPU 412-1). Je vybavený komunikačným procesorom (CP443-1), technologickou kartou FM 458-1 DP s rozširujúcim modulom EXM 438-1 a digitálnymi vstupno-výstupnými modulmi. Hardvérová konfigurácia riadenia modelu je na Obr. 4.



Obr. 4 Hardvérová konfigurácia riadenia fyzikálneho modelu

Vzhľadom na to, že súčasná zostava riadenia linky síce spĺňa potreby linky, no v praxi je už nahrádzaná novšími komponentmi, aj tu sa plánuje náhrada jestvujúceho PLC za novšie, technologické, radu S7-1500.

Pohon jednotlivých valcov zabezpečujú synchronne servomotory. Kvôli unifikácii bol zvolený jeden typ, ktorý spolu s prevodovkou pokryje nároky na požadovanú rýchlosť a ťahy v linke. Model linky obsahuje celkovo 6 servomotorov Siemens 1FK7032-2AK71-1QG0 s integrovaným enkodérom. Použitý je absolútny jednootáčkový enkodér s rozlíšením 20 bitov. Enkodér prenáša signály do riadenia meniča prostredníctvom rozhrania Drive-Cliq. Každý motor má vlastnú priamu prevodovku Apex Dynamics PG II 060-025. Pri pohone zásobníka je použitá uhlová prevodovka PG IIR 060-025. Parametre motorov sú uvedené v Tab.1, parametre prevodoviek sú v Tab.2.

Tabuľka 1
 Parametre motorov

Menovitá otáčky [rpm]	Počet polpárov [-]	Menovitý moment [Nm]	Menovitý prúd [A]	Moment zotrvačnosti [kgcm ²]	Účinnosť [%]
6000	6	0,8	1,3	0,650	88,0

Tabuľka 2
 Parametre prevodovky

Prevodový pomer [-]	Menovitý výstupný moment [Nm]	Maximálne vstupné otáčky [rpm]	Vôľa v zuboch [arcmin] PG II / PG IIR	Účinnosť [%] PG II / PG IIR
25,0	40,0	6000	≤ 9,0 / ≤ 13,0	≤ 94 / ≤ 90

Servomotory sú napájané z frekvenčných meničov Sinamics S120. Sinamics S120 je modulárny systém určený pre servopohony. Na výber sú riadené alebo neriadené usmerňovače, vrátane rekuperačných, ktoré napájajú spoločný jednosmerný medziobvod (DC-link). Na medziobvod sú pripojené striedače, ktoré napájajú jednotlivé servomotory. Uvedená zostava je výhodná aj pri linkách, kde umožňuje prelievanie energie medzi pohonmi bez toho, aby sa energia vracala do napájacej siete a opätovne z nej odoberala. Riadenie striedačov zabezpečuje riadiaca jednotka, v našom prípade CU320-2 DP, ktorá dokáže obslúžiť naraz šesť pohonov pracujúcich v režime Servo. Na napájanie fyzikálneho modelu bol použitý rekuperačný

usmerňovač typu Smart Line Module. Na spoločný medziobvod sú pripojené tri dvojmotorové striedače typu Double Motor Module. Riadiaca jednotka komunikuje s PLC prostredníctvom zbernice Profibus. Parametre usmerňovača sú uvedené v Tab.3, parametre motorových modulov v Tab.4.

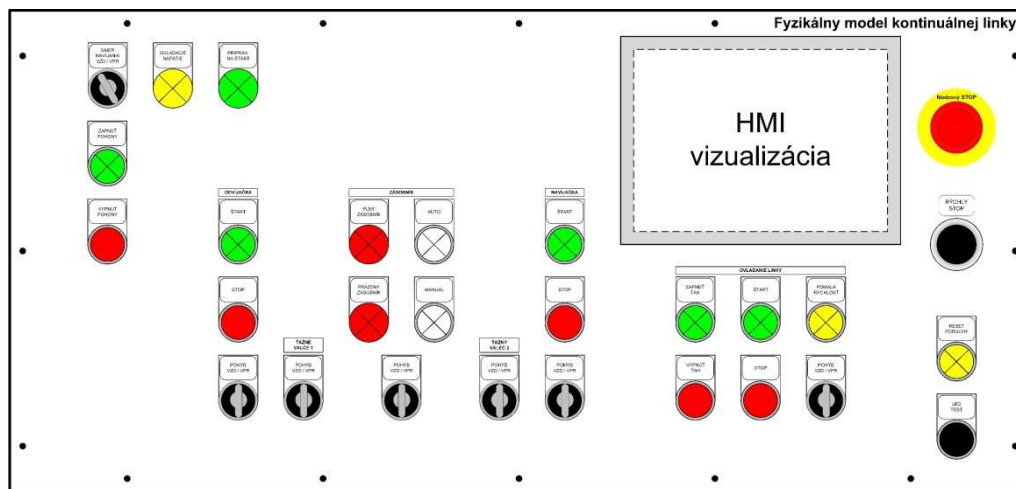
Tabuľka 3
Parametre usmerňovača

Vstupné napätie [V]	Menovitý výkon [kW]	Výstupné napätie medziobvodu [V]	Menovitý výstupný prúd usmerňovača [A]
3AC 380 - 480	5,0	600,0 DC	8,3 DC

Tabuľka 4
Parametre dvojitého motorového modulu

Menovitý výstupný výkon [kW]	Výstupné napätie [V]	Menovitý výstupný prúd [A]	Maximálny výstupný prúd [A]
2 x 1,6	0 – 480 AC	2 x 3,0 AC	2 x 9,0 AC

Ovládanie modelu linky je z operátorského pultu (Obr. 5). Okrem tlačidiel a prepínačov sa tu nachádza aj HMI rozhranie, tvorené dotykovým panelom Siemens KTP700 Basic. Pomocou vizualizácie sa dajú prepínať jednotlivé riadiace štruktúry, nastavovať požadované parametre riadenia ako je ťah v jednotlivých sekciách, rýchlosť linky, nastavenie priemerov navijacky, odvíjačky a pod. Operátorský pult umožňuje individuálne ovládanie každého pohonu a zobrazovanie aktuálnych hodnôt vybraných veličín ako sú rýchlosti, prúdy, momenty, teploty a pod. Vizualizácia zobrazuje aj chybové hlásenia z diagnostiky pohonu. Monitorovanie a archivácia priebehov z technologickej regulácie sa dá robiť jednak prostredníctvom vizualizácie v programe WinCC alebo s využitím dataloga iba-PDA [5].



Obr. 5 Operátorský pult

C. Parametre linky

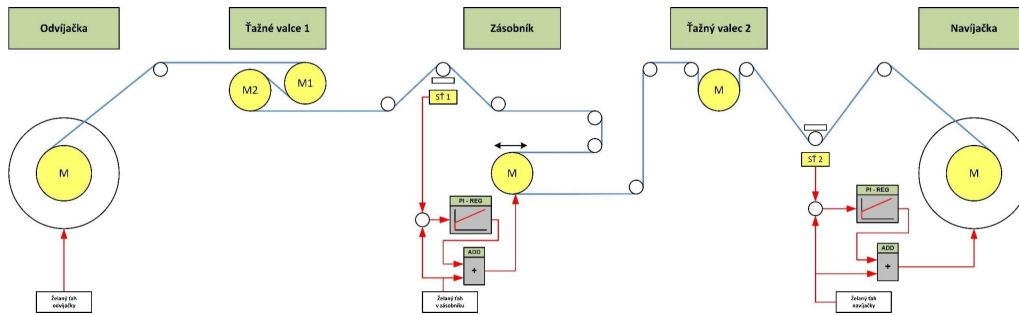
Mechanická konštrukcia linky bola navrhnutá tak, aby vydržala aj menej šetrné zaobchádzanie, ku ktorému nechceme dochádzať pri experimentovaní s riadením na cvičeniach. Na základe toho, ako boli ukotvené motory a snímače, boli definované základné parametre linky. Materiál, ktorý sa v linke prevíja, je celuloidový filmový pás šírky 35 mm s perforáciou na okrajoch. Rýchlosť linky je ohraničená veľkosťou zásobníka a maximálnou konštrukčnou rýchlosťou lineárneho vedenia. Základné mechanické parametre linky sú uvedené v Tab.5, parametre prevíjaného materiálu sú v Tab.6.

Tabuľka 5
Parametre fyzikálneho modelu linky

Max. rýchlosť linky [m/min]	Zrýchlenie a spomalenie linky [s]	Max. ťah v páse [N]	Dĺžka zásobníka [m]	Max. rýchlosť zásobníka [m/min]	Priemer valcov [mm]	Navijacka min. priemer [mm]	Navijacka max. priemer [mm]
10,0	5,0	200,0	0,6	5,0	80,0	100,0	195,0

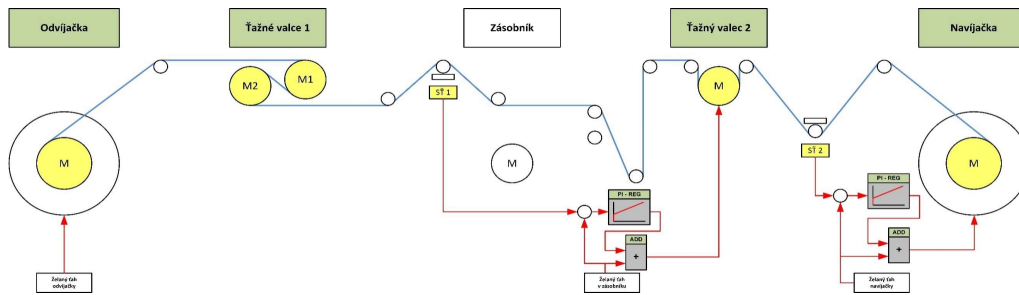
D. Riadiace štruktúry

Model linky má celkovo šesť regulovaných pohonov, pričom modul S-valcov (Ťažné valce 1) má dva pohony. Pokiaľ sa využijú všetky komponenty, na modeli sa dá nastaviť štruktúra linky podľa Obr. 6. Pokiaľ sa vynechá Zásobník materiálu, dostaneme štruktúru na Obr. 7, a ak sa vynechá aj Ťažný valec 2, dostaneme štruktúru na Obr. 8.

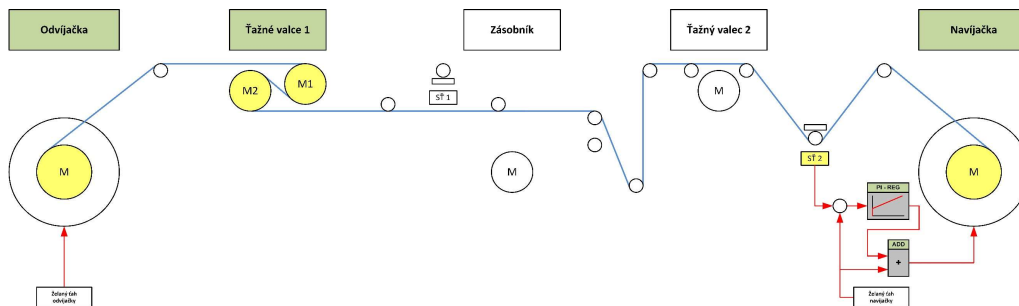


Obr. 6 Kompletná riadiaca štruktúra

Okrem toho je možné voliť, či v regulácii využijeme snímače ťahu, alebo sa vynechajú aj tie, a ťah sa bude regulovať nepriamo prostredníctvom momentu odvíjačky resp. navíjačky. Pomocou snímačov ťahu sa dá potom kontrolovať presnosť nepriamej regulácie.



Obr. 7 Riadiaca štruktúra bez zásobníka materiálu



Obr. 8 Riadiaca štruktúra iba s ťažnými valcami 1 ako hlavným rýchlostným pohonom

Vo všetkých štruktúrach sa musí zvoliť jeden pohon, ktorý bude plniť funkciu hlavného rýchlostného pohonu, a ten bude určovať rýchlosť celej linky. Voči tejto rýchlosti sa budú nastavovať rýchlosti ostatných pohonov. Funkciu hlavného pohonu môže plniť každý z pohonov okrem zásobníka. Vďaka tejto univerzálnosti je možné na jednom modeli skúšať rôzne konfigurácie kontinuálnej linky, čo je pre výučbu výhodné.

IV. PRÍKLAD RIADIACEJ ŠTRUKTÚRY

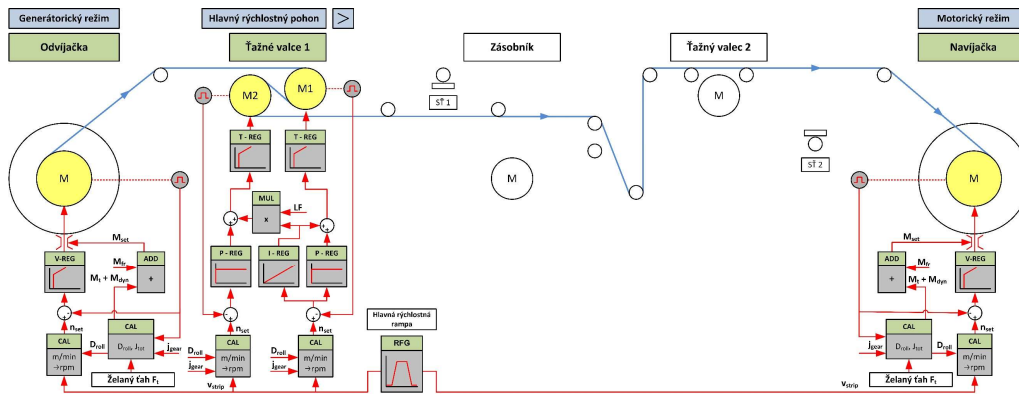
A. Realizovaná štruktúra

Ako príklad realizácie bola zvolená štruktúra podľa Obr. 9. Je to jednoduchá štruktúra s odvíjačkou, navíjačkou a S-valcami. Hlavným rýchlostným pohonom sú S-valce. Odvíjačka aj navíjačka pracujú v ťahovom režime a udržiavajú voči S-valcom želaný ťah. Keďže sa v regulácii nevyužívajú snímače ťahu, oba pohony majú nepriamu reguláciu ťahu. Ťah je

udržiavaný momentom motora s kompenzáciou trenia a dynamického momentu. Želaný ťah je v oboch prípadoch 20 N.

Nastavenie linky vyžaduje výpočet referencií otáčok a ťahu. Rýchlosť v_{strip} a zrýchlenie ε_{set} pre celú linku generuje spoločný hlavný rampový generátor (RFG). Navíjačka aj odvíjačka pracujú v ťahovom režime. Celkový želaný moment M_{set} obsahuje okrem ťahového momentu M_t aj kompenzáciu trenia M_{fr} a dynamického momentu pohonu M_{dyn} . Moment motora sa riadi cez obmedzovač na výstupe regulátora otáčok.

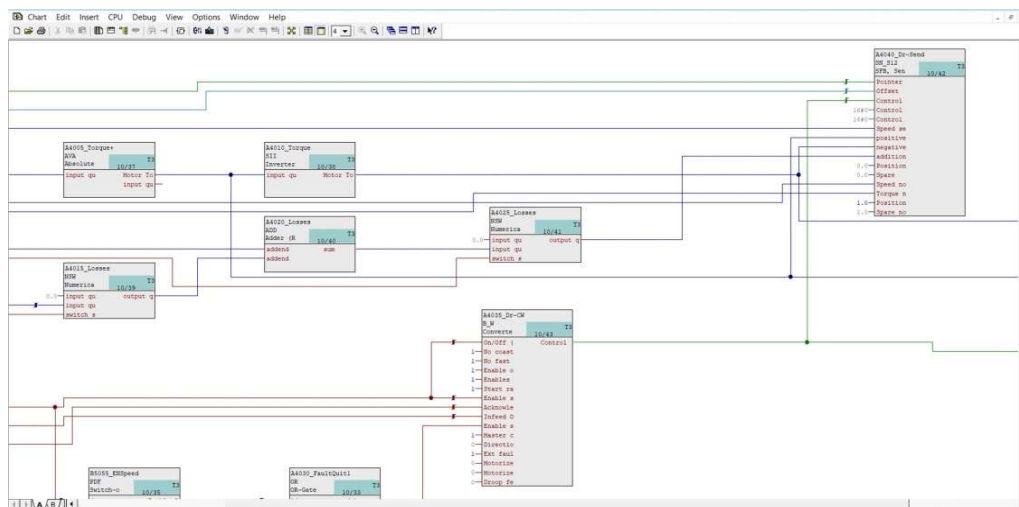
Ťahový moment je úmerný želanéj ťahovej sile F_t , dynamický moment M_{dyn} je úmerný celkovému momentu zotrvačnosti pohonu J_{tot} . Celkový moment zotrvačnosti pohonu je zasa daný momentom zotrvačnosti servomotora J_{mot} a momentom zotrvačnosti prevodovky J_{gear} . Pri navíjačke a odvíjačke obsahuje celkový moment zotrvačnosti aj moment zotrvačnosti tŕňa.



Obr. 9 Realizovaná štruktúra

S-valce sú zaraďované do linky tam, kde sa vyžaduje veľká zmena ťahu. Model obsahuje S-valce (Ťažné valce 1) zložený z dvoch motorov. Ak predpokladáme, že oba pohony budú mať kompenzované trenie a dynamický krútiaci moment, I-zložka regulátora otáčok oboch pohonov bude úmerná len ťahovému momentu. S-valce sú riadené tak, že jeden z motorov (M1) je vodiaci (master). Jeho regulátor otáčok je typu PI. Druhý motor (M2) je vedený (slave) a jeho regulátor otáčok je typu P. Celkový moment sa rozdelí medzi pohony S-valcov tak, aby sa pri danom koeficiente trenia a príslušnom opásaní dosiahla požadovaná zmena ťahu. Rozdelenie ťahov sa robí prostredníctvom parametra LF . I-zložka z mastra je prepočítaná cez LF a pridaná na výstup regulátora otáčok slave pohonu.

Technologické riadenie je realizované na technologickej karte FM 458-1 DP s využitím programu Continuous Function Chart (CFC) [6], ktorý je súčasťou balíka STEP 7. CFC je grafický editor, ktorý umožňuje skladat' regulačné schémy z jednoduchých blokov (Obr. 10). Výsledná schéma je tak prehľadná.



Obr. 10 Príklad schémy kreslenej v programe CFC

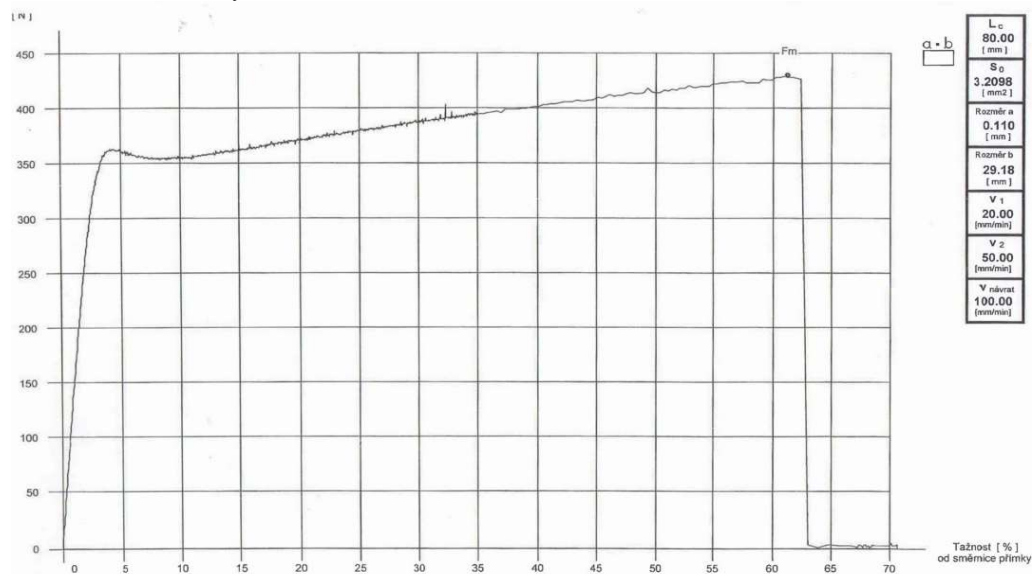
B. Parametre spracovávaného materiálu

Ako spracovávaný materiál bol použitý klasický filmový pás s postrannou perforáciou šírky 35 mm. Na overenie pevnosti pásu bola vykonaná trhacia skúška. Vlastnosti materiálu sú uvedené v Tab.6, kde a_0 je hrúbka pásu, b_0 je šírka pásu, S_0 je prierez pásu, F_m je sila, pri ktorej došlo k roztrhnutiu pásu a σ_m je mechanická pevnosť v ťahu. Pri trhacej skúške bola odstránená perforácia, ktorá skresľovala hranicu pevnosti materiálu, preto je šírka pásu v Tab.6 menšia.

Tabuľka 6
Parametre prevíjaného materiálu

a_0 [mm]	b_0 [mm]	F_m [N]	S_0 [mm ²]	σ_m [MPa]
0,11	29,18	431	3,21	134

Keď sa robila trhacia skúška s pásom s perforáciou, došlo k roztrhnutiu už pri menšej sile, ako udáva Tab.6, a to cca 350 N. Preto sa pri práci s linkou využíva len oblasť do 200 N, kedy sa pás nachádza v oblasti pružnej deformácie, a zároveň má dostatočnú rezervu v pevnosti. Graf z trhacej skúšky pásu je na Obr. 11. Tvar krivky je podobný ako pri oceľových pásoch, len s nižšími hodnotami sily.



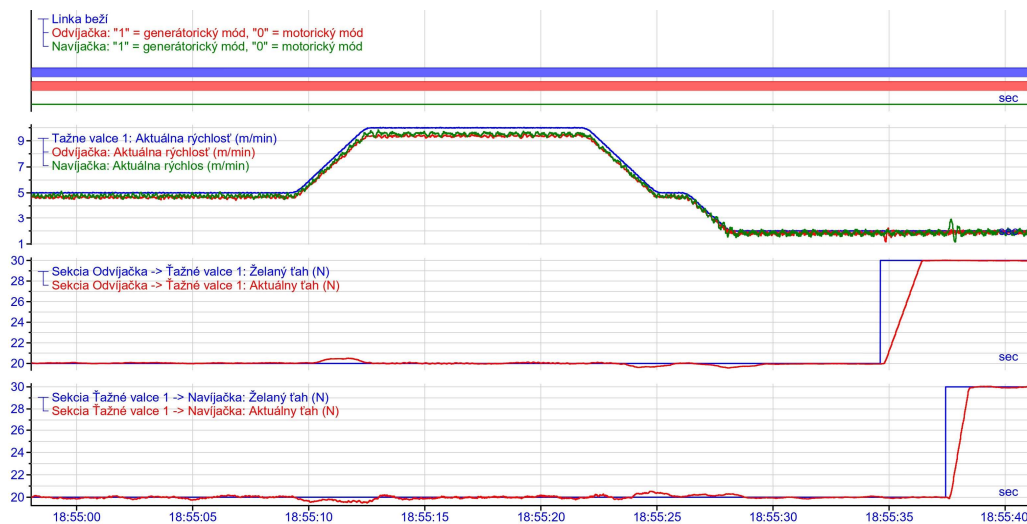
Obr. 11 Graf z trhacej skúšky pásu

C. Namerané priebehy

Regulačná štruktúra podľa konfigurácie na Obr. 9 bola testovaná v nasledovnom režime: linka bola rozbehnutá na rýchlosť 5.0 m/min. V čase 18:55:09 bola rýchlosť zvýšená na 10.0 m/min a následne znížená na 6.0 m/min a v čase 18:55:25 na rýchlosť 3.0 m/min. Monitorovanie linky bolo pomocou programu iba-PDA. Namerané priebehy sú na Obr. 12. Horný graf ukazuje logické signály, ktoré indikujú chod linky (log. hodnota 1), generátorický chod odvíjačky (log. hodnota 1, t.j. odvíjačka brzdí pás) a motorický chod navíjačky (log. hodnota 0, t.j. navíjačka napína pás). Oba pohony pracujú v ťahovom režime. Na ďalšom grafe sú okamžité obvodové rýchlosti jednotlivých pohonov a pod ním ťahy medzi odvíjačkou a S-valcami a S-valcami a navíjačkou.

Z priebehov aktuálnych ťahov v jednotlivých sekciách môžeme pozorovať prekmity ťahu v páse počas zrýchľovania a spomaľovania. Tieto prekmity predstavujú približne 2.5% želaného ťahu, čo môže byť spôsobené nepresným nastavením kompenzácie dynamického momentu. Nakoľko sa jedná o malé rozdiely, v reálnej regulácii to nepredstavuje problém.

Rozdiel aktuálnej rýchlosti medzi odvíjačkou, S-valcami a navíjačkou je spôsobený nepresným meraním priemerov zvitkov.



Obr. 12 Namerané priebehy realizovanej riadiacej štruktúry

V. ZÁVER

Fyzikálny model kontinuálnej linky obsahuje základné komponenty, ktoré sa vyskytujú na reálnych linkách. Tieto komponenty možno usporiadať do rôznych konfigurácií a tak testovať rôzne typy liniek. Použité zariadenia, programové vybavenie aj spôsob ovládania linky je podobný ako v reálnej prevádzke. Vďaka tomu študenti získajú základné informácie i zručnosti s riadením a ovládaním takýchto liniek v praxi.

POĎAKOVANIE

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23 a s podporou Agentúry na podporu výskumu a vývoja na základe Zmluvy č. APVV-15-0750.

POUŽITÁ LITERATÚRA

- [1] BWG Bergwerk- und Walzwerk- Maschinenbau GmbH. *Anniversary Book*. Duisburg, Germany. 2015
- [2] J. Jadlovský, M. Čopík, P. Papcun: *Distribúované systémy riadenia*. Elfa Košice, 2013. ISBN 978-80-8086-227-5.
- [3] *SIMOTION Line Tension Control*. Entry-ID: 56293754, V2.2.1, 01/2019. Prístup: <https://support.industry.siemens.com/cs/ww/en/view/56293754>
- [4] <https://www.br-automation.com/en/products/software/automation-software/automation-studio/>
- [5] <https://www.iba-ag.com/en/ibapda>
- [6] CFC for SIMATIC S7. 11/2006. A5E00807939-01. SIEMENS 2006.

Vývoj výskumnej hardvérovej platformy na báze multikoptéry

¹Ján BAČÍK, ²Pavol FEDOR, ³Daniela PERDUKOVÁ

^{1,2,3}Katedra Elektrotechniky a Mechatroniky, Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská republika

¹jan.bacik.2@tuke.sk, ²pavol.fedor@tuke.sk, ³daniela.perdukova@tuke.sk

Abstract—V tomto príspevku je opísaný vývoj univerzálnej hardvérovej platformy na báze multikoptéry, určenej pre výskum a vývoj navigačných a riadiacich algoritmov z oblasti mobilnej robotiky. Vyvinutá platforma predstavuje malú multikoptéru, umožňujúcu akrobatický let, vybavenú senzorickým systémom schopný presnej detekcie polohy s presnosťou na pár jednotiek milimetrov, a pozemnú stanicu, ktorá slúži na korekciu chyby navigácie a prenos riadiacich a navigačných signálov.

Keywords— mobilná robotika, senzorický systém, multikoptéra, RTK

I. ÚVOD

S mobilnými robotickými systémami sa v súčasnosti človek v bežnom živote stretáva čoraz častejšie. Tieto robotické systémy si našli svoje uplatnenie nie len v priemysle ale aj v bežných domácnostiach v podobe automatizovaných vysávačov a kosačiek [1]. Taktiež sa tieto domáce roboty stali obľúbenou platformou rôznych výskumných inštitúcií zaoberajúcich sa vývojom mobilných robotických systémov [2]. Výhodou týchto platforiem je ich univerzálnosť a možnosť doplnenia systému o ďalšie senzory a výpočtové jednotky. V oblasti priemyslu si mobilné robotické systémy našli uplatnenie ako logistické ťahače vo veľkých výrobných halách, kde je ich úlohou autonómny transport materiálu z miesta na miesto [3].

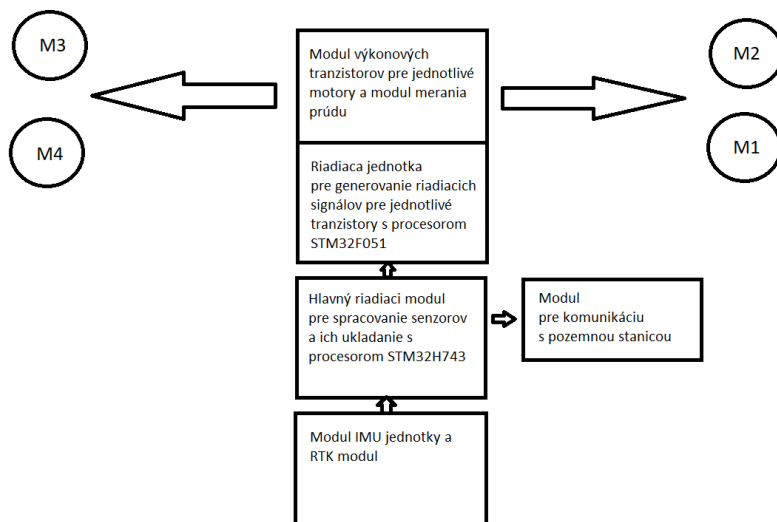
Okrem klasickej kolesovej mobilnej robotiky sa čoraz viac vývoj sústreďuje aj na využitie multikoptér, poprípade lietadiel s pevným krídlom, na plnenie rôznych komerčných úloh. Či už ide o autonómny prieskum a pozorovanie alebo poprípade presun ľahkého materiálu z miesta na miesto. Pre zabezpečenie týchto typov úloh je potrebné, aby zariadenia boli vybavené presným navigačným systémom, ktorý by samotný autonómny let umožňoval.

V článku sme zamerali pozornosť na vytvorenie malej lietajúcej hardvérovej platformy, ktorá by obsahovala všetky komponenty potrebné pre riadený let, získavania dát zo senzorov a ich následné ukladanie pre ďalšiu analýzu. Jednotlivé hardvérové moduly tvoria podsystémy riadenia motorov, podsystémy senzorického systému, podsystémy riadenia a vyhodnocovania a ukladania dát. Samostatný hardvérový modul tvorí aj pozemná stanica, ktorej úlohou je zabezpečiť korekčné navigačné dáta a odosielanie riadiacich signálov pre multikoptéru.

II. POPIS MODULOV SYSTÉMU MULTIKOPTÉRY

Hlavný riadiaci modul pozostáva z procesora typu ARM, konkrétne STM32H743. Tento procesor obsahuje dostatočný počet komunikačných periférií, aby dokázal komunikačne prepojiť všetky zvyšné moduly a senzory (Obr. 1). Zároveň umožňuje ukladať dáta na pamäťovú SD kartu. Výmena dát medzi procesorom a pozemnou stanicou je zabezpečená modulom pre komunikáciu s pozemnou stanicou.

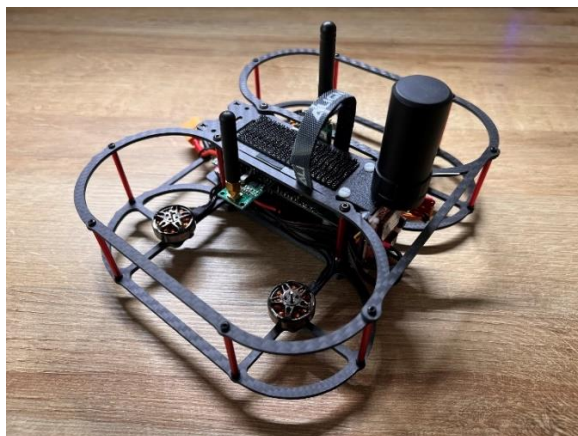
Komunikačný modul pozostáva z čipu NRF905, ktorý pracuje na nosnej frekvencii 868MHz a je schopný komunikovať prenosovou rýchlosťou 100kbps. Pre zrobustnenie komunikácie je okrem čipu NRF905 v module použitý aj sekundárny čip SX1278 komunikujúci na frekvencii 433MHz. Dáta, ktoré sa vymieňajú medzi multikoptérou a pozemnou stanicou sú paralelne vysielané a prijímané cez obe komunikačné čipy naraz a následne kontrované procesorom hlavného riadiaceho modulu.



Obr. 1 Bloková schéma zapojenia modulov multikoptéry

Dôležitou súčasťou celého systému sú senzory určené pre určenie aktuálnej polohy a orientácie multikoptéry počas letu. Okrem nevyhnutného troj-osého gyroskopu spojeného s akcelometrom a magnetického senzoru, ktoré slúžia prevažne na určenie orientácie v priestore [4], je celý systém doplnený o RTK (Real Time Kinematics) modul [5]. RTK modul obsahuje GNSS snímač, ktorý prijíma signály z globálnych navigačných systémov a zároveň má komunikačný kanál pre príjem korekčných dát. Zdrojom korekčných dát môže byť RTCM korekčný tok vysielaný poskytovateľom tejto služby, napr. UNAVCO, alebo zdrojom korekčného signálu môže byť separátne RTK modul, nastavený ako pozemná stanica, s podmienkou jeho presného umiestnenia v rámci globálneho pozičného systému. Práve z tohto dôvodu je celá hardvérová platforma pozostávajúca nie len z multikoptéry samotnej (Obr. 2), ale aj pozemnej stanice.

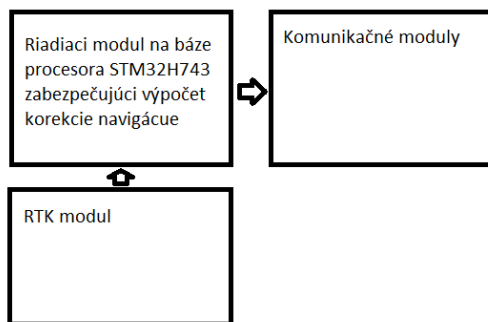
Posledný z modulov multikoptéry je riadiaci obvod pre rýchlostné riadenie jednotlivých BLDC motorov pozostávajúci z výkonových MOSFET tranzistorov. Pre generovanie hradlových signálov tento modul obsahuje samostatný procesor typu ARM, STM32F051.



Obr. 2 Multikoptéra s osadenými modulmi

III. POZEMNÁ STANICA

Ako už bolo spomenuté, hlavnou úlohou pozemnej stanice je spracovanie korekčných dát z RTK modulu, a ich následne odoslanie do riadiaceho procesora multikoptéry. Rovnako, ako multikoptéra, obsahuje pozemná stanica dvojicu komunikačných čipov pre odosielanie a príjem dát, a RTK modul (Obr. 3).



Obr. 3 Bloková schéma zapojenia modulov pozemnej stanice

Jednotlivé moduly pozemnej stanice sú uložené v karbónovom šasi (Obr. 4), ktoré bolo vyrobené presne na mieru. Veľké červené tlačidlo centrálného zastavenia v súčasnosti nemá využitie ale v budúcnosti je plánované ako povel pre zastavenia multikoptéry na mieste počas automatického letu. Súčasťou stanice sú aj USB prevodníky umožňujúce komunikačné prepojenie stanice a počítača.



Obr. 4 Pozemná stanica

IV. ZÁVER

V článku sme stručne predstavili nami vyvinutú univerzálnu hardvérovú platformu na báze multikoptéry, ktorá obsahuje všetky potrebné elektronické a senzorické komponenty potrebné pre určenie presnej polohy a orientácie multikoptéry počas letu. Táto platforma umožňuje ďalší vývoj v oblasti spracovania dát zo senzorov a následne vývoj algoritmov umožňujúcich automatický let.

V. POĎAKOVANIE

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23.

REFERENCIE

- [1] J. Forlizzi, C. DiSalvo, Service robots in the domestic environment: a study of the roomba vacuum in the home, proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction, New York, 2006.
- [2] M. Dekan, L. Chovanec, A. Babinec, A. Vitko, New modules for the iRobot Create platform, Procedia Engineering, 2012.
- [3] Ceit Logistic Trucks: available on web: <http://www.ceittechnnovation.eu/index.php/sk/agv-system>
- [4] What is RTK, anyway?: available on web: <https://www.sparkfun.com/rtk>
- [5] D.H. Titterton, J.L. Weston, Strapdown Inertial Navigation Technology, second edition, Institution of Engineering & Technology, Reston, 2005.

Využitie Technológie Blockchain v Procese Smerovania v Mobilných Viacpreskokových Siet'ach

¹Maroš BAUMGARTNER, ²Ján PAPAĽ

^{1,2} Katedra elektroniky a multimediálnych telekomunikácií, Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská Republika

¹maros.baumgartner@tuke.sk, ²jan.papaj@tuke.sk

Abstrakt — Proces smerovania v mobilných viacpreskokových siet'ach je jeden z najdôležitejších mechanizmov, ktorý vplýva na robustnosť a efektívnosť siete. Z vlastností MANET sietí vyplýva, že sa do tejto siete a procesu smerovania môže pripojiť akýkoľvek mobilný uzol, ktorý disponuje základným technickým vybavením potrebným na prenos dát, avšak do siete sa môže pripojiť aj mobilný uzol so zlomyseľným správaním, ktorý úmyselne zabraňuje komunikácii v rámci siete prostredníctvom zahadzovania smerovacích a dátových paketov. Z tohto dôvodu je potrebné navrhnúť mechanizmus pre výber vhodných mobilných uzlov pomocou technológie blockchain. Tento príspevok pojednáva o analyzovaní smerovacieho protokolu AODV s implementovanou technológiou blockchain v procese smerovania pri simulovaných útokoch na sieť.

Kľúčové slová — MANET, blockchain, smerovanie, robustnosť

I. ÚVOD

Mobilné bezdrôtové technológie a bezdrôtová komunikácia sú v súčasnosti najdynamickejšie rozvíjajúca sa metóda slúžiaca na komunikáciu ktorú využívajú ľudia na celom svete a mimo neho. V dnešnej dobe má pri sebe takmer každý obyvateľ v civilizovanom svete zariadenie slúžiace na komunikáciu, či už to sú mobilné telefóny, laptopy, tablety, chytré hodinky a pod. Každé z týchto zariadení je vybavené technológiami WiFi alebo Bluetooth, ktoré sú určené na komunikáciu v rámci krátkej vzdialenosti, avšak s pokrokom telekomunikačných technológií sa tieto zariadenia stávajú súčasťou mobilných bezdrôtových sietí nazývaných MANET (Mobile Ad-Hoc Network) vďaka ktorým sa vieme pripojiť k internetovému pripojeniu a komunikovať s kýmkoľvek bez ohľadu na vzájomnú vzdialenosť. MANET siete [1] disponujú určitými výhodami, ktoré sú zároveň aj ich základnými vlastnosťami. V MANET sieti môže byť každý účastník, ktorí je v tejto terminológii nazývaný uzol, zdrojom, cieľom alebo akousi prestupnou stanicou cez ktorú prúdia dátové toky od zdrojového uzla k cieľovému uzlu [2].

MANET siete sa svojmu rozmachu tešia najmä kvôli začínajúcemu trendu nasadzovania 5G sietí do komerčného využitia. S nástupom sietí piatej generácie vedci začali intenzívne implementovať do tejto architektúry aj iné typy mobilných sietí. VANET (Vehicle Ad-Hoc Network) [3][4] vozidlové bezdrôtové siete, ktoré svoje využitie uplatňujú v automobilovej komunikácii počas premávky, DroNet (Drone Network) [5] sieť v ktorej sa jednotlivé uzly nazývajú UAV (Unmanned Aerial Vehicle) a jej využitie je zamerané najmä pre bezpilotné siete. Určitou obdobou DroNet je FANET (Flying Ad-Hoc Network) [5], ktorá je však určená pre dopravné, súkromné a komerčné lietadlá. Okrem rôznych typov mobilných bezdrôtových sietí sú spolu so sieťami 5G spájané aj komunikácie typu D2D (Device to Device) [6] a viacpreskokové D2D. Tie majú svojimi vlastnosťami prevyšovať MANET siete najmä v rýchlosti dátového prenosu a priepustnosti siete [7].

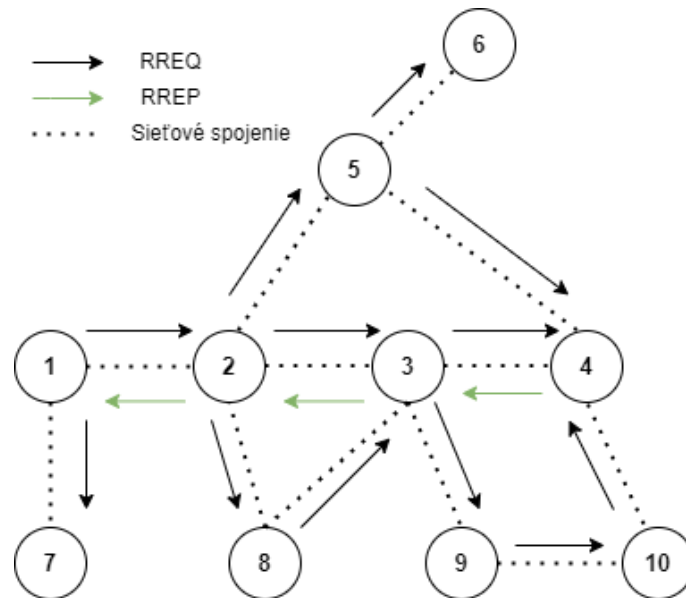
Na vyriešenie týchto a ďalších bezpečnostných problémov sa v dnešnej dobe používajú rôzne metódy a technológie. Veľmi účinnou metódou, ktorá dokáže zlepšiť odolnosť siete je technológia blockchain. Blockchain je určitá forma databázy, ktorá sa od klasickej databázy údajov líši niekoľkými zásadnými vlastnosťami. Je to databáza, ktorá v sebe ukladá informácie o smerovaní, preposlaných paketoch, dôveryhodnosti uzlov a pod., tieto informácie sú ukladané do blokov, ktoré tvoria vzájomnú nadväznosť čím neumožňujú pozmenenie alebo vymazania dát. Takáto databáza blockchain je zdieľaná medzi všetkými účastníkmi siete a platí princíp, že každý

vie o každom všetko. Teda ak by sa medzi uzlami objavil zlomyseľný uzol, jeho dôveryhodnosť klesne a bude vyradený z procesu komunikácie. Takýmto spôsobom sa dá jednoducho zabrániť vniknutiu akýchkoľvek narušiteľov do siete a tým výrazne zlepšiť samotnú odolnosť siete.

II. PROCES SMEROVANIA V MOBILNÝCH VIACPRESKOOVÝCH SIEŤACH

A. Smerovací protokol Ad Hoc On-Demand Distance Vector

Podobne ako všetky smerovacie protokoly patriace do skupiny reaktívnych smerovacích protokolov aj AODV (Ad Hoc On Demand Distance Vector) [8] protokol komunikuje až na základe požiadavky. Smerovacie informácie v AODV sa prenášajú v tzv. Route Request (RREQ) a Route Reply (RREP) paketoch (Obr. 1). Keď zdrojový uzol iniciuje požiadavku komunikovať s cieľovým uzlom, pošle RREQ paket smerom k cieľovému uzlu. Tento paket sa prenáša cez všetky susedné uzly, ktoré ju opätovne vysielajú k svojim susedom, až kým nebude prijatý uzlom, ktorý má informáciu o cieľovom uzle alebo o uzle, ktorý je bližšie k cieľovému uzlu. V prípade, že cieľový uzol prijme RREQ paket a overí IP adresu uloženú v tomto pakete vyšle RREP paket prostredníctvom unicastového vysielania po vzniknutej komunikačnej trase späť k zdrojovému uzlu. Na kontrolu a aktualizáciu smerovacích informácií sa využívajú tzv. Sekvenčné čísla (angl. Sequence numbers). Každý uzol má vlastný zoznam, ktorý zaznamenáva čísla posledných použitých RREQ a RREP paketov. Keď uzol dostane nový RREQ správu, porovná ju s posledným použitým RREQ číslom. Ak je toto číslo menšie alebo rovné číslu v RREQ pakete, uzol prijme RREQ paket a aktualizuje svoj poradovník. V opačnom prípade správu zahodí.



Obr. 1 Príklad funkcionality smerovacieho protokolu AODV

Za účelom navýšenia výkonu a zlepšenie optimalizácie sú v AODV protokole implementované rôzne mechanizmy. Okrem sekvenčných čísel je v RREQ pakete parameter TTL (Time To Live). Parameter TTL je prednastavený na určitú hodnotu, ktorá s pribúdajúcim počtom skokov klesá, ak sa tomuto paketu nepodariť nadviazať komunikačnú trasu s cieľovým uzlom bude tento paket zahodený a zdrojový uzol bude nútený opätovne iniciovať proces smerovania. Táto funkcionality siete môže znížiť réžiu AODV protokolu, ale taktiež negatívne vplyva na latenciu pre získanie komunikačnej trasy. Ďalším mechanizmom slúžiacim na udržanie komunikačnej trasy je periodické vysielanie HELLO paketov. Tieto typy paketov dopĺňajú pasívne a aktívne ACK (Acknowledgment) pakety na detekciu prerušenia komunikačnej trasy, teda AODV využíva aj proaktívny prístup k detekcii prerušenia s myšlienkou zvýšenia stability siete.

B. Technológia blockchain

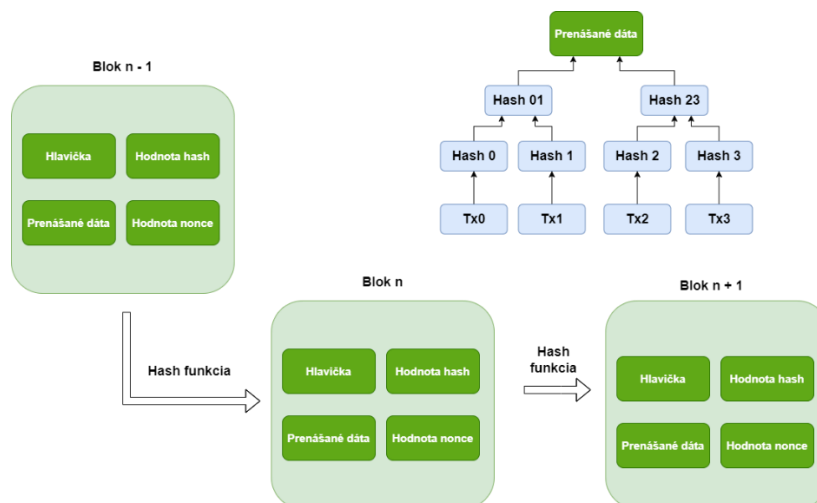
Poskytovanie bezpečného smerovania v bezdrôtových sieťach, ako sú MANET siete, je náročná a zložitá záležitosť z dôvodu mobilného charakteru týchto sietí. Hlavnou výzvou pre aplikovanie technológie blockchain (BC) v MANET sieťach je prispôbenie MANET siete na výpočtovú náročnosť blokovej validácie pri zachovaní charakteristík BC a zahrnúť uzly do procesu overovania. Úspešné realizovanie hodnoty hašovacej funkcie na prepojenie údajov a zdieľanie prepojených dát s uzlami v sieti BC ponúka autentifikáciu pre viacpreskokové siete.

Nový blok, ktorý bol vytvorený, má aktuálne informácie o všetkých ostatných blokoch v reťazci. Tieto bloky sú chránené digitálnym podpisom dôveryhodných uzlov, ktoré riadia a zabezpečujú systém proti útokom a zlomyseľným uzlom. Inteligentná zmluva je použiteľná na vytvorenie Blockchain Contractual Routing protokolu (BCR) pre zariadenia vo viacpreskokových sieťach. Využíva sa na zistenie cesty zo zdrojového uzla do cieľa (namiesto pôvodného RREQ v typických smerovacích protokoloch). Vo viacpreskokových sieťach môžu byť uzly v sieti rozdelené do troch častí podľa prenosu dátového paketu:

- Zdrojový uzol, ktorý je zodpovedný za prenos dátového paketu
- Medziľahlý uzol, ktorý je zodpovedný za preposielanie dátového paketu
- Treťou časťou je cieľový uzol, ktorý je zodpovedný za príjem dátových paketov.

Tieto tri časti sú podľa prenosu paketov vo viacpreskokových sieťach mapované do vzťahu BC a uzlov [10].

BC je reťazec blokov, kde každý blok je braný ako transakcia na celom reťazci, z ktorého tieto dané bloky nemôžu byť odobraté, ale iba pridávané. Každý blok prítomný v reťazci pozostáva z údajov v bloku. Pozostáva z hodnoty hašovacej funkcie, ktorá sa odkazuje na tieto konkrétne údaje na konkrétnom bloku, a predchádzajúcej hodnoty hašovacej funkcie predchádzajúceho bloku. Údaje zaznamenané v BC závisia od typu daného BC. Ak je nový blok vytvorený, vygeneruje sa aj hodnota hašovacej funkcie tohto konkrétneho bloku. Hodnota hašovacej funkcie bloku bude zmenená úpravami vykonanými v bloku [11]. Preto je táto hodnota veľmi dôležitým faktorom pri vykonávaní úprav. Ak je hodnota hašovacej funkcie z ktoréhokoľvek bloku zmenená, potom sa nebude považovať za súčasť toho istého bloku. Okrem hodnoty hašovacej funkcie aktuálneho bloku obsahuje blok aj hodnotu hašovacej funkcie predchádzajúceho bloku. To pomáha vytvoriť reťaz, ktorá sa prepája medzi aktuálnym blokom a predchádzajúcim blokom.



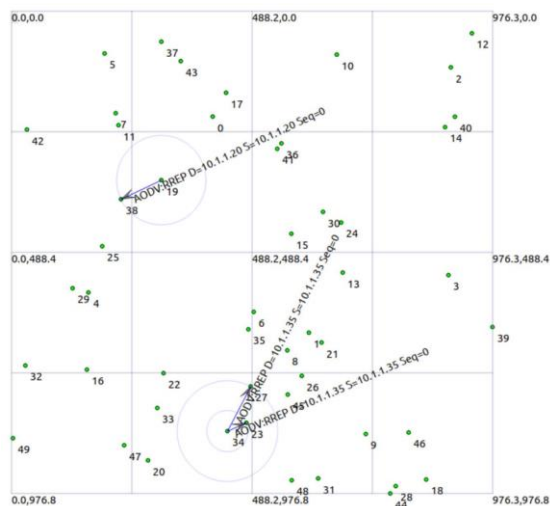
Obr. 2 Štruktúra technológie blockchain

Na Obr. 2 je zobrazený príklad reťazca s tromi blokmi. Každý blok sa skladá z hodnoty hašovacej funkcie aktuálneho bloku a predchádzajúceho bloku. Blok číslo 2 odkazuje na hodnotu hašovacej funkcie bloku číslo 1, blok číslo 3 odkazuje na hodnotu hašovacej funkcie bloku číslo 2 pomocou hodnoty predchádzajúcej hašovacej funkcie. Hodnota hašovacej funkcie prvého bloku je 0000, pretože ide o špeciálny blok, ktorý neodkazuje späť na žiadny blok. Tento blok je známy ako pôvodný blok [9].

III. SMEROVANIE V MOBILNÝCH VIACPRESKOKOVÝCH SIEŤACH NA BÁZE TECHNOLOGIE BLOCKCHAIN

A. Simulačné parametre

Navrhnutý spôsob smerovania v MANET sieťach na báze technológie blockchain a smerovacieho protokolu AODV bol simulovaný vo viacerých testovacích scenároch. V rámci simulácií bolo sledované správanie sa siete (Obr. 3) pri zmene počtu uzlov, rýchlosti pohybu uzlov a hodnoty vysielacieho signálu. Úplné simulačné parametre sú zobrazené v Tab. 1.



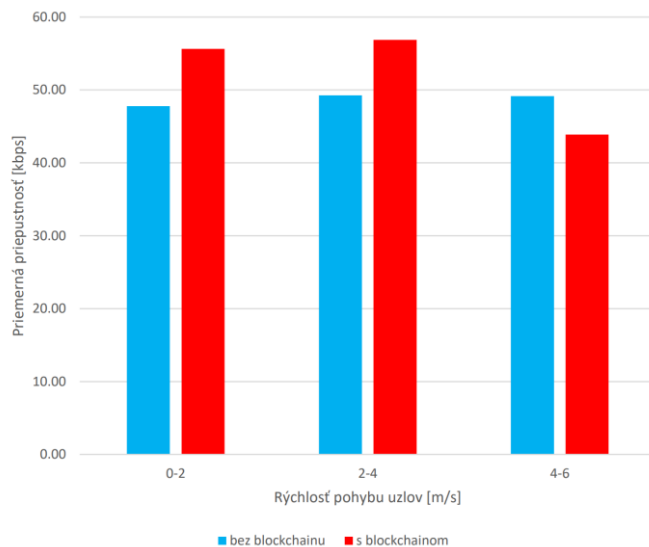
Obr. 3 Navrhnutá MANET sieť

Tab. 1 Simulačné parametre

Simulačný parameter	Hodnota
Počet škodlivých uzlov [%]	60
Rýchlosť pohybu uzlov [m/s]	0 – 2; 2 – 4; 4 – 6
Počet uzlov	50
Simulovaná oblasť [m]	1000 x 1000
Prenosové médium	IEEE 802.11b
Model mobility	Model náhodného pohybu
Simulačné prostredie	Network Simulator 3
Počet simulácií	100
Trvanie simulácie [s]	90

B. Dosiahnuté výsledky

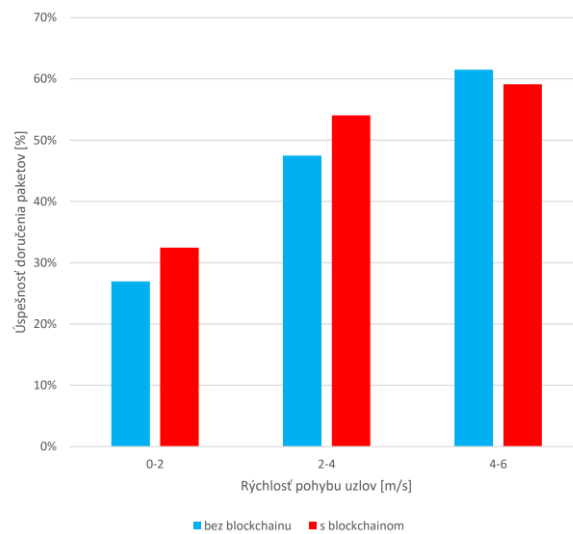
V rámci dosiahnutých výsledkov boli analyzované priemerné hodnoty QoS (Quality of Service) parametrov, priepustnosť, PDR (Počet Doručených Paketov), oneskorenia a kolísania oneskorenia.



Obr. 4 Priemerná priepustnosť siete v závislosti na rýchlosti pohybu uzlov

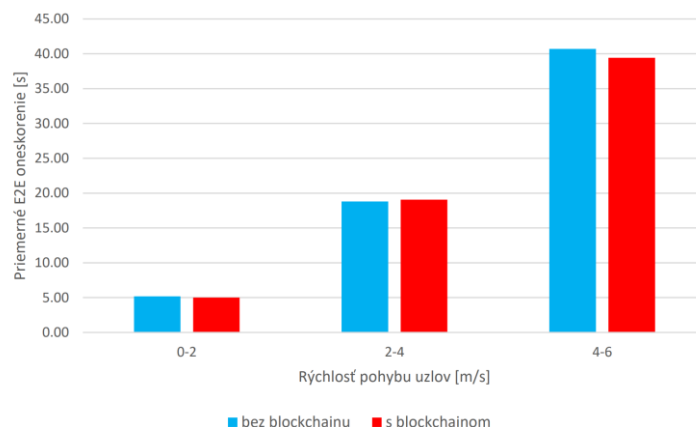
Na Obr. 4 sú zobrazené dosiahnuté priemerné hodnoty priepustnosti siete v závislosti na počte škodlivých uzlov siete a rýchlosti pohybu uzlov. Na základe týchto výsledkov je možné konštatovať, že pri nízkej až miernej rýchlosti mobility 0 – 4 m/s, dokáže smerovací protokol s implementovanou technológiou blockchain zvýšiť priepustnosť siete v priemere o 7 kb/s, čo predstavuje približne 12%. Naopak pri väčšej rýchlosti 4 – 6 m/s bol zaznamenaný negatívny vplyv blockchainu na priepustnosť siete spôsobený najmä výpočtovým zaťažením overovacích uzlov.

Následne bol analyzovaný vplyv rýchlosti pohybu uzlov na parameter PDR (Obr. 5). Zistili sme, že zvyšovaním rýchlosti pohybu uzlov sa zvyšovala aj percentuálna úspešnosť doručenia paketov. To môžeme odôvodniť tým, že na simulačnej ploche 1000x1000 metrov sa pri danej dĺžke simulácie uzly s nízkou rýchlosťou nestihli priblížiť k sebe natoľko, aby mali rádiový dosah. Taktiež bolo zistené, že pri rýchlostiach 0 – 4 m/s implementovaná technológia blockchain pozitívne vplývala na parameter PDR, a tým pádom aj na celkovú odolnosť a efektivitu siete. Pozorované zlepšenie predstavovalo rozdiel 5 – 6%. Pri vyšších rýchlostiach mala technológia blockchain opačný vplyv, pozorovali sme priemerné zníženie úspešnosti doručenia paketov o 3%.



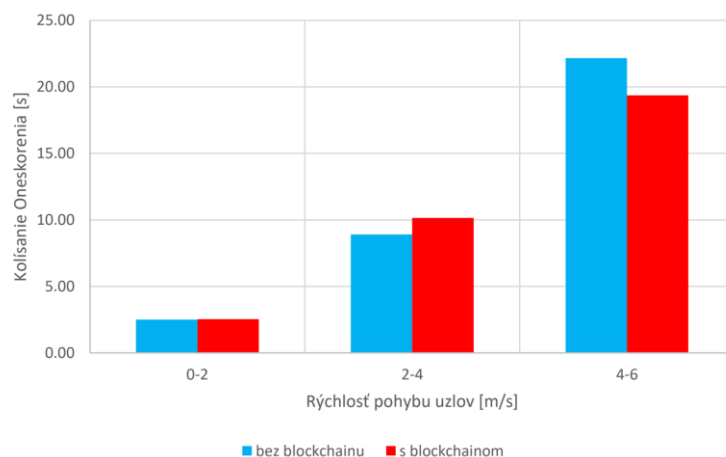
Obr. 5 Priemerný počet doručených paketov v závislosti na rýchlosti pohybu uzlov

V prípade celkového E2E oneskorenia (Obr. 6) pri rýchlosti 4 – 6 m/s je z analyzovaných výsledkov možné konštatovať, že okrem zvýšenia úspešnosti doručenia paketov sa znížilo celkové E2E oneskorenie avšak iba v minimálnej miere. Z toho vyplýva, že technológia blockchain neovplyvňuje celkové oneskorenie siete, výsledky boli vo všetkých troch scenároch takmer totožné.



Obr. 6 Priemerné celkové oneskorenie v závislosti na rýchlosti pohybu uzlov

Keďže so zvyšovaním rýchlosti pohybu uzlov sa zvyšovalo aj celkové E2E oneskorenie, zvyšovalo sa aj kolísanie oneskorenia (jitter). Protokol AODV, ktorý využíval technológiu blockchain, nadobúdala nižšie hodnoty len pri rýchlosti 4 – 6 m/s, pričom rozdiel bol 2,8s. V opačnom prípade boli hodnoty kolísania oneskorenia vyššie, resp. podobné smerovaciemu protokolu AODV bez implementovanej technológie blockchain. Graf kolísania oneskorenia v závislosti na rýchlosti pohybu uzlov je zobrazený na Obr. 7.



Obr. 7 Priemerné kolísanie oneskorenia (Jitter) v závislosti na rýchlosti pohybu uzlov

IV. ZÁVER

Tento príspevok popisuje návrh procesu smerovania v mobilných viacpreskokových sieťach na báze smerovacieho protokolu AODV a implementovanej technológie blockchain. Navrhnutý model bol analyzovaný na základe QoS parametrov oneskorenia, počtu doručených paketov, kolísania oneskorenia a priepustnosti siete. Z dosiahnutých výsledkov je možné konštatovať, že vplyvom implementovanej technológie blockchain sa sieť s postupným zvyšovaním rýchlosti pohybu uzlov a miery škodlivých uzlov v sieti stáva efektívnejšou a robustnejšou. Tento trend je zabezpečený výberom vhodných uzlov na smerovanie, ktoré následne negatívne neovplyvujú na smerovanie a dátový prenos.

POĎAKOVANIE

Tento výskum bol financovaný agetúrou na podporu výskumu a vývoja SR, výskumný grant č. APVV-17-0208 a VEGA 1/0260/23.

POUŽITÁ LITERATÚRA

- [1] R. GOTZHEIN, a kol., 2020. Real-time Communication Protocols for Multi-hop Ad-hoc Networks Wireless Networking in Production and Control Systems. Springer, Computer Communications and Networks. ISSN: 2197-8433, ISBN: 978-3-030-33319-5.
- [2] M. M. A. AZIM, X. JIANG, 2016. Wireless Sensor Multimedia Networks Architectures, Protocols, and Applications. CRC Press. ISBN – 13: 978-1-4822-5313-9.
- [3] K. SUGANTHI, a kol., 2020. Machine Learning and Deep Learning Techniques in Wireless and Mobile Networking Systems. CRC Press, ISBN: 978-1-003-10747-7.
- [4] R. HENRIQUE, a kol., 2001. The Law of Accelerating Returns An Kurzweil. Available: <https://www.kurzweilai.net/the-law-of-accelerating-returns>.
- [5] A. KUMARI, a kol., 2020. A taxonomy of blockchain-enabled softwarization for secure UAV network. ELSEVIER. Journal Computer Communications 161 (2020) 304-323, <https://doi.org/10.1016/j.comcom.2020.07.042>.
- [6] Q. N. TRAN, a kol., 2020. D2D Multi-hop Multi-path Communications in B5G Networks: A Survey on Models, Techniques, and Applications. EAI Endorsed Transactions on Industrial Networks and Intelligent Systems.
- [7] F. F. SHAIKH, a kol., 2018. Routing in Multi-Hop Cellular Device-to-Device(D2D) Networks: A Survey. IEEE COMMUNICATIONS SURVEYS & TUTORIAL.
- [8] C. Perkins, E. Royer, "Ad-hoc on-demand distance vector routing," WMCSA'99. Second IEEE Workshop on Mobile Computing Systems and Applications, New Orleans, LA, USA: IEEE, 1999, pp. 90–100, ISBN: 978-0-7695-0025-6. doi: 10.1109/MCSA.1999.749281. [Online]. Available: <http://ieeexplore.ieee.org/document/749281/>.
- [9] L. BAHRI, a kol., 2019. Blockchain Technology: Practical P2P Computing (Tutorial). IEEE 4th International Workshops on Foundations and Applications of Self* Systems (FAS*W).
- [10] D. YAGA, a kol., 2018. Blockchain Technology Overview. National Institute of Standards and Technology, 2018. Dostupné: <https://doi.org/10.6028/NIST.IR.8202>.
- [11] Z. ZHENG, et al., 2018. Blockchain challenges and opportunities: a survey. Int. J. Web and Grid Services, 2018.

Využitie Blockchainu v procese smerovania vo viacpreskokovej sieti

¹Maroš BAUMGARTNER, ²Ján PAPAJ, ³Matúš SABOL

^{1,2,3} Katedra elektroniky a multimediálnych telekomunikácií, Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská Republika

¹maros.baumgartner@tuke.sk, ²jan.papaj@tuke.sk, ³matus.sabol@student.tuke.sk

Abstrakt — Blockchain je najnovšou a najperspektívnejšou technológiou v dnešnej dobe. Táto technológia dokáže vyriešiť rôzne druhy problémov vo sfére informačných technológií ako napríklad nedostatok dôveryhodnosti, transparentnosti, bezpečnosti a spoľahlivosti spracovania dát. Technológia blockchain teoreticky prináša pozitívne výsledky, ale realita v praxi sa môže líšiť. Táto publikácia sa zaoberá opisom technológie blockchain, MANET sieťami a ich kombináciou za účelom pozorovania odolnosti siete. Boli opísané základné výhody a nevýhody technológie blockchain. Analyzovali sme viaceré implementované riešenia blockchainu v MANET a ich prínosy a nedostatky z perspektívy parametrov kvality služieb. Pomocou sieťového simulátora NS-3 a skriptov v programovacom jazyku C++ bola implementovaná technológia blockchain do mobilného ad-hoc smerovacieho protokolu AODV, ktorý bol následne analyzovaný. Za účelom analyzovania odolnosti smerovacieho protokolu a odolnosti siete vo všeobecnosti bol vytvorený model škodlivých uzlov, ktorý úmyselne zhoršuje parametre kvality služieb siete.

KLúčové slová — blockchain, manet, aodv, simulácia

I. ÚVOD

V tejto publikácii boli opísané typy smerovacích protokolov v MANET. Keď sa spomenie blockchain, väčšina ľudí si predstaví kryptomeny. Technológia blockchain má však omnoho viac využití a kryptomeny sú len jedným z mála existujúcich aplikácií tejto technológie. Blížšie sa blockchainu venujeme v tretej kapitole, kde sme tento pojem objasnili. Bol analyzovaný jeho princíp fungovania a používané parametre. Definovali sme jeho základné stavebné prvky. Štvrtá kapitola sa zaoberá analýzou implementácie blockchainu v MANET jej výsledkami. Na implementáciu bol použitý protokol AODV. Jeho vlastnosti sú odsimulované v sieťovom simulátore NS-3. Analyzovali sme výsledné parametre kvality služieb v závislosti od počtu škodlivých uzlov a rýchlosti pohybu uzlov. V záverečnej časti sú zhrnuté výsledky a tiež je predstavená potenciálna budúcnosť blockchainu v MANET sieťach.

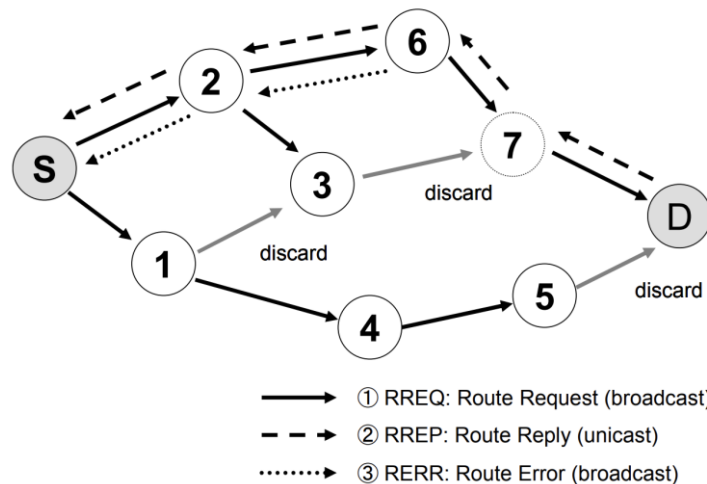
II. SMEROVANIE V MANET SIEŤACH

Smerovanie v MANET sieťach rozlišujeme na základe vlastností konkrétnej implementácie použitého protokolu. V súčasnosti rozdelíme smerovanie v MANET na proaktívne, reaktívne a hybridné. **Proaktívne smerovanie** je založené na fakte, že všetky uzly si udržiavajú kompletne informácie o celej topológii vo svojej smerovacej tabuľke. Aby mohli mať jednotlivé uzly stále aktuálnu topológiu siete, je potrebné pravidelné zaplavovanie siete paketmi. Podstatou **reaktívneho** smerovania je fakt, že mobilný uzol si uchováva informáciu len o komunikačnej ceste medzi zdrojovým a cieľovým uzlom. Samotný proces hľadania novej cesty je aktivovaný iba vtedy, ak nastane situácia, kedy sa aktuálne používaná komunikačná cesta preruší. **Hybridné smerovanie** využíva dobré vlastnosti proaktívneho smerovania a spája ich s dobrými vlastnosťami reaktívneho smerovania, preto je možné povedať, že hybridné smerovanie je kombináciou proaktívneho a reaktívneho smerovania. Hybridné smerovacie protokoly môžu obsahovať aj iné parametre, ktoré ovplyvňujú rozhodovanie pri smerovaní. V tejto práci je týmto parametrom blockchain a jeho stavebné prvky.

A. Smerovací protokol AODV

Protokol AODV je reaktívny MANET smerovací protokol. Pojem reaktívny protokol je taktiež možné zameniť za tzv. protokol na vyžiadanie (on-demand). Tento protokol bol implementovaný v praktickej časti tejto práce, preto bolo potrebné potrebné jeho objasnenie.

Protokol AODV (Ad Hoc On-Demand Distance Vector) kombinuje základné mechanizmy pre objavenie cesty (Route Discovery) a udržiavanie cesty (Route Maintenance) z protokolov DSR a DSDV. AODV vykonáva proces objavovania cesty použitím kontrolných správ typu Route Request (RREQ) a Route Reply (RREP). Ak zdrojový uzol S chce komunikovať s cieľovým uzlom D a nevie nájsť trasu vo svojej smerovacej tabuľke, zdrojový uzol vyšle RREQ správu všetkým susedným uzlom s posledným známym sekvenčným číslom destinácie. Ak žiaden zo susedných uzlov nemá vo svojich smerovacích tabuľkách vyhovujúcu trasu k cieľovému uzlu, všetky susedné uzly rovnako rozpošlú paket všetkým uzlom s ktorými susedia. Tento proces sa opakuje, až kým sa RREQ správa nedostane do cieľového uzla alebo kým niektorý z prostredných uzlov medzi zdrojom a cieľovým uzlom nemá vo svojej smerovacej tabuľke vhodnú cestu [1]. Na Obrázok 1 je možné vidieť zjednodušenú ukážku protokolu AODV, kde S je zdrojový uzol a D je cieľový uzol.



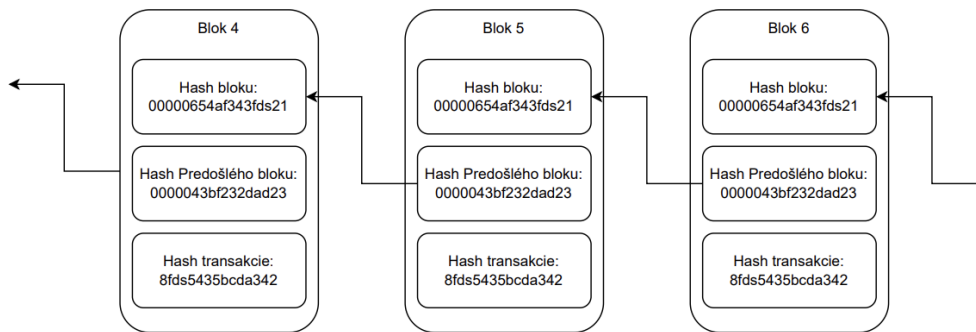
Obrázok 1 Smerovací protokol AODV [2]

III. TECHNOLÓGIA BLOCKCHAIN

Blockchain je zdieľaná, nemenná účtovná kniha (DL = Distributed Ledger), ktorá uľahčuje proces zaznamenávania transakcií a sledovania aktív v obchodnej sieti. Aktívum môže byť hmotné (dom, auto, hotovosť, pozemok) alebo nehmotné (informácie o uzle v sieti, duševné vlastníctvo, patenty, autorské práva, značka). V blockchainovej sieti je možné ukladať, sledovať a obchodovať prakticky čokoľvek, čo má hodnotu, čím sa znižuje riziko a náklady pre všetkých zúčastnených.

Prečo je teda blockchain dôležitý? Obchod funguje na informáciách. Čím rýchlejšie ich vie prijať a čím sú presnejšie, tým lepšie. Blockchain je ideálny na poskytovanie týchto informácií, pretože poskytuje okamžité, zdieľané a úplne transparentné informácie uložené v nemennej účtovnej knihe, ku ktorej majú prístup iba oprávnení členovia siete. Blockchain sieť môže sledovať objednávky, platby, účty, produkciu a oveľa viac. A keďže členovia zdieľajú jednotný pohľad na pravdu, je možné vidieť všetky podrobnosti transakcie od začiatku až do jej konca, čo im dáva väčšiu dôveryhodnosť, ako aj nové príležitosti a možnosti efektívnosti.

Iná definícia technológie blockchain hovorí, že blockchain je nový druh databázy. Táto technológia je veľmi zaujímavá pre ľudí, pretože dokáže riešiť problémy spojené s duplikátnymi záznamami dát. Blockchain vytvára bloky s rôznymi informáciami. Každý z týchto blokov súvisí so susednými blokmi v blockchaine. Existujúce bloky je takmer nemožné zmeniť alebo zmazať. Ak by útočník chcel napadnúť blockchain, potreboval by nato enormné množstvo výpočtového výkonu. Ťažiar (miners) sú ľudia, ktorí vypočítavajú hash hodnotu nového bloku, aby ho bolo možné pridať do blockchainu. Tento výpočet prebieha pomocou opakovaného hashovania rovnakého reťazca, pričom sa skúšajú meniť hodnoty nonce, až kým nedostaneme akceptovateľnú hodnotu hash reťazca [3]. Každá transakcia je zaznamenaná ako blok dát. Tieto transakcie ukazujú pohyb aktíva, ktoré môže byť hmotné (produkt) alebo nehmotné (intelektuálne). Dátový blok môže zaznamenať rôzne informácie, napr. kto, kedy, kde, ako a za akých okolností vykonal činnosť – v našom prípade informácie o jednotlivých uzloch siete. Každý blok nadväzuje na predošlý blok tým, že si uchováva jeho hash hodnotu. Na Obrázok 2 je možné vidieť schému technológie blockchain.



Obrázok 2 Schéma blockchainu

Každý blok pozostáva z hash hodnoty predošlého bloku, samotných dát (transakcie), hash hodnoty samotných dát, čísla nonce a časovej známky (timestamp).

Nonce je číslo generované pri vytvorení nového bloku. Toto číslo musia ostatné uzly v blockchaine uhádnuť a použiť ho vo výpočte novej hash hodnoty za účelom pridania bloku do blockchainu, čím sa zabezpečí verifikácia, integrita a nemennosť daného bloku. Iba ak je na začiatku hash hodnoty nového bloku určitý počet núl, iba vtedy môže byť blok pridaný do blockchainu. To znamená, že číslo Nonce bolo správne uhádnuté. V opačnom prípade sa musí vybrať iné číslo nonce, pomocou ktorého sa blok zahashuje.

Každý blok je pripojený k ostatným blokom pred ním a za ním. Tieto bloky tvoria Reťazec (Chain) dát počas pohybu aktíva z jedného miesta na druhé alebo sa mení vlastníctvo. Bloky potvrdzujú presný čas a postupnosť transakcií. Aby sa zabránilo narušeniu blokov alebo vloženiu nového bloku medzi dva existujúce bloky v reťazci, bloky sú bezpečne spájané dokopy so zachovaním integrity. Transakcie sú teda spájané do nenavrátiteľného reťazca – blockchainu. Každý pridaný blok zosilňuje verifikáciu predošlého bloku a teda celého blockchainu. To robí blockchain odolným voči neoprávneným zásahom, čo prináša kľúčovú silu nemennosti. Tým sa odstráni možnosť manipulácie zo strany škodlivého aktéra a vytvorí sa kniha transakcií (ledger), ktorým môžete vy a ostatní členovia siete dôverovať.

Blockchainové uzly sú zainteresované strany (stakeholders) siete a ich zariadenia, ktoré sú oprávnené sledovať distribuovanú účtovnú knihu a, ktoré slúžia ako komunikačné uzly pre rôzne sieťové úlohy. Hlavnou úlohou uzla je potvrdiť zákonnosť každej nasledujúcej dávky sieťových transakcií (blokov). Okrem toho, každý uzol má pridelený jedinečný identifikátor, ktorý zabezpečuje ľahko rozlíšiť jednotlivé uzly v sieti. V koncepcii Proof-of-Work blockchainu, akým je napr. Bitcoin, sú súčasťou siete aj ťažiče – minery. Iba od úplných uzlov (full nodes) je vyžadované ukladať všetky blockchainové transakcie na ich zariadeniach. Tieto uzly majú na starosti overovanie blokov a transakcií. Na druhej strane, ľahkovážne (lightweight) uzly majú menšie požiadavky na úložisko, pretože potrebujú stiahnuť iba hlavičku bloku na overenie transakcie. Odmena za spracovanie bloku nemusí byť istotou v ani jednej verzii uzla. Keď miner pridáva nový blok transakcie do blockchainu, blok je vysielaný ako broadcast všetkým uzlom. Na základe legitimity bloku uzly môžu prijať alebo odmietnuť jeho pridanie. Keď uzol prijme nový blok, uloží si ho a umiestni ho na vrchol/koniec existujúcich blokov. V blockchaine protokoly zachovávajú decentralizovaný prístup, eliminujú potrebu centrálny authority zdieľaním údajov v celej sieti. Protokoly zaisťujú, že dáta prenášané cez sieť sú efektívne, zabezpečené a spoľahlivé. Zodpovednosťou blockchain protokolu je udržiavať a riadiť bezpečnosť, konsenzus a sieťové prvky blockchainu. Nasledujúce pojmy sú kľúčové pre porozumenie Blockchain protokolov.

Konsenzus je jeden z najdôležitejších termínov v oblasti blockchainu. Určuje kto je zodpovedný za overovanie (validating) záznamov blokov a ako ich môžu ostatní členovia autentifikovať. Algortimus zaisťuje štandardnú dohodu medzi účastníkmi ohľadom stavu distribuovanej peňaženky (ledgera).

Technológia Distribovanej databázy (Distributed Ledger Technology - DLT) je infraštruktúra, ktorá umožňuje účastníkom komunikovať skrz sieť na dosiahnutie konsenzu. Distribuovaná peňaženka odpovedá, zdieľa a synchronizuje dáta digitálne a zdieľa ich skrz sieť na rôznych miestach.

Smart kontrakty (Smart Contracts) sú dohody medzi dvomi účastníkmi uloženými na blockchaine. Je to kus kódu, ktorý je vykonaný automaticky, keď sú splnené predurčené

podmienky. Smart kontrakt je uložený vo verejnej databáze a teda nemôže byť presunutý alebo zmenený [4].

IV. VYUŽITIE TECHNOLOGIE BLOCKCHAIN V PROTOKOLE AODV

Logika implementovaná do zdrojového kódu bola založená na existujúcej implementácii blockchainu do smerovacieho protokolu AODV [5]. Implementáciu blockchainu do smerovacieho protokolu AODV je možné opísať nasledovne:

1. pri každom prijatom pakete je generovaný nový blok dát,
2. blok dát pozostáva z informácií o pakete a trase opísaných v Tabuľka 1,
3. celý blockchain sa postupne zapisuje do súboru blockchain.csv,
4. uzol prijme paket len v prípade, ak obsahuje vhodnú hash hodnotu bloku.

Tabuľka 1 Parametre implementovaného blockchainu

Parameter	Hodnota
Hashovacia Funkcia	SHA256
Veľkosť bloku [B]	64
Hashované dáta	Zdrojová IP adresa (IPV4)
	Cieľová IP adresa (IPV4)
	ID odosielacieho uzla
	predošlá Hash hodnota v hexadecimálnom tvare
	Trasa = Zoznam ID uzlov
	veľkosť trasy = počet preskokov
Nová hodnota hash	64 hexadecimálnych znakov (256 bitov)

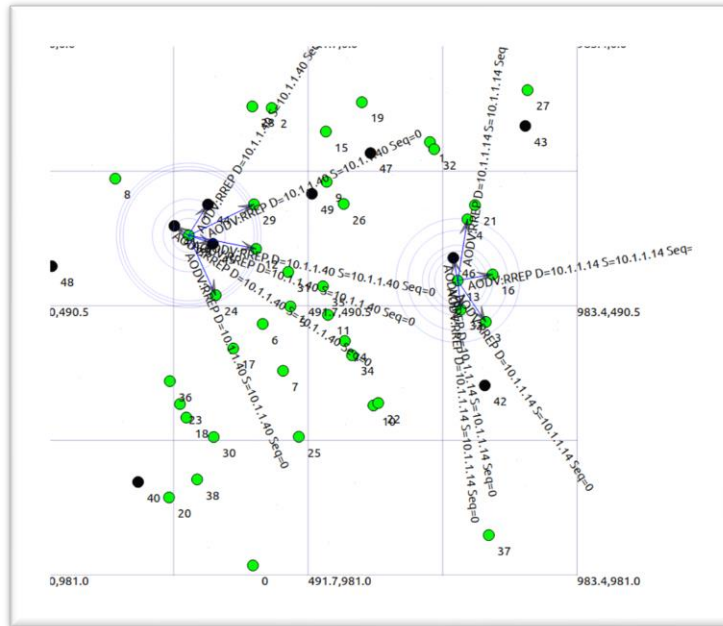
Pre praktickú časť bol zvolený protokol AODV, ktorý bol obohatený o technológiu blockchain. Tento náš novovytvorený hybridný protokol bol následne analyzovaný samostatne a zároveň porovnávaný s tradičnou implementáciou AODV za rovnakých simulačných podmienok. Pre každý simulačný scenár, v ktorom sme analyzovali vplyv škodlivých uzlov, bolo vykonaných 60 simulácií. Zároveň sme pozorovali správanie vybraných protokolov v závislosti na rýchlosti pohybu uzlov. Tu sme nastavili tri intervaly, a to: nula až dva, dva až štyri a štyri až šesť metrov za sekundu. Všetky ostatné parametre je možné vidieť v Tabuľka 2. V tomto prípade bolo vykonaných až 60 simulácií pre každý simulačný scenár. Z výsledkov každej jednej simulácie boli vytvorené záznamy do .csv súboru, z ktorého sa následovne vytvorili priemerné hodnoty, ktoré boli analyzované.

Tabuľka 2 Parametre simulácii

Premenná	Hodnota
Veľkosť simulačnej plochy [m x m]	1000x1000
Počet MANET uzlov	50
Počet Škodlivých uzlov	10%, 20%,30%,40%,50%,60%
Rýchlosť uzlov [m/s]	0-2; 2-4; 4-6
Pohybový model	Model náhodného pohybu
Čas simulácie [s]	90
Vysielací výkon uzlov [db/mW]	7.5
Počet Simulácii pre každý scenár	60
Propagačný model	Voľný priestor
Technológia vysielania	IEEE802.11

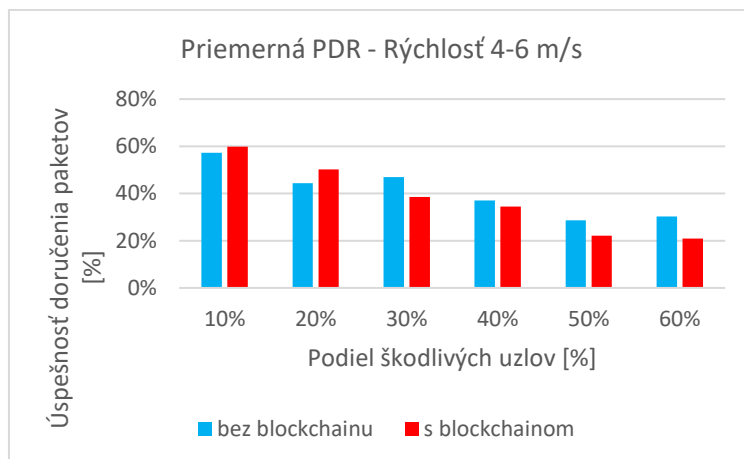
Simulácie boli zamerané na pozorovanie a potencionálne zlepšenie odolnosti protokolov pomocou technológie blockchain. Preto boli vytvorené škodlivé uzly, ktoré mali za úlohu negatívne ovplyvňovať sieť. Pridanie škodlivých uzlov do simulácie je možné vidieť na nasledujúcom obrázku. Tieto uzly boli pomenované ako blackholes, čiže v preklade čierne diery. Blackhole uzol síce vie prijať paket, ale ďalej ho nepreposiela a zahodí ho, čím znižuje celkovú kvalitu prenosu, zvyšuje oneskorenie a negatívne vplýva na ostatné parametre siete. Vo vytvorených simuláciách bol počet týchto uzlov vyjadrený v percentuálnom počte voči celkovému počtu uzlov v pomere 10-60%. To znamená, že napríklad pri 10% zaťažení siete škodlivými uzlami sa počet uzlov rovnal piatim, ak berieme do úvahy, že celkový počet všetkých

uzlov bol päťdesiat. Škodlivivé uzly sú označené čiernou farbou a je možné ich vidieť na Obrázok 3.



Obrázok 3 Vizualizácia simulácie so škodlivými uzlami v simulátore NS-3

Najdôležitejším ukazovateľom pri QoS parametroch je pre nás PDR = Packet Delivery Ratio, čiže úspešnosť doručenia paketov. So zvyšovaním podielu škodlivých uzlov voči zdravým uzlom v sieti sme pozorovali klesajúcu úspešnosť doručenia paketov. Zistili sme, že implementovaný blockchain v protokole AODV pri rýchlosti 4 až 6 metrov za sekundu zvyšuje úspešnosť doručenia paketov len pri 10 a 20 percentnom zaťažení. Zlepšenie predstavovalo zvýšenie úspešnosti doručenia paketov o tri až šesť percent. Vo zvyšných prípadoch bol prejav opačný a pozorovali sme pokles o päť až osem percent, čo je možné vidieť na Obrázok 4.



Obrázok 4 Priemerná percentuálna úspešnosť doručenia paketov v závislosti na počet škodlivých uzlov pri rýchlosti 4-6 m/s

V Tabuľka 3 je možné vidieť výsledky celkového porovnania priemerného E2E oneskorenia pri všetkých možných simulačných scenároch, pričom každý simulačný scenár prebehol práve 60 krát. Skratka BL znamená, že simulácia používala smerovací protokol AODV s technológiou blockchain. Stĺpce, ktoré neobsahujú skratku BL, reprezentujú klasický smerovací protokol AODV.

Tabuľka 3 End-to-End oneskorenie v závislosti od rýchlosti pohybu uzlov a počtu škodlivých uzlov

Oneskorenie [s]		Rýchlosť pohybu uzlov [m/s]					
		0-2	BL 0-2	2-4	BL 2-4	4-6	BL 4-6
Podiel škodlivých uzlov [%]	10%	5.19	3.66	17.00	12.01	27.05	33.67
	20%	2.86	4.03	8.38	8.59	17.44	25.61
	30%	1.41	0.56	11.44	5.55	15.17	13.28
	40%	0.68	2.44	2.67	8.32	9.18	22.68
	50%	4.58	0.31	9.91	5.30	16.16	9.02
	60%	1.21	0.17	10.53	8.07	11.63	17.68

V. ZÁVER

Vytvorili sme potrebné simulačné scenáre zložené z päťdesiatich uzlov, v ktorých sme menili percentuálny podiel škodlivých uzlov a rýchlosť pohybu uzlov. Z dosiahnutých výsledkov sme pozorovali nasledovné fakty. Postupným zvyšovaním percentuálneho podielu škodlivých uzlov klesala priemerná percentuálna úspešnosť doručenia paketov. Pri využití blockchainu sme pozorovali zlepšenie percentuálnej úspešnosti paketov o päť až šesť percent pri rýchlosti uzlov nula až štyri metre za sekundu. Zlepšenie priemernej úspešnosti paketov o tri až šesť percent bolo pozorované pri desať, dvadsať, štyridsať a päťdesiat percentnom zaťažení siete škodlivými uzlami. Pre zvyšné percentuálne zaťaženia už blockchain neprinášal benefit a jeho percentuálna úspešnosť doručenia paketov klesala o päť až osem percent v porovnaní s protokolom AODV, ktorý nevyužíva blockchain. Úspešnosť doručenia paketov pri rýchlosti štyri až šesť metrov za sekundu bola pri využití blockchainu nižšia o tri percentá v porovnaní s protokolom bez blockchainu. V nižších rýchlostiach implementovaný protokol s blockchainom zvyšoval úspešnosť doručenia paketov o päť až šesť percent pri porovnaní s klasickým AODV protokolom.

V budúcnosti môžeme vidieť nárast implementácii blockchainu práve v oblasti mobilných ad-hoc sietí, kde blockchain môže priniesť nový pohľad na MANET siete a ich udržateľnosť z pohľadu bezpečnosti a odolnosti smerovania.

POĎAKOVANIE

Tento výskum bol financovaný agetúrou na podporu výskumu a vývoja SR, výskumný grant č. APVV-17-0208 a VEGA 1/0260/23.

LITERATÚRA

- [1] G. O FANG, X. - YANG, D. - GUNDECHA, P. - XUE, G. 2010. Multi-constrained anypath routing in wireless mesh networks [online]. [2023-04-06]. In 7th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks. Dostupné na: <https://ieeexplore.ieee.org/document/5508266>
- [2] KATTO, J.- SAKURAI, Y. : AODV multipath extension using source route lists with optimized route establishment [online]. 2021. [2023-04-11]. Dostupné na internete: https://www.researchgate.net/publication/4183879_AODV_multipath_extension_using_source_route_lists_with_optimized_route_establishment
- [3] GOLOSOVA, J. – ROMANOV, A. : The Advantages and Disadvantages of the Blockchain Technology [online]. Vilnius, Litva. 2018. [2023-04-12]. Dostupné na internete: <https://ieeexplore.ieee.org/document/8592253> . ISBN 978-1-7281-1999-1
- [4] 3801-2022 : 2022, Standard for Blockchain-based Electronic Contracts [online]. IEEE. [2023-04-12]. Dostupné na internete: <https://ieeexplore.ieee.org/document/9745868/>
- [5] JARJIS, A. – GOVAND, K. : Blockchain Authentication for AODV Routing Protocol [online]. Erbil, Irak. 2020. [2023-04-13]. Dostupné na internete: <https://ieeexplore.ieee.org/document/9274452> . ISBN: 978-1-7281-8370-1.

Smerovanie vo viacpreskokovej D2D sieti a Blockchain

¹Maroš BAUMGARTNER, ²Ján PAPAJ, ³Peter SÁROSI

^{1,2,3}Katedra Elektroniky a Multimediálnych Telekomunikácií,
Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach,
Slovenská Republika

¹maros.baumgartner@tuke.sk, ²jan.papaj@tuke.sk, ³peter.sarosi@student.tuke.sk

Abstract — Článok predstavuje úvodnú štúdiu využitia technológie Blockchain pre viacpreskokové D2D siete. Cieľom článku je poskytnúť základný popis využívaných mobilných sietí, popis komunikácie medzi zariadeniami D2D a jej viacpreskokovým variantom MD2D, opis technológie Blockchain a jej význam v smerovaní a taktiež možnosti využitia technológie Blockchain s dôrazom kladeným na zvýšenie sieťovej bezpečnosti a robustnosti v D2D a viacpreskokových D2D sieťach.

Kľúčové slová — smerovanie, blockchain, D2D, MD2D, bezpečnosť

I. ÚVOD

S narastajúcim počtom inteligentných zariadení v dnešnej dobe, ktoré sú pripojené do mobilnej siete, spôsobil, že súčasné mobilné siete sa začali zahlcovať a strácať svoje funkcie, resp. sa začali spomaľovať, pretože neboli navrhnuté pre tak masívny počet zariadení. Jedným z riešení tohoto problému je mobilná komunikácia medzi zariadeniami (D2D), ktorá umožňuje zariadeniam priamu komunikáciu medzi sebou bez potreby spoliehania sa na tradičnú sieťovú infraštruktúru. Komunikácia medzi zariadeniami je technika priamej komunikácie v riadenom prostredí, ktorá vedie k lepšiemu výkonu siete. Medzi výhody komunikácie medzi zariadeniami patrí zvýšená spektrálna účinnosť, zvýšený výkon a energetická účinnosť, zvýšená kapacita, zabezpečenie kvality služieb, zníženie rušenia, lepšia spoľahlivosť, lepšie pokrytie a nákladová efektívnosť. Avšak používanie technológie komunikácie medzi zariadeniami v celulárnych sieťach spôsobuje rušenie, pretože rovnaké zdroje zdieľajú dva typy používateľov. Preto musíme znížiť rušenie medzi zariadeniami v mobilnej sieti. Keď sú mobilné zdroje v downlink spojení zdieľané medzi komunikáciou medzi zariadeniami a používateľmi mobilných telefónov, aj prijímač komunikácie medzi zariadeniami toleruje rušenie spôsobené základňovou stanicou. Na druhej strane, keď sú zdroje mobilnej siete zdieľané v uplinku, potom základňová stanica ruší komunikáciu medzi zariadeniami. Technológia Blockchain predstavuje jednu z najvýznamnejších zmien v technologickom svete. Technológia, ktorá začala ako základná infraštruktúra Bitcoinu, sa teraz používa v rôznych aplikáciách, ktoré fungujú na jej základoch a siahajú ďaleko za hranice kryptomien. Základom tejto technológie je, že Blockchain môže ponúknuť dôveru v sieťovom uzle a zároveň výrazne zlepšiť bezpečnosť siete. Ide o distribuovanú a decentralizovanú účtovnú knihu, ktorá ponúka transparentnosť, bezpečnosť a dôveru pri zaznamenávaní transakcií a správe údajov. Koncept Blockchainu bol predstavený v roku 2008 vytvorením prvej kryptomeny založenej na technológii Blockchain, Bitcoin.

II. MOBILNÁ KOMUNIKÁCIA

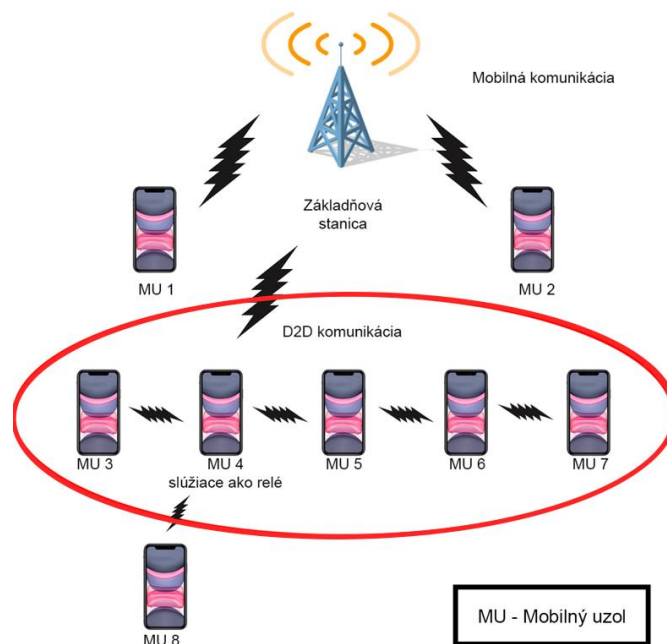
A. Piata generácia mobilných sietí

5G sieť je piata generácia technológie mobilnej siete, ktorá sľubuje vyššiu rýchlosť internetu, nižšie oneskorenie a vyššiu kapacitu v porovnaní s jej predchodcom 4G. Jedným z najvýznamnejších vylepšení, ktoré 5G sieť prináša, je jej rýchlosť. Zatiaľ čo siete 4G zvyčajne ponúkajú rýchlosť sťahovania až 100 Mb/s, siete 5G môžu dosiahnuť až 10 Gb/s. To znamená, že sťahovanie veľkých súborov, streamovanie videí vo vysokom rozlíšení a iné dátovo náročné úlohy sú rýchlejšie spracované. Ďalšou dôležitou výhodou 5G siete je jej nižšia odozva. Odozva označuje čas, ktorý trvá prenosu údajov z jedného zariadenia do druhého. S 5G sieťou sa očakáva, že oneskorenie klesne až na jednu milisekundu, čo umožní bezproblémový chod aplikácií v

reálnom čase, ako sú online hry a virtuálna realita. 5G sieť tiež sľubuje podporu viacerých pripojených zariadení súčasne, a to vďaka zvýšenej kapacite. To je dôležité najmä pre Internet vecí (IoT), kde sú zariadenia čoraz viac prepojené a potrebujú medzi sebou bezproblémovo komunikovať. Zavádzanie sietí 5G však stále prebieha a v súčasnosti je dostupné v niektorých krajinách a mestách. Okrem toho si 5G sieť vyžaduje novú infraštruktúru vrátane nových veží a antén, ktorých implementácia môže byť nákladná. Preto môže chvíľu trvať, kým sa táto sieť stane široko dostupnou a prístupnou pre každého. Sieť piatej generácie môže pracovať na frekvenciách do 100 GHz, pre Európu je k dispozícii frekvenčné pásmo medzi 24,2 a 27,5 GHz. Tieto pásma umožňujú použitie jednotlivých kanálov až do veľkosti 400 MHz, čo znamená, že teoretická rýchlosť je 20 Gbps. Skutočná prenosová rýchlosť sa však bude pohybovať medzi 100 Mbps a 1 Gbps, v závislosti predovšetkým od veľkosti oblasti, hustoty zariadení a oblasti pokrytia prístupovým bodom. Latencia sa zníži na 1 milisekundu zvýšením šírky kanála a prenosovej rýchlosti. Je možné podporovať až 10^6 zariadení na km^2 a zároveň zlepšiť mobilitu a stabilitu pripojenia až do rýchlosti 500 km/h. Celkové zlepšenia spektrálnej účinnosti by mali byť až päťkrát väčšie ako pri súčasných sieťach [1].

B. Priama komunikácia medzi zariadeniami

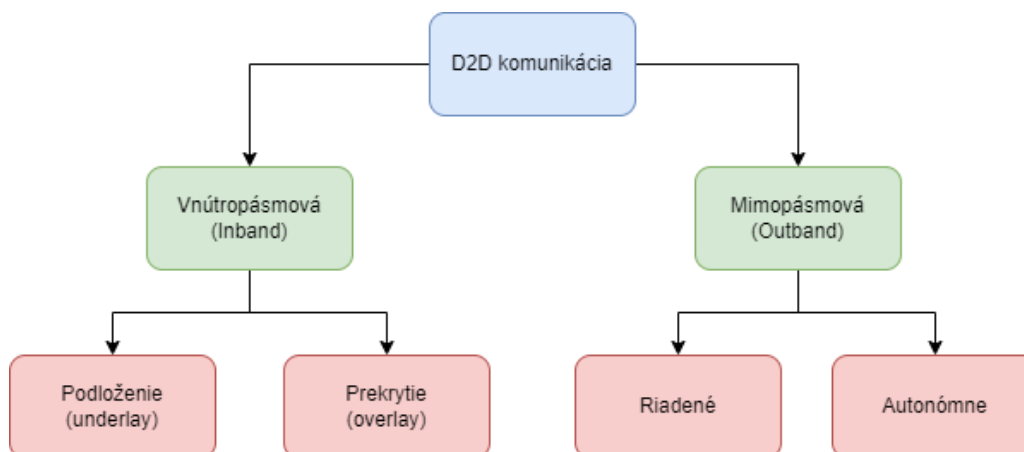
Nová technológia prenosu dát v sieťach piatej generácie je priama komunikácia medzi zariadeniami (D2D). Táto technológia si v sieťach štvrtej generácie nenašla uplatnenie. Jej použitie v nových sieťach piatej generácie však vylepšuje mobilné služby a zlepšuje kvalitu služieb. Device-to-Device komunikácia umožňuje zariadeniam komunikovať medzi sebou priamo bez prechodu komunikácie cez základňovú stanicu alebo prístupové body. Je takto možné prenášať dáta väčšou rýchlosťou a v kratšom čase. Prínosom D2D komunikácie je aj zvýšenie spektrálnej účinnosti základňovej stanice, cez ktorú komunikácia môže, ale nemusí smerovať. D2D môžeme rozdeliť na priamu D2D komunikáciu, komunikáciu asistovanú základňovou stanicou a na komunikáciu pomocou smerovačov.



Obrázok 1 Mobilná a D2D komunikácia

D2D môže využívať licencované spektrum (vnútro pásmove) – D2D Inband, alebo nelicencované spektrum (mimo pásma) – D2D Outband, na vytvorenie priameho spojenia [2]. Inband D2D komunikácia je kategorizovaná ako prekrytie (overlay) a podloženie (underlay). Podkladová D2D komunikácia umožňuje nastavenie priamych a bunkových spojení v bunkovom spektre. Na druhej strane v prekryvanom D2D sa vyhradená časť dostupného spektra používa na komunikáciu medzi zariadeniami, pričom zvyšok spektra sa používa na mobilnú komunikáciu. Keďže mimopásmová D2D komunikácia využíva nelicencované spektrum na vytváranie priamych spojení, je kategorizovaná ako autonómna a riadená. Pri ovládaní sú rádiové rozhrania v D2D riadené uNB, zatiaľ čo v autonómnom režime sú koordinované samotnými používateľskými zariadeniami. Rušenie medzi používateľmi D2D a mobilnými používateľmi v mimopásmovej D2D nie je problémom, zatiaľ čo koordinácia komunikácie v nelicencovanom pásme vyžaduje druhé rádiové rozhranie (napríklad Wi-Fi Direct, Bluetooth,

ZigBee) [3], [4]. Kategorizácia D2D komunikácie je znázornená na obr. č. 2. Aby sme čo najefektívnejšie využili obmedzené dostupné spektrum, musíme vedieť, kde a akú kategóriu D2D komunikácie použiť.



Obrázok 2 Typy D2D komunikácie

C. Technológia Blockchain

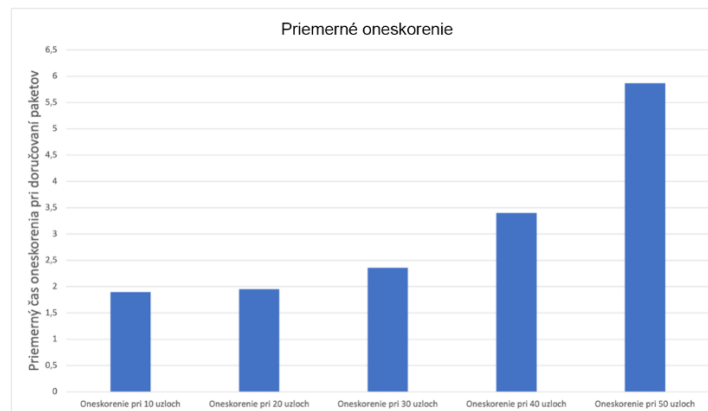
Technológia Blockchain je distribuovaná databáza, ktorá je zdieľaná medzi uzlami počítačovej siete. Blockchain ako databáza uchováva informácie elektronicky v digitálnom formáte. Blockchain je známy pre svoju kľúčovú úlohu v kryptomenových systémoch, ako je Bitcoin, pre udržiavanie bezpečného a decentralizovaného záznamu transakcií. Inováciou Blockchainu je, že zaručuje vernosť a bezpečnosť záznamu údajov a vytvára dôveru bez potreby dôveryhodnej tretej strany. Jedným z kľúčových rozdielov medzi typickou databázou a Blockchainom je štruktúra údajov. Blockchain zhromažďuje informácie spoločne v skupinách, známych ako bloky, ktoré obsahujú súbory informácií [5]. Bloky majú určité úložné kapacity a po naplnení sú zatvorené a prepojené s predtým vyplneným blokom, čím sa vytvorí reťazec údajov známy ako Blockchain. Všetky nové informácie, ktoré nasledujú po tom, čo sa vytvoril nový čerstvo pridaný blok sa uložia do novovytvoreného bloku, ktorý sa po naplnení tiež pridá do reťazca. Databáza zvyčajne štrukturuje svoje údaje do tabuliek, zatiaľ čo Blockchain, ako naznačuje jeho názov, štrukturuje svoje údaje do blokov, ktoré sú spojené dohromady. Táto štruktúra údajov vo svojej podstate vytvára nezvratnú časovú os údajov, pokiaľ sa implementuje v decentralizovanom charaktere. Keď je blok naplnený, následne zasadne do reťazca a stáva sa súčasťou tejto časovej osi. Každý blok v reťazci dostane pri pridávaní do reťazca presnú časovú pečiatku [6]. Štruktúra technológie Blockchain sa skladá z dát, hash hodnoty, časovej pečiatky, hodnoty nonce a stromu merkle. Každý blok v rámci Blockchainu je identifikovaný hashom vygenerovaným pomocou kryptografického hashového algoritmu SHA256 v hlavičke bloku. Každý blok tiež odkazuje na predchádzajúci blok, známy ako rodičovský blok, prostredníctvom poľa „hash predchádzajúceho bloku“ v hlavičke bloku. Inými slovami, každý blok obsahuje hash svojho rodiča vo svojej vlastnej hlavičke.

III. NÁVRH SIMULÁCIE

Tabuľka 1
 Simulačné parametre

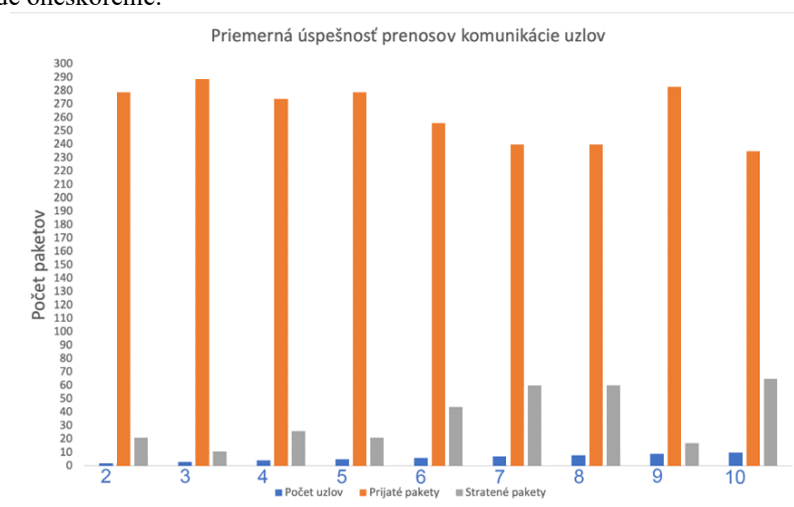
Premenná	Hodnota
Veľkosť oblasti	1000x1000 metrov
Počet uzlov	10, 20, 30, 40, 50
Šírka pásma	20 MHz
Pomer signálu k šumu (SNR)	25 dB
Rádiový dosah	50 metrov
Veľkosť paketov	300 kb
Prostredie	Otvorený priestor bez prekážok
Komunikačný štandard	802.11

Návrh simulácie viacpreskokovej D2D komunikácie je realizovaný v programovacom jazyku Python. Pre simuláciu bolo potrebné navrhnuť možnosti zvýšenia robustnosti komunikácie, a preto cieľom bolo zlepšiť prenosovú rýchlosť, znížiť oneskorenie paketov. Najskôr sa stanoví rozmer oblasti, kde sa budú jednotlivé zariadenia nachádzať, na 1000x1000 metrov. V oblasti sa následne náhodne vygeneruje pozícia uzlov, ktorých počet je vopred zadaný. Po nájdení trasy sa medzi zdrojovým a cieľovým uzlom začnú posielat' pakety o veľkosti 300 kb. Merá sa čas potrebný k prenosu, prenosová rýchlosť, strata paketov, počet skokov. Navrhnutý model komunikácie bol testovaný prostredníctvom parametrov zobrazených v tabuľke 1. Celkový počet simulácií bol 500. Experimentálne výsledky simulácii sme rozdelili na celkový počet doručených paketov, priemerné oneskorenie paketov, priemernú priepustnosť uzlov a na priemerný čas doručenia paketov cez viac skokov.



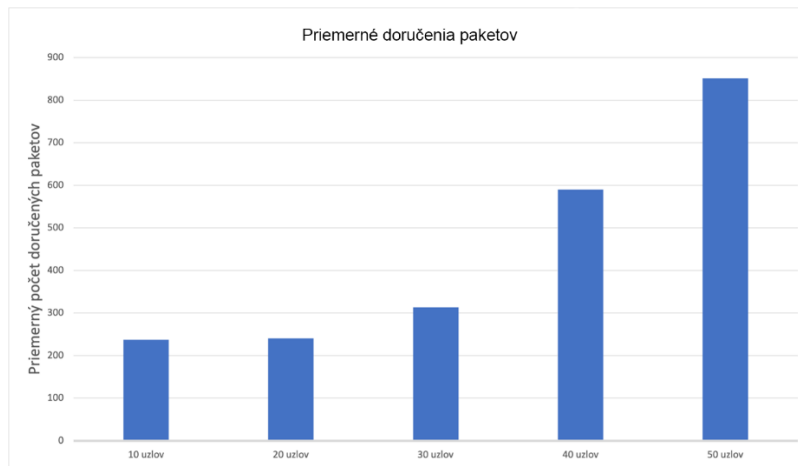
Obrázok 3 Experimentálne výsledky oneskorenia

Na obrázku 3 je znázornené priemerné oneskorenie komunikácie uzlov. Pri 10 uzloch bolo oneskorenie najmenšie, keďže bola pri simuláciách malá šanca, že pakety budú pri smerovaní stratené či poškodené. Nárastom počtu uzlov ale narastalo aj oneskorenie. Najčastejšou príčinou tohto oneskorenia bola strata paketov, kde ak uzol daný paket stratil či nejakým spôsobom bol paket poškodený, musel uzol paket opätovne preposielat'. Najväčšie oneskorenie bolo namerané pri simulácii s 50 uzlami. Na trasách, ktoré boli nájdené, sa vyskytovalo viac uzlov, ktoré preposielali pakety, teda sa podieľali na smerovaní. Týmto nárastom počtu uzlov, ktoré sa podieľali na komunikácii nám vznikla MD2D sieť. Teda čím viac skokov musí paket prejsť, tým väčšie bude oneskorenie.



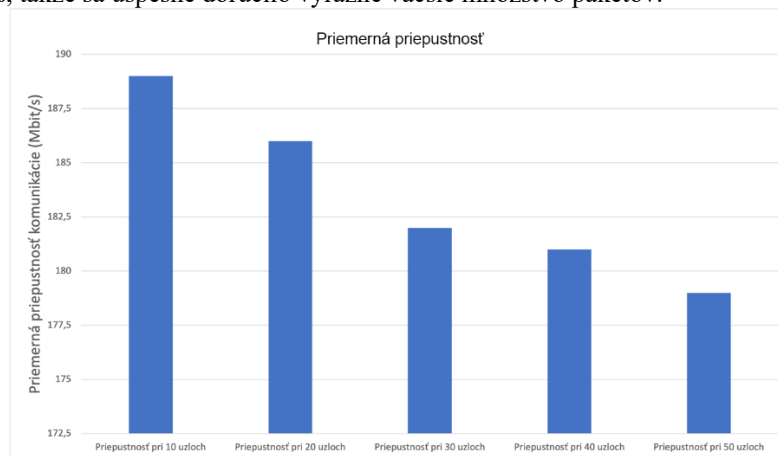
Obrázok 4 Experimentálne výsledky oneskorenia úspešnosti prenosu paketov

Môžeme vidieť, že najefektívnejší variant MD2D je cesta s tromi uzlami, teda dvoma skokmi. Stratovosť paketov pri troch uzloch predstavuje hodnotu 4%, teda úspešnosť prenosu dvoma skokmi bola 96%. Naopak, varianty so siedmimi a ôsmimi uzlami, teda 7 a 8 skokov, sa ukázali byť najmenej úspešné, a to na hranici 80% úspešne prijatých paketov, teda straty paketov predstavovali hodnotu 20%. Druhé najúspešnejšie riešenie v MD2D bolo riešenie s deviatimi uzlami, teda ôsmimi skokmi, kde hodnota prijatých paketov predstavovala 94%.



Obrázok 5 Experimentálne výsledky priemerného počtu doručených paketov

Na obrázku 5 sú znázornené priemerné doručenia paketov pri rôznych počtoch uzlov v sieti. Nízky počet doručených paketov pre 10, 20 a 30 uzlov bol spôsobený tým, že spomedzi všetkých simulácií sa málo krát našla trasa pre komunikáciu. V priemere bola chybovosť nájdenia trasy veľmi vysoká, teda cez 70%. Taktiež vidíme, že počet doručených paketov sa začal zvyšovať pri komunikácii pomocou 40 uzlov a najviac paketov sa doručilo pri komunikácii použitím 50 uzlov. Opäť je to spôsobené tým, že pri 50 uzloch sa percentuálna úspešnosť nájdenia trasy pohybovala okolo 60%, takže sa úspešne doručilo výrazne väčšie množstvo paketov.



Obrázok 6 Experimentálne výsledky priepustnosti

Priemerná priepustnosť bola meraná v Mbit/s. Ako vidíme na obrázku 6, simulácie s 10 uzlami sa ukázali ako najviac priepustné. Je to zapríčinené už spomínaným nízkym počtom uzlov, a teda malým počtom viacskokových trás. Vo všeobecnosti môže mať sieť s menším počtom uzlov vyššiu potenciálnu priepustnosť ako sieť s viacerými uzlami, pretože je menšia šanca na rušenie a preťaženie. Simulácie s 50 uzlami sa ukázali ako najmenej priepustné, opäť je to spôsobené trasami s viacero skokmi. Sieť s 50 uzlami dosiahla priemernú priepustnosť 179 Mbit/s.

IV. ZÁVER

D2D komunikácia je sľubnou technológiou v nových 5G sieťach. Výhody, ktoré so sebou prináša zlepšujú efektivitu mobilných služieb. Na základe rozdelenia ju je možné použiť nielen v súčasných mobilných sieťach, ale aj mimo nich, čo prispeje k väčšiemu dosahu signálu. Využitie D2D komunikácie v rôznej sfére spraví siete pre každého dostupnejšie. S využitím technológie Blockchain je možné zlepšiť robustnosť siete. Teória spísaná v tomto príspevku oboznamuje o moderných mobilných sieťach a o technológii Blockchain. S touto technológiou je potrebné pochopiť aj princíp k porozumeniu využitia v návrhu modelu viacpreskokovej D2D komunikácie. Podľa tvrdení overených v simuláciách bolo zistené, že je možné zlepšenie bežnej komunikácie cez technológiu D2D.

POĎAKOVANIE

Tento výskum bol financovaný agetúrou na podporu výskumu a vývoja SR, výskumný grant č. APVV-17-0208 a VEGA 1/0260/23.

POUŽITÁ LITERATÚRA

- [1] Yu Tang, Sathian Dananjayan, Chaojun Hou, Qiwei Guo, Shaoming Luo, Yong He, A survey on the 5G network and its impact on agriculture: Challenges and opportunities, *Computers and Electronics in Agriculture*, Volume 180, 2021
- [2] A. Asadi, Q. Wang and V. Mancuso, "A Survey on Device-to-Device Communication in Cellular Networks," in *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 1801-1819, Fourthquarter 2014, doi: 10.1109/COMST.2014.2319555.
- [3] W. Lee, J. Kim and S. -W. Choi, "New D2D Peer Discovery Scheme Based on Spatial Correlation of Wireless Channel," in *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10120-10125, Dec. 2016, doi: 10.1109/TVT.2016.2530139.
- [4] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, S. Li and G. Feng, "Device-to-device communications in cellular networks," in *IEEE Communications Magazine*, vol. 52, no. 4, pp. 49-55, April 2014, doi: 10.1109/MCOM.2014.6807946.
- [5] M. Di Pierro, "What Is the Blockchain?," in *Computing in Science & Engineering*, vol. 19, no. 5, pp. 92-95, 2017, doi: 10.1109/MCSE.2017.3421554.
- [6] P. R. Nair and D. R. Dorai, "Evaluation of Performance and Security of Proof of Work and Proof of Stake using Blockchain," *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, Tirunelveli, India, 2021, pp. 279-283, doi: 10.1109/ICICV50876.2021.9388487.

Strojové učenie v rozpoznávaní rôznych foriem antisociálneho chovania

¹Kristína MACHOVÁ, ²Alexandra PEJKOVÁ

^{1,2} Katedra Kybernetiky and Umelej Inteligencie, Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská Republika

¹kristina.machova@tuke.sk, ²alexandra.pejkova@student.tuke.sk

Abstrakt — Článok sa skúma možnosti použitia metód strojového učenia v rozpoznávaní najrozšírenejších foriem antisociálneho chovania v online priestore. Anonymita a neosobnosť online priestoru rapídne uľahčila možnosť pôsobenia aktérov antisociálneho správania v dobe zmietaanej pandémie ako aj v čase jej odznievania. Je viac ako zrejmé, že manipulácia s informáciami tykajúcimi sa pandémie sa stali pre ľudí veľkým problémom. Príspevok popisuje tréning ako aj hodnotenie modelov na detekciu spamových príspevkov, podozrivých komentárov, dezinformácií a falošných správ. Na tréning modelov boli použité metódy ako Naive Bayes klasifikátor, SVM (support Vector Machine) a RF (Random Forest).

Kľúčové slová — detekcia antisociálneho chovania, spracovanie textových dokumentov, dolovanie webu, strojové učenie, Naive Bayes, Random Forest, SVM

I. ÚVOD

Rozmach antisociálneho správania sa so zväčšovaním množstva používateľov sociálnych sietí stáva ťažko riešiteľným doteraz používanými klasickými metódami. Giganti ako je napríklad Meta sa snažia používateľov svojich platforiem od takéhoto správania uchrániť používaním rôznych algoritmov a nástrojov, tento proces je žiaľ zatiaľ iba v zárodkoch a jeho výsledky ako aj detekčné modely sú nedostačujúce, zvyčajne charakteristické znepokojujúcim počtom nesprávne vyhodnotených správ.

Cieľom tohto príspevku je ponúknuť jedno z možných riešení vo forme implementácie nástrojov a modelov strojového učenia na vyhodnotenie a analýzu textov, so zameraním na rôzne druhy toxického správania. Prezentovaná práca teoreticky podáva stručný pohľad na problematiku antisociálneho správania v online priestore, poukazuje na rôzne druhy takéhoto správania a taktiež na ich pôvod. Ďalej popisuje tréning modelov strojového učenia, ktoré sú vhodné na detekciu tohto správania. Článok obsahuje aj výsledky testovania týchto modelov zameraných na odhaľovanie niekoľkých druhov toxického správania.

Motiváciou pre túto prácu je jej možný pozitívny dopad na všetkých používateľov sociálnych sietí, či už sa jedná o dospelých ľudí alebo zraniteľnejšie skupiny ako sú deti, tínedžeri či starší ľudia.

II. ANTISOCIÁLNE CHOVANIE

Antisociálne správanie môže predstavovať rôzne formy správania sa v online priestore, ktoré negatívne ovplyvňujú používateľov sociálnych sietí a narúšajú pravidlá, normy a hodnoty skupín či spoločností. Jeho zámerom je zvyčajne ublížiť iným ľuďom. Narúša taktiež aj sociálny status jednotlivca, prípadne povzbudzuje k podobnému správaniu sa iných jedincov pod falošnou domnienkou, že je toto správanie správne alebo nebude potrestané. Hranične sa radí taktiež k psychickým poruchám a vzhľadom na fakt, že jedinec si v mnohých prípadoch nemusí uvedomovať ako jeho správanie škodí ľuďom v jeho okolí, je táto porucha veľmi rezistentná na liečbu. Je typickejšia pre mužov ako pre ženy. Prejavy antisociálneho správania sú ľahko zameniteľné s narcizmom a bipolárnou poruchou, avšak pri hlbšej analýze je možné od nich odlíšiť [1].

Antisociálne správanie môže byť typické pre skupinu či jednotlivca a jeho novodobým teritóriom je online priestor. Pod rúškom takzvanej anonymity, ktorú poskytuje absencia stretu „zoči-voči“ s osobou či osobami na druhej strane obrazovky počítača či možnosť skryť svoju identitu za avatar a prezývku, ľudia strácajú zábrany, zabúdajú na morálne hodnoty a častokrát aj strácajú schopnosť triezvo uvažovať. Mnohé z týchto činov sa považujú za priestupky až trestné činy, a preto je morálne aj právne dôležité, aby bolo takéto konanie obmedzované, prípadne

sankcionované a neopakovalo sa. Keďže je internet globálne prístupné médium a jeho používateľmi sú aj nepľnoleté osoby, vysoko senzitivne osoby či osoby s emocionálnou nestabilitou, môže ich takéto správanie ohroziť viac ako bežného používateľa. Deti sú obeťami kyberšikany, epileptici sú obeťami trollův a osoby trpiace depresiami sú povzbudzované k samovraždě v pochybných anonymných skupinách [2].

Existuje viacero druhův antisociálneho správania v online priestore. Ten svojim globálnym dosahom ponúka rozmanité druhy obetí, a možnosti, ako im uškodiť, teda každým dňom pribúdajú. Rozvojom internetu sa cieľová skupina rozšírila na deti, pri ktorých je okrem kyberšikany rozšírená aj krádež identity, šírenie detskej pornografie či obťažovanie. Rizikovým faktorom je najmä ich nízka úroveň digitálnej gramotnosti. O nebezpečenstvách na internete vznikol v roku 2020 dokument „V sítí“ [3]. V tomto dokumente sa okrem iného uvádzajú nasledovné údaje: 60 % detí nemá rodičmi nijako limitovaný čas strávený na internete, každé šieste dieťa zdieľa svoje intímne fotografie a videá, tretina detí by sa pri zneužití ich fotografií zdráhala niekomu zdôveriť, každé druhé dieťa chatuje s cudzími ľuďmi, pätina by neodmietla osobné stretnutie, 41 % detí potvrdilo, že dostalo od druhej osoby pornografickú fotografiu.

Najčastejšie sa stretávame s nasledovnými formami:

- šírenie detskej pornografie
- kyberšikana
- trolovanie
- falošné účty
- obťažovanie
- šírenie poplašných správ
- predaj nelegálnych produktův
- spamové správy,
- podozrivé komentáre,
- dezinformácie a falošné správy.

Hanblivosť, nízke sebavedómie, sociálna úzkosť, depresia, no najmä osamelosť sú veľmi úzko spájané so závislosťou na internete. Nedostatok sociálneho kontaktu, nepochopenie od okolia, slabé či narušené spoločenské väzby pomáhajú uzatvárať človeka do seba. Introverti ďalej používajú internet na sociálny kontakt s inými jednotlivcami, namiesto začleňovania sa do komunit v reálnom svete, ako je práca alebo škola. Mentálne problematickí či narušení jedinci dokážu lepšie fungovať v online priestore ako v reálnom svete, pretože dokážu kamuflovať či potlačiť aspekty, ktoré ich vylučujú z bežných komunit. Typickým správaním je aj kopírovací mechanizmus, pri ktorom kopírujú správanie oponenta aby pôsobili vyrovnaně a normálne [4].

Pri boji s aktérmi antisociálneho správania máme už dnes nejaké možnosti. Takmer každá platforma sociálneho webu si niektorú osvojila, alebo používa kombináciu viacerých. Pri menej vyťažených platformách a stránkach, ako sú napríklad stránky e-shopov, zvyčajne nepozorujeme veľkú automatizáciu a filtrovanie nevhodného obsahu, ako sú nerelevantné komentáre, recenzie či spam sú skôr na pleciach administrátora. V posledných rokoch pozorujeme čoraz väčší záujem o automatizované filtrovanie obsahu a používateľov. Video platforma YouTube takto automatizovala detekciu obsahu, ktorý neprináleží sprostredkujúcej osobe. Takýto používateľ je sankcionovaný demonetizovaním videa a pri opakovaní takéhoto správania zablokovaním a následným vymazaním účtu. V súčasnosti sa najviac používajú nasledovné formy obmedzovania antisociálneho chovania:

- ručné nahlasovanie ostatnými používateľmi platformy
- moderovanie online diskusií správcom
- strojové učenie sa detekčných modelov
- analýza sentimentu

Používanie metód strojového učenia už teraz dosahuje zaujímavé výsledky. V prípade falošných správ a dezinformácií sa v práci [5] podarilo za pomoci Naivného Bayesa dosiahnuť presnosť až 93%, pričom za metódami hlbokého učenia zaostával len 2%. Ďalšia práca [6] sa zamerala na použitie Strojov s podpornými vektormi (SVM) a najvyššia dosiahnutá presnosť bola 94%. Experimentovali aj s metódou Náhodných lesov (RF) a dosiahli 92% presnosť.

III. PRÍSTUP K BUDOVANIU DETEKČÝCH MODELOV

A. Popis Dátových Množín

Všetky použité dátové sady boli prevzaté z portálu Kaggle, kde sú voľne dostupné.

Dataset „spamových správ“ obsahoval údaje o 5172 náhodne vybraných správ, ktoré buď obsahovali alebo neobsahovali spam. Dataset obsahuje 3002 stĺpcov, pričom prvý indikuje emailové meno, z dôvodu ochrany osobných údajov pozmenené na číselné označenie. Posledný stĺpec obsahoval labelling - označenie, či sa jedná o spam (1) alebo nejedná (0). Zvyšných 3000 stĺpcov – atribútov obsahovalo najfrekventovanejšie slová vo všetkých správach, pričom hodnota atribútu prezentovala početnosť výskytu slova v správe.

Dataset „podozrivá komunikácia“ pochádza z Twitteru a Facebookových. Príspevky zahŕňajú podozrivé aktivity ako sú rasizmus, diskriminácia, urážlivé slová, vyhrážanie a iné prejavy kyberšikany. Dataset obsahuje len 2 atribúty - atribút „tagging“ obsahuje hodnoty 1 pre podozrivé správy a 0 pre nepodozrivé správy a atribút „comments“ konkrétne správy získané zo sociálnych sietí. Dataset obsahuje 20001 správ, pričom 7822 z nich sú podozrivé správy a zvyšných 12179 správ je neutrálnych.

Dataset „dezinformácie a falošné správy“ bol vytvorený v rámci systému CONSTRAINT-2021 na detekciu nepriateľských príspevkov. Zameriava sa na odhaľovanie falošných správ súvisiacich s COVID19 v angličtine. Zdrojom údajov sú rôzne platformy sociálnych médií, ako napríklad Twitter, Facebook, Instagram a iné. Príspevky získané zo sociálnych sietí boli označené ako falošné alebo skutočné správy. Dataset obsahoval 3 atribúty – atribút „id“ s číslom riadku, atribút „tweet“ s textom správy a atribút „label“ s hodnotami „fake“ a „real“ podľa toho, či je daná správa pravdivá alebo falošná. Tvorí ho 6420 správ, pričom 3360 z nich sú pravdivé a 3060 z nich sú falošné správy.

B. Použité Metódy

Na základe analýzy podobných prác boli zvolené nasledovné algoritmy: Naivný Bayes, metóda podporných vektorov (SVM) a náhodný les (RF).

V porovnaní s inými metódami strojového učenia, je naivná bayesová klasifikácia jednoduchá metóda na implementáciu, pričom dosahuje relatívne nízku chybovosť. Pri vytváraní modelu, teda pri tréovaní nepotrebuje veľké množstvo dát. Klasifikátor vychádza z predpokladu, že efekt, ktorý ma hodnota každého atribútu na danú triedu, nie je ovplyvnený hodnotami ostatných atribútov. Čiže požaduje aby predikátory boli medzi sebou nezávislé. Kvôli tomuto zjednodušeniu sa tento klasifikátor nazýva aj ako „naivný“. Naivný bayesov klasifikátor našiel svoje uplatnenie pri analýze sentimentu, filtrovanie spamu, klasifikácii textu alebo doručovacích systémov. Medzi používané naivné bayesové klasifikátory patrí napríklad Gaussovský Naivný Bayes, Multinomický Naivný Bayes alebo Bernoulliho Naivný Bayes [7].

Metóda podporných vektorov (SVM) klasifikuje data spravidla do dvoch tried. V prípade, že sú triedy lineárne separovateľné, hovoríme o lineárnej SVM, v prípade, že nie sú, hovoríme o nelineárnej SVM alebo kernelovej SVM. Na správne fungovanie algoritmu potrebujeme labelovanú tréovaciu množinu. Cieľom SVM algoritmu je nájsť nadrovinu (ang. Hyperplane) v N -rozmernom priestore, kde N je počet príznakov, ktorá rozdeľuje jednotlivé triedy a určuje ktoré body patria do ktorej triedy. Nadrovina môže byť tiež označená ako rozhodovacia hranica (ang. Decision boundary). Vo všeobecnosti na oddelenie tried môže existovať nekonečne veľa oddeľovacích nadrovin, avšak pri metóde SVM hľadáme takú hranicu, ktorá vymedzuje po celej dĺžke čo najširšiu vzdialenosť medzi hraničnými bodmi jednotlivých tried. Pri tréovaní SVM, body, ktoré sa nachádzajú najbližšie k nadrovine sa nazývajú podporné vektory (ang. Support vectors) a pomocou týchto bodov maximalizujeme šírku hranice, tieto body nám pomáhajú budovať SVM. Šírka hranice reprezentuje kritériálnu funkciu [8].

Náhodný les (angl. Random forest) je metóda založená na kombinácii viacerých dekorelovaných rozhodovacích stromov, kde každý strom poskytne unikátny pohľad na riešenie problematiky. Každý strom závisí na hodnotách náhodného vektora hodnôt vzorkovaného nezávisle a s rovnakým rozložením pre všetky stromy v tzv. lese stromov. Používa sa ak jeden rozhodovací strom nie je dostatočne presný. Ide teda o metódu, ktorá vytvára viacero rozhodovacích stromov, pomocou ktorých získavame presnejší výsledok [9].

IV. NÁVRH RIEŠENIA

Tréovanie modelov, experiment a testovanie boli vykonávané vo voľne dostupnom softvéri Weka. Tréovanie všetkých modelov prebehlo za pomoci 5-násobnej krížovej validácie.

Pri tréovaní modelov na prvej dátovej množine „spamových správ“ nebolo nutné dáta upravovať a čistiť. V programe Weka sa definoval atribút „Prediction“ označujúci hodnoty pre spam či nie-spam ako nominálny atribút (program Weka ho automaticky registroval ako numerický atribút). Ide však o veľkú dátovú množinu (31,24 MB), a jej spracovanie bolo veľmi náročné na čas a procesor. Na týchto dátach boli tréované modely použitím Naivnej Bayesovej metódy (NB), metódy podporných vektorov (SVM) a metódy generovania náhodného lesa (RF). Výsledky testovania týchto modelov sú uvedené v Tabuľke 1.

Výsledky testovania ukázali, že najvhodnejšia metóda na generovanie detekčného modelu pre rozpoznávanie spamových správ je prekvapujúco Naivný Bayes. Model SVM má v tomto prípade najnižšiu úspešnosť pri správnej klasifikácii tried. SVM je model známy svojou ťažkou interpretáciou, a teda spätnou analýzou výsledkov. Jedným z dôvodov môže byť nevyvážený dataset, kde 3672 vzoriek je označených ako príklady bez spamu a 1500 ako spam. Model SVM sa podvedome snaží predvídať zatriedenie vzorky do triedy, ktorú zastupuje viac hodnôt, aby dosiahol čo najvyššiu percentuálnu úspešnosť.

Tabuľka 1

Výsledky testov modelov generovaných metódami strojového učenia NB, SVM a RF nad dátami „spamových správ“ pomocou mier efektívnosti: Úspešnosť, Presnosť, Návratnosť a F1-skóre.

	NB	SVM	RF
Úspešnosť	0.850	0.724	0.804
Presnosť	0.846	0.801	0.838
Návratnosť	0.850	0.724	0.807
F1 skóre	0.846	0.621	0.774

Pri tréovaní modelov na dátovej množine „podozrivá komunikácia“ bola táto množina pred spracovaním programom Weka vyčistená od znakov – „‘,“ a následne boli hodnoty atribútu „tweet“ komentované do dvojitéch úvodzoviek, aby ich Weka dokázala spracovať ako text. V samotnom programe Weka bol následne zmenený typ atribútu „tagging“ z numerickej hodnoty na nominálnu hodnotu. Výsledky testovania týchto modelov sú uvedené v Tabuľke 2.

Tabuľka 2

Výsledky testov modelov generovaných metódami strojového učenia NB, SVM a RF nad dátami „podozrivá komunikácia“ pomocou mier efektívnosti: Úspešnosť, Presnosť, Návratnosť a F1-skóre.

	NB	SVM	RF
Úspešnosť	0.956	0.608	0.965
Presnosť	0.967	0.609	0.967
Návratnosť	0.966	1.000	0.963
F1 skóre	0.965	0.757	0.965

Výsledky testovania ukázali, že Modely Naivný Bayes a Random Forests pri tréovaní na datasete „podozrivá komunikácia“ dosiahli veľmi dobré výsledky. Opäť však zaostáva model SVM.

Dataset „dezinformácie a falošné správy“ bol taktiež pred spracovaním programom Weka vyčistený od znakov – „‘,“ a následne boli hodnoty atribútu „tweet“ komentované do dvojitémi úvodzoviek, aby ich Weka dokázala spracovať ako text. Aj na tomto datasete bili tréované modely pomocou algoritmov NB, SVM a FR. Výsledky testovania týchto modelov sú uvedené v Tabuľke 3.

Tabuľka 3

Výsledky testov modelov generovaných metódami strojového učenia NB, SVM a RF nad dátami „dezinformácie a falošné správy“ pomocou mier efektívnosti: Úspešnosť, Presnosť, Návratnosť a F1-skóre.

	NB	SVM	RF
Úspešnosť	0.525	0.520	0.525
Presnosť	0.751	0.514	0.751
Návratnosť	0.525	0.519	0.525
F1 skóre	0.618	0.516	0.618

Výsledky testovania ukázali, že tréovanie modelov na datasete „dezinformácie a falošné správy“ o CoViD-19 malo veľmi slabú úspešnosť predikcie pri všetkých použitých modeloch. Tento výsledok je pravdepodobne dôsledkom ťažšieho rozoznania dát či ide alebo nejde o dezinformáciu, vzhľadom na dôležitosť poznania kontextu.

Sumarizácia výsledkov testovania uvedená v Tabuľke 4 ukazuje, že najlepšiu úspešnosť pri detekcii toxického správania globálne vykazuje model Naivný Bayes, za ním je model Náhodných lesov (RF). Pričom model SVM sa v tejto úlohe.

V. ZÁVER

Prínosom tohto príspevku so zameraním na problematiku detekcie toxických správ je porovnanie detekcie rôznych druhov antisociálneho správania, ako aj ukážka úspešnosti rôznych modelov strojového učenia pri detekcii rozdielnych druhov toxického správania. Poukazuje na problém nízkej univerzálnosti využitia jedného modelu strojového učenia, ako aj na neschopnosť

modelov adekvátne spracovať zložitejšie dáta. Možnosťou zlepšenia dosiahnutých výsledkov by mohlo byť použitie nástroja na transformáciu textov do takej foriem reprezentácie, ktoré by následne dokázali modely strojového učenia spracovať a vyhodnotiť s vyššou úspešnosťou, alebo preskúmať metódy strojového učenia a prípadne nájsť vhodnú kombináciu metód na čo najpresnejšiu identifikáciu toxického správania v online priestore.

V budúcnosti by bolo zaujímavé bližšie sa pozrieť na problematiku detekcie falošných a dezinformačných správ, v ktorej použité modely strojového učenia výrazne zaostávali, ako aj problematiku toxických správ rôzneho druhu v slovenskom jazyku, vzhľadom na jeho odlišnosť a komplikovanú gramatiku v porovnaní s jednoduchším anglickým jazykom.

Tabuľka 4

Sumarizácia výsledkov testov modelov generovaných metódami strojového učenia NB, SVM a RF nad tromi dátovými množinami pomocou miery Úspešnosti.

	NB	SVM	RF
Spamové správy	0.850	0.724	0.804
Podozrivá komunikácia	0.956	0.608	0.965
Falošné správy	0.525	0.520	0.525

POĎAKOVANIE

Táto práca vznikla s podporou Slovenskej grantovej agentúry Ministerstva školstva a SAV Slovenskej republiky v rámci projektu VEGA 1/0685/21.

LITERATÚRA

- [1] MARQUES, Paulo Barbosa – PAULINO, Mauro – ALHO, Laura: *Psychopathy and Criminal Behavior*. Academic Press, 2021. 536 s. ISBN: 978-0-12-811419-3.
- [2] KUMAR, Srijan – CHENG, Justin – LESKOVEC, Jure: *Antisocial Behavior on the Web: Characterization and Detection*. In: WWW '17 Companion: Proceedings of the 26th International Conference on World Wide Web Companion. Republic and Canton of Geneva: International World Wide Web Conferences Steering Committee, 2017. s. 947-950. ISBN: 978-1-4503-4914-7
- [3] KOLLÁROVÁ, Jana: *Kto chytá v sieti? Vít Klusák má nový šokujúci dokument o zneužívaní detí na internete* [online]. Bratislava: Pravda, 2020. [cit. 2023-5-18]. Dostupné na internete: <<https://kultura.pravda.sk/film-a-televizia/clanok/544818-kto-chyta-v-sieti-vit-klusak-ma-novy-sokujuci-dokument-o-zneuzivani-deti-na-internete/>>.
- [4] WONGPAKARAN, N. et al: *Loneliness and problematic internet use: testing the role of interpersonal problems and motivation for internet use*. In: BMC Psychiatry. Vol. 21. No. 9. 2021. 447 s.
- [5] KHAN, Junaed Younus et al.: *A benchmark study of machine learning models for online fake news detection*. In: Machine Learning with Applications. Vol. 4. No. 1. 2021. ISSN: 2666-8270.
- [6] PROBIERZ, Barbara – STEFANSKI, Piotr – KOZAK, Jan: *Rapid detection of fake news based on machine learning methods*. In: Procedia Computer Science. Vol. 192. No. 25. 2021. ISSN: 1877-0509.
- [7] HAN, Jiawei – KAMBER, Micheline – PEI, Jian. *Data Mining Concepts and Techniques*. Third Edition. ISBN: 978-0-12-381479-1. [cit. 2023-05-15]. Dostupné na internete: <<http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>>.
- [8] GANDHI, Rohith. 2018. *Support Vector Machine — Introduction to Machine Learning Algorithms, towards data science*. [cit. 2023-05-15]. Dostupné na internete: <<https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>>.
- [9] DÚBRAVA, Juraj. *Využití strojového učení pro predikci vlivu mutací na stabilitu proteinů*. Brno: Fakulta Informačních technologií, Vysoké Učení Technické v Brne, 2018, [cit. 2023-05-15]. Dostupné na internete: <https://www.vut.cz/www_base/zav_prace_soubor_verejne.php?file_id=181194>.

Compensation of the waveform distortion of transformers to match the desired output voltage

¹Šimon GANS, ²Ján MOLNÁR

^{1,2} Department of Theoretical and Industrial Electrical Engineering, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovak Republic

¹simon.gans@tuke.sk, ²jan.molnar@tuke.sk

Abstract — This work describes a method to determine the changes required to the input voltage of a transformer to obtain a desired output voltage of a transform waveform. Ferromagnetic materials usually have significant permeability and conductivity that damps the electromagnetic wave amplitude and dissipate energy. This damping is frequency dependent and because of that the higher frequencies are attenuated more than the lower ones. This can distort fast changing waveforms of magnetic flux flowing through such media. The Fourier analysis of the output voltage is computed for various frequencies of a linearized system. The Fourier analysis of the desired output voltage is computed and then the frequency spectrum of the input voltage that creates the desired output voltage is determined. The analysis only works for linear materials before saturation occurs.

Keywords — distortion, simulation, transformers, transient

I. INTRODUCTION

When dealing with higher frequency signals or non-harmonic signals with fast rising edges, like in the case of the square wave signal, the damping behavior of conductive ferromagnetic materials can attenuate the higher frequency components of the magnetic flux [1] and therefore change the waveform of the signal. Magnetic damping occurs also with lower frequency signals [2]. The diffusive nature of the materials can also be modeled via fractional differential models [3]. Supposed that the desired signal is connected to the primary coil and is also expected to be present on the secondary coil (for example for providing galvanic isolation in a 1:1 transformer). The attenuation is determined by various factors, not just material characteristics, but also the transformer shape and the placement of the windings [2]. In some cases, like providing galvanic isolation, where the windings cannot be placed closely together because of the risk of arcing between them in high voltage applications, the damping can be significant. The damping is one source of error of current transformers [6].

When a transformer is used to drive the gates of power transistors, the damping could cause that the switching times are longer than desired which is unwanted because it creates higher power losses [5]. A toroid core transformer is supposed operating in the linear region of its BH curve before magnetic saturation where the permeability can to some degree of error be supposed to be constant and therefore the material can be assumed to be linear. Material isotropy is also assumed.

II. SIMULATION MODEL

For the modelling part an isolation transformer from a microwave oven transformer control board was selected which has a ferrite core and the windings are placed on the opposite sides of the core. Because 3D simulations of magnetic systems tend to be rather challenging, the toroid is modeled in 2D, where only a cut plane perpendicular to the toroid axis of rotation is taken. The neglected z dimension of the toroid was set up in such a way, that the cross-section of the toroid is the same as its simplified 2D model. The 2D model created in the COMSOL Multiphysics 6.0 FEM simulation software can be seen below (Fig. 1)[7].

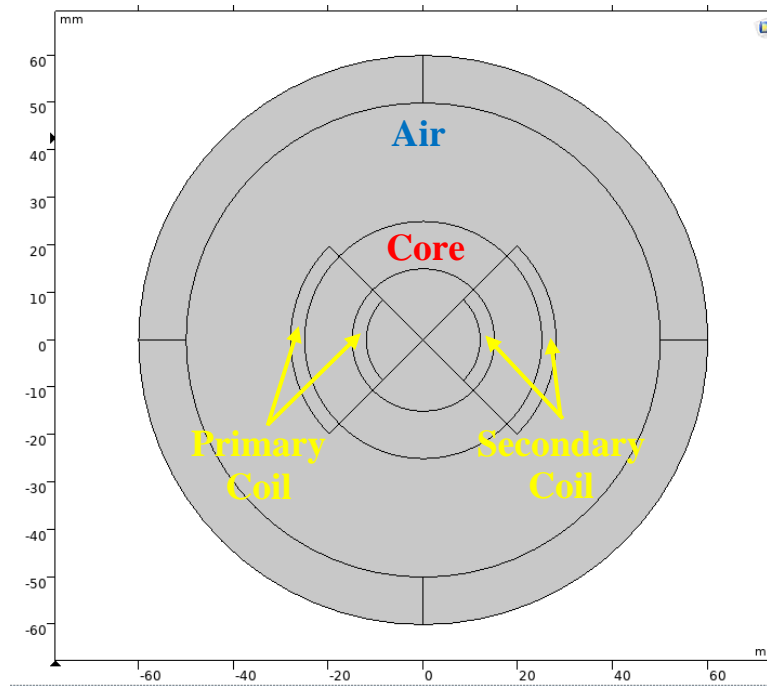


Fig. 1: Simulation model of the toroid core inductor

The material characteristics that have been added to the respective domains in the simulation are shown in the table below (Table 1). The primary and secondary coil were given copper material characteristics and the coil model was said to be homogenized so the effect of eddy currents was not considered. The magnetic permeability of the transformer steel was not written in the table, because it is magnetic field strength dependent and the BH curve graph is shown below (Fig. 2).

Table 1: Material characteristics of the material models in the model

Material	Conductivity [$S \cdot m^{-1}$]	Relative permeability [-]	Relative permittivity [-]
Air	1	1	1
Copper	5.998×10^7	1	1
Iron Powder SMP 1321	2000	N/A (H dependent)	1

The magnetization curve (or the BH curve) of the core defines the magnetic flux density B for each magnetic field intensity H value and therefore defines the permeability. Until a certain point the material permeability is nearly constant, which is shown in the graph below by a red line that has a relative permeability of approximately 51 and coincides with the magnetization curve until a field strength of approximately 8000 A/m.

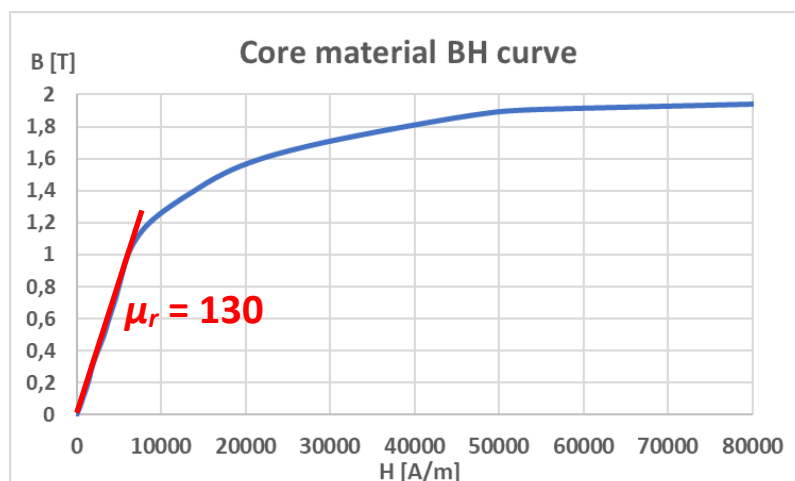


Fig. 2: BH curve of the Amoflux alloy powder core

A frequency domain analysis was set up to test the frequencies between 500 kHz and 200 MHz in steps of 500 kHz. The output of the frequency domain was graph (Fig. 3) which shows the attenuation of the secondary voltage compared to the input voltage and the phase-shift the frequency components create. In the phasor notation, the y-axis of the graphs can be described by equation (1),

$$\dot{y}(f) = \frac{\dot{U}_2(f)}{\dot{U}_1(f)} \quad (1)$$

where U_2 (dot) is the phasor of the secondary harmonic voltage at the specific frequency, U_1 (dot) is the phasor of the primary voltage harmonic and y (dot) is the phasor of their ratio.

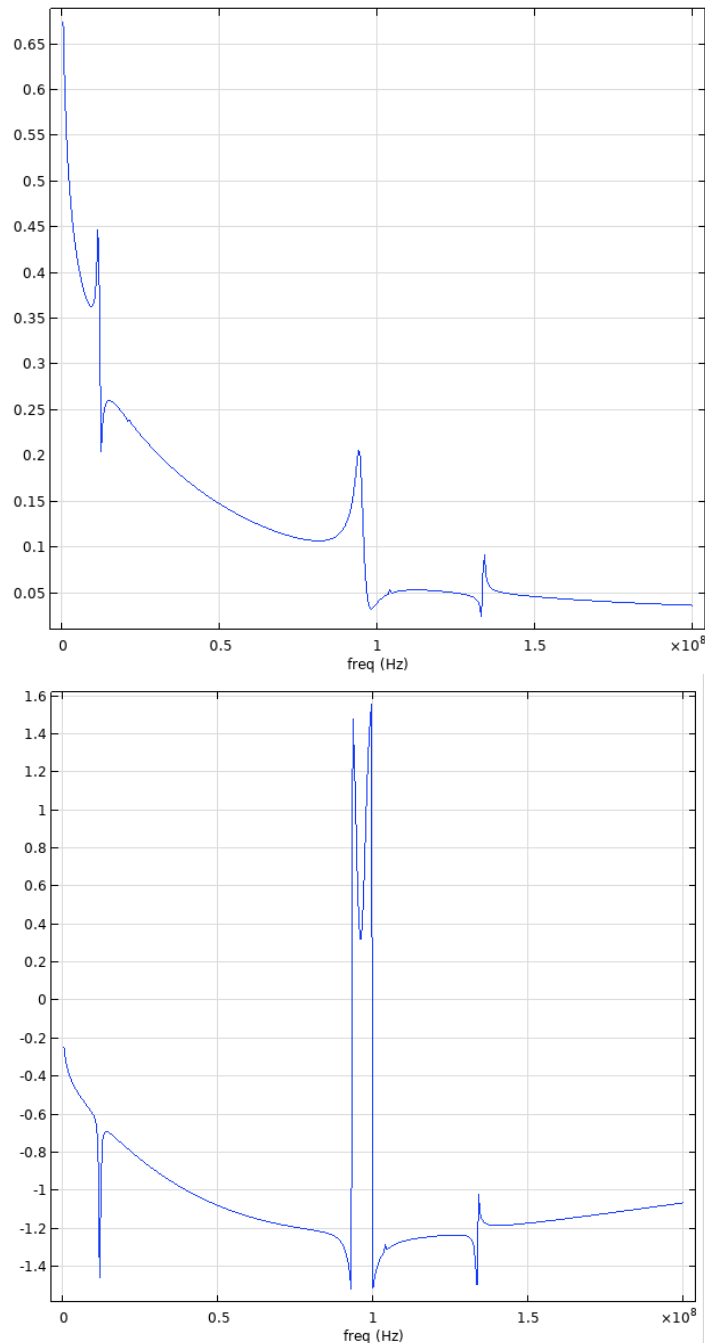


Fig. 3: The amplitude frequency spectrum and the phase shift frequency spectrum of the ratio between the secondary and primary voltage y

The input voltage supplied to the primary coil is a 500 kHz triangle wave with an amplitude of 1 V and it is shown below (Fig. 4). The secondary voltage waveform is distorted compared to the primary voltage and it got attenuated by the dissipative character of the core material. Eddy currents also arise in the material and generate their own magnetic field that works against the magnetizing field and therefore slows down the high frequency flux through the core and hinder the amount of generated voltage.

If the problem is generating a specific secondary voltage waveform like it is in the case for the use in the output of a galvanically isolated frequency generator, this attenuation and waveform

change must be compensated for. Since the attenuation amplitude and phase spectrum of the secondary voltage in reference to the primary voltage were calculated, we can calculate the frequency spectrum of the desired voltage waveform and from that we can calculate the needed primary voltage to create it on the secondary side.

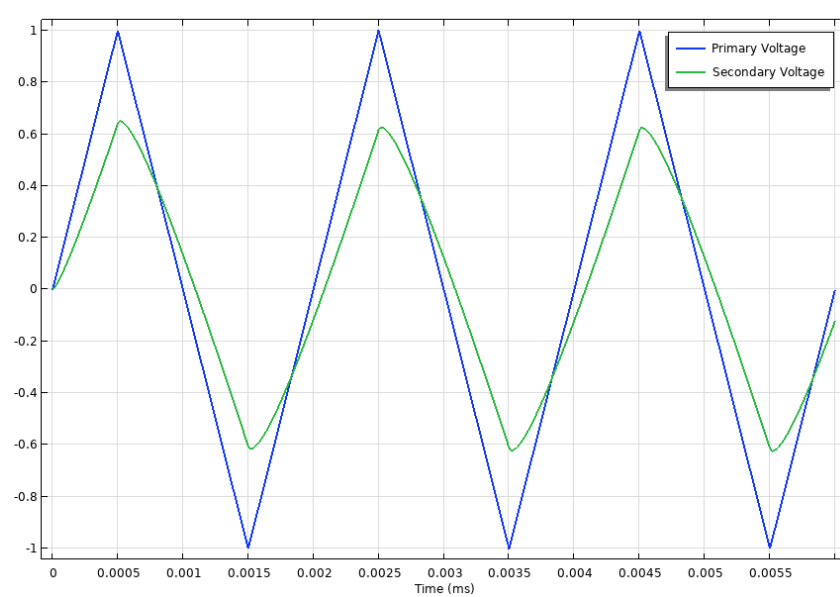


Fig. 4: The primary and secondary voltage of the simulated transformer

III. SPECTRUM COMPUTATION OF THE DESIRED SIGNAL

The frequency spectrum was computed via the MATLAB programming language via the `fft()` function that computes the frequency spectrum via a faster algorithm compared to the standard discrete fourier transform, but requires an input vector size that is a whole number power of 2. A size of 4096 was used[8].

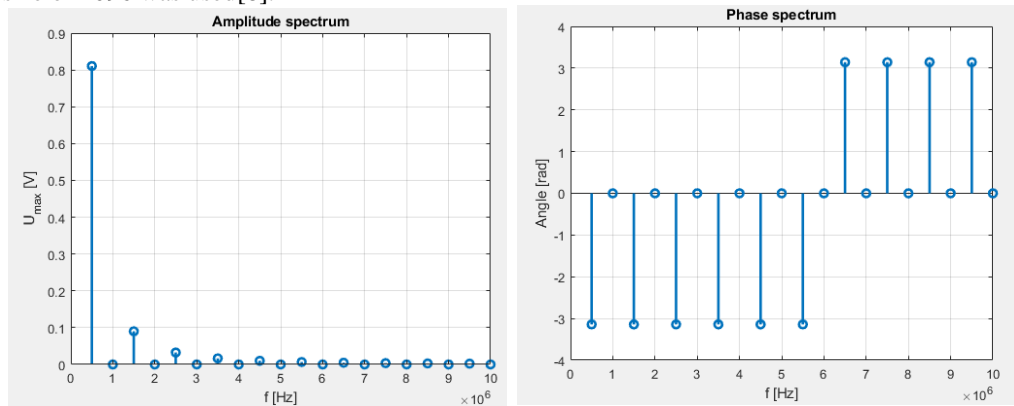


Fig. 5: The amplitude and phase spectrum of the desired 500 kHz triangle wave. The frequency components between 0 and 10 MHz are shown. Higher frequencies have neglectable amplitudes that do not need to be considered.

If we export the data of the frequency analysis in the phasor notation and write the i -th frequency component as $x_i(\text{dot})$, then if we multiply this spectrum by the inverse of the frequency spectrum of the attenuation from (1), we get equation (2),

$$\dot{z}(f) = \dot{x}(f) \cdot \frac{1}{\dot{y}(f)} \quad (2)$$

where $z(\text{dot})$ is the frequency spectrum of the signal that would create the desired triangle wave signal on the secondary coil. To obtain the waveform of the primary voltage $u_1(t)$ we must perform an inverse Fourier transform of the z phasor spectrum (3).

$$u_1(t) = \mathcal{F}^{-1}\{\dot{z}(f)\} \quad (3)$$

For simplicity's sake we can only take the frequency components which amplitude is larger than 0,01, since very low amplitudes do not affect the shape much. The amplitude and phase spectrum are shown below (Fig. 6).

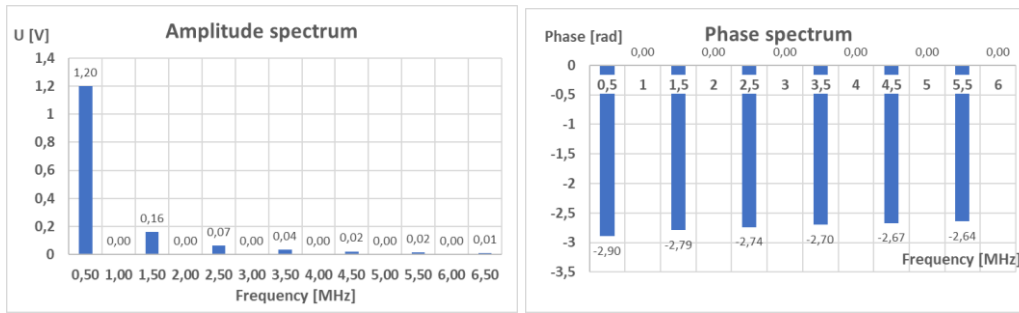


Fig. 6: The amplitude and phase spectrum of the desired signal

When applying the inverse Fourier transform on the given spectrum we obtain the time-dependent waveform (4),

$$u_1(t) \approx 1,201 \cos(\omega t - 2,895) + 0,162 \cos(3\omega t - 2,792) + 0,066 \cos(5\omega t - 2,738) + 0,0363 \cos(7\omega t - 2,7) + 0,0235 \cos(9\omega t - 2,67) + 0,0165 \cos(11\omega t - 2,644) \text{ [V]} \quad (4)$$

where the angular frequency ω is $2 \cdot \pi \cdot 500$ kHz. When we input this waveform into the previous simulation model as the primary voltage, we obtain the following graphs of the waveforms (Fig. 7). The secondary voltage is now a nearly a perfect triangle wave with a frequency of 500 kHz. It does not have sharp edges, because not the whole frequency spectrum was converted to the time domain. If we would include more harmonics, the closer the output will become. Also the output voltage was computed with the nonlinearity of the material BH curve. The linear behavior is just an approximation here and therefore the linear approach of the frequency study holds here to some extent. The primary coil of the transformer has 20 turns of copper wire, and a current was flowing through it with a maximum value of 41 mA. The secondary voltage also had 20 turns. Using equation (5),

$$H = \frac{N \cdot I}{l_m} \quad (5)$$

where N is the number of turns of the primary coil, I is the current flowing through the coil and l_m is the average magnetic field length we can calculate that the average magnetic field strength in the toroid core is 6,528 A/m ($l_m = 0,1256$ m) and so is well within the linear region of the magnetization curve, which holds up to the field strength of 5500 A/m for the given material (in fact it is closer to the Rayleigh region, where the magnetization tends to behave quadratically).

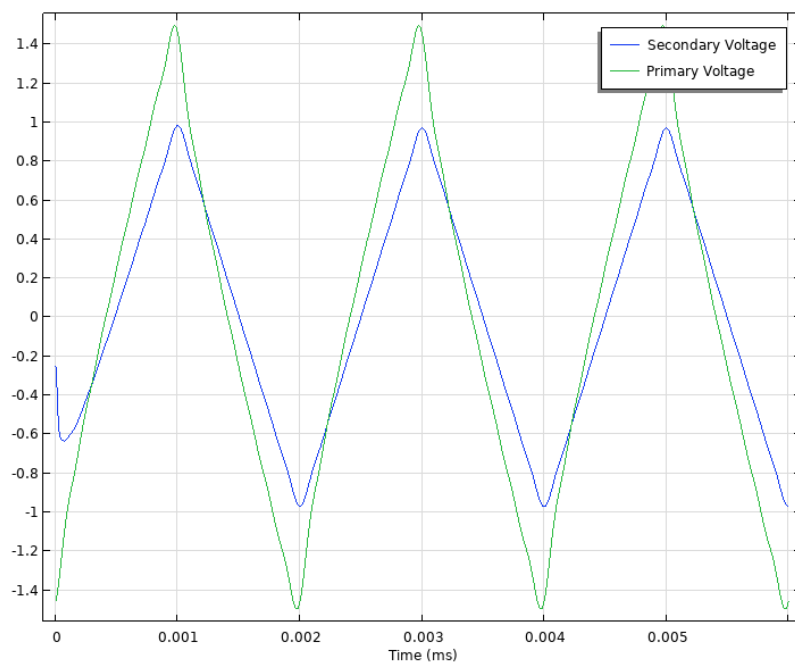


Fig. 7: The primary and the secondary voltage of the transformer after the shape compensation of the primary voltage has been applied.

IV. CONCLUSION

In this work we compensated the waveform distortion that occurs due to the dissipative and

diffusive nature of conductive ferromagnetic materials. This kind of behavior is prominent for example in current sensors too in cases where the measured frequency current spectrum of current transformers exceeds the specified range and if possible should be compensated for. The method of this work can also be used in the case when the output voltage of a frequency generator is electrically isolated via a 1:1 transformer. The attenuation of the voltage waveform compared to the input waveform could be stored in a look up table for various frequencies and then the desired waveform by the user would be compensated to create the given waveform on the secondary side.

ACKNOWLEDGMENT

The authors are grateful for the financial support from the GRANT FEI 2023 project that was specified by the index FEI-2023-91.

REFERENCES

- [1] L. O. Hoefft, Experimental and Theoretical Analysis of the Magnetic Field Attenuation of Enclosures, 1988, IEEE Transactions on Electromagnetic Compatibility, vol. 30, no. 3. DOI: 10.1109/15.3312. Available on the internet at: <https://ieeexplore.ieee.org/document/3312>
- [2] L. Hasselgren, J. Luomi, Geometrical aspects of magnetic shielding at extremely low frequencies, 1995, IEEE Transactions on Electromagnetic Compatibility, vol. 37, issue 3. DOI: 10.1109/15.406530. Available on the internet at: <https://ieeexplore.ieee.org/document/406530>
- [3] B. Ducharne, P. Tsafack, Y. A. Tene Deffo, B. Zhang, G. Sebald, Anomalous fractional magnetic field diffusion through cross-section of a massive toroidal ferromagnetic core, 2021, Communications in Nonlinear Science and Numerical Simulation, Volume 92, 1005450. DOI: 10.1016/j.cnsns.2020.105450. Available on the internet at: <https://www.sciencedirect.com/science/article/pii/S100757042030280X>
- [4] H. A. Haus, J. R. Melcher, Electromagnetic Fields and Energy, Chapter 11. 5. Electromagnetic Dissipation, Hypermedia Teaching Facility, 1998. Available on the internet at: https://web.mit.edu/6.013_book/www/chapter11/11.5.html
- [5] R. Hou, Y. Shen, H. Zhao, H. Hu, J. Lu, T. Long, Power Loss Characterization and Modeling for GaN-Based Hard-Switching Half-Bridges Considering Dynamic ON-State Resistance, 2020, IEEE Transactions on Transportation Electrification, Vol. 6, No. 2. DOI: 10.1109/TTE.2020.2989036. Available on the internet at: <https://ieeexplore.ieee.org/document/9072419>
- [6] A. E. Emanuel, J. A. Orr, Current Harmonics Measurement by Means of Current Transformers, 2007, IEEE Transactions on Power Delivery, Vol. 22, No. 3. DOI: 10.1109/TPWRD.2007.900108. Available on the internet at: <https://ieeexplore.ieee.org/document/4265673>
- [7] COMSOL, AC/DC Module user's guide for Comsol Multiphysics 5.4., 2018, pages 214-260. Available on the internet at: <https://doc.comsol.com/5.4/doc/com.comsol.help.acdc/ACDCModuleUsersGuide.pdf>
- [8] MathWorks, MATLAB Documentation, fft – Fast Fourier transform. Available on the internet at: <https://www.mathworks.com/help/matlab/ref/fft.html>

Analýza výkonnosti smerovacieho protokolu AODV a jeho modifikácií pre MANET siete

¹Natalia KURKINA, ²Ján PAPAĽ

^{1,2} Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹natalia.kurkina@tuke.sk, ²jan.papaj@tuke.sk

Abstrakt — Tento článok sa zaoberá výskumom MANET siete a reaktívneho protokolu AODV a jeho vylepšeniam pomocou algoritmov umelej inteligencie K-means a GMM (Gaussian Mixture Model). Cieľom štúdie bolo preskúmať vylepšenia týchto protokolov z hľadiska výkonnosti v MANET prostredí. V analýze sme zistili, že implementácia K-means algoritmu v protokole AODV umožňuje efektívne rozdelenie uzlov do skupín, čo viedlo k zníženiu počtu služobných paketov v sieti pozostávajúcej s veľkého množstva mobilných zariadení. Vylepšenie protokolu AODV pomocou algoritmu GMM nevedlo k zlepšeniu výkonnosti siete.

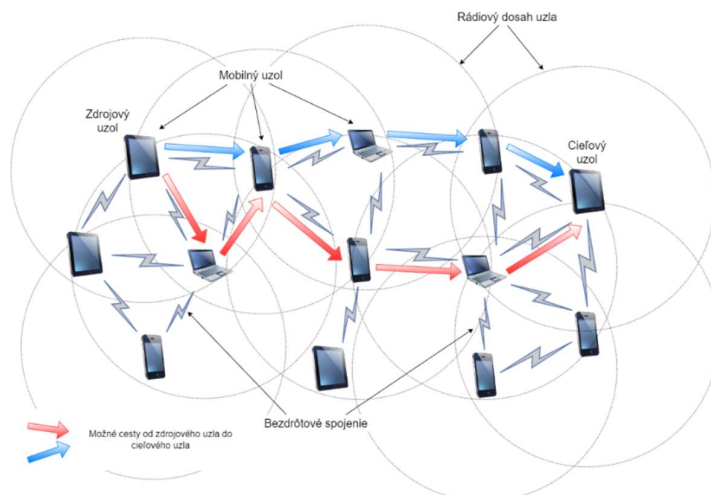
KLúčové slová — MANET, AODV, strojové učenie, smerovací protokol, K-means, GMM

I. ÚVOD

Mobilná ad hoc sieť (MANET) je bezdrôtová sieť, ktorá sa vytvára bez použitia statickej infraštruktúry pomocou mobilných uzlov. Tieto siete sú decentralizované, konfigurovateľné a majú dynamickú topológiu. Smerovanie je jednou zo základných funkcií v MANET, ktorá umožňuje uzlom navzájom komunikovať pomocou nájdenia optimálnej cesty zo zdrojového uzla do cieľového uzla. [1]

Tradičné smerovacie protokoly nie sú pre MANET vhodné. Vedci preto vyvinuli niekoľko smerovacích protokolov špeciálne pre tento typ siete. Tieto protokoly však majú aj určité nevýhody, takže smerovanie v MANET stále predstavuje niekoľko výziev. Obmedzený výkon batérie mobilných zariadení, dynamická povaha topológie siete a nedostatok centralizovaného riadenia sú niektoré z významných problémov, ktoré ovplyvňujú efektívnosť a spoľahlivosť smerovania v MANET sieťach. [2]

Na riešenie týchto výziev vedci navrhli použitie techník umelej inteligencie (UI) pri smerovaní v MANET sieťach. Smerovacie protokoly založené na UI používajú algoritmy strojového učenia na prijímanie inteligentných rozhodnutí o smerovaní na základe dynamických podmienok siete. Tieto protokoly sa môžu prispôbiť zmenám siete, optimalizovať proces výberu vhodných na smerovanie uzlov a znížiť počet služobných správ v MANET sieťach. [3]



Obr. 1 Základný princíp fungovania MANET.

II. SMEROVACIE PROTOKOLY V SIEŤACH MANET

Smerovacie protokoly zohrávajú rozhodujúcu úlohu pri vytváraní a udržiavaní komunikácie medzi mobilnými zariadeniami v sieťach bez pevnej infraštruktúry. Tieto protokoly umožňujú efektívny prenos dát v dynamickom a samokonfigurujúcom sieťovom prostredí. V MANET sa bežne používa niekoľko smerovacích protokolov, z ktorých každý má svoje vlastné charakteristiky a vhodnosť pre rôzne scenáre. [2]

A. Proaktívne protokoly.

Proaktívne smerovacie protokoly udržiavajú aktuálne informácie o smerovaní nepretržitou výmenou riadiacich správ medzi sieťovými uzlami. Tento prístup zaisťuje, že cesty sú v prípade potreby ľahko dostupné, ale v dôsledku častých výmen správ sieť môže byť vyťažena preposielaním veľkého počtu služobných správ, preto môžu byť neefektívne vo veľkých sieťach alebo vysoko mobilných prostrediach. [3]

B. Reaktívne protokoly.

Reaktívne smerovacie protokoly, vytvárajú cesty len v prípade, keď uzol potrebuje poslať údaje. Tieto protokoly posielajú služobné pakety na objavovanie ciest iba v prípade potreby. Na druhej strane tento prístup sa zavádza zvýšené oneskorenie, pretože sa vyžaduje čas na zisťovanie cesty od zdroja do cieľa. [3]

C. Hybridné protokoly.

Hybridné protokoly kombinujú proaktívne a reaktívne prístupy, aby využili výhody oboch. Napríklad Zone Routing Protocol (ZRP) rozdeľuje sieť na zóny, kde sa na smerovanie v rámci zóny používa proaktívny protokol a na medzizónové smerovanie sa používa reaktívny protokol. Tento hybridný prístup zlepšuje škálovateľnosť a znižuje počet posielaných služobných správ pri zachovaní efektívneho smerovania. [3]

D. Geografické protokoly.

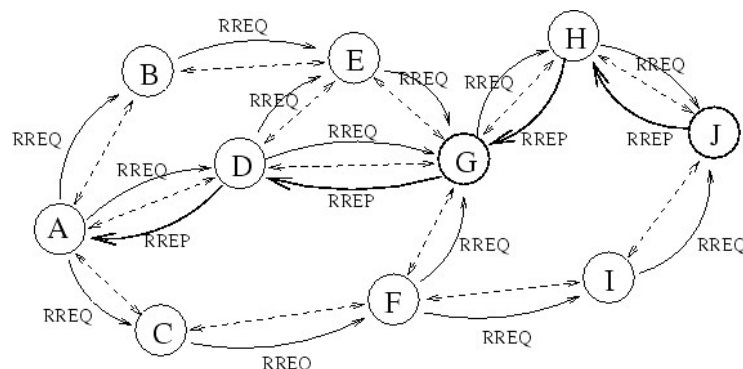
Geografické smerovacie protokoly využívajú informácie o polohe uzlov na rozhodovanie o smerovaní. Uzly si vymieňajú súradnice polohy a používajú ich na určenie ďalšieho hopu smerom k cieľu. Tieto protokoly sú vhodné pre scenáre, kde sú informácie o polohe ľahko dostupné a môžu dosiahnuť efektívne smerovanie bez zachovania smerovacích tabuliek. [3]

Stojí za zmienku, že MANET sú vysoko dynamické siete a topológia siete sa môže často meniť v dôsledku mobility uzlov. Táto dynamická povaha predstavuje ďalšie výzvy pre smerovacie protokoly, pokiaľ ide o stabilitu cesty, prispôsobivosť a energetickú účinnosť. [4]

Na riešenie týchto výziev boli navrhnuté rôzne vylepšenia a optimalizácie. Jednou z možností vylepšenia protokolov je použitie algoritmov umelej inteligencie. Cieľom týchto rozšírení je zlepšiť spoľahlivosť, výkon a bezpečnosť komunikácie v MANET, ktoré zahŕňajú špecifické požiadavky a obmedzenia týchto siete.

III. AD HOC ON-DEMAND DISTANCE VECTOR SMEROVACÍ PROTOKOL A JEHO MODIFIKÁCIE

Ad hoc On-Demand Distance Vector (AODV) je reaktívny smerovací protokol široko používaný v mobilných ad hoc sieťach. AODV je navrhnutý tak, aby minimalizoval počet služobných paketov, potrebných na smerovanie, a zároveň efektívne objavoval a udržiaval cesty v sieti. [5]



Obr. 2 Komunikácia v MANET sieti pri využití smerovacieho protokolu AODV [6]

Kľúčové mechanizmy fungovania AODV protokolu sú:

1. Zisťovanie cesty: keď chce zdrojový uzol prenášať údaje do cieľového uzla, iniciuje proces zisťovania cesty vysielať paket žiadosti o cestu (RREQ). Stredné uzly preposielajú RREQ všetkým svojim susedom. Tento proces pokračuje, kým RREQ nedosiahne cieľ alebo stredný uzol s platnou cestou do cieľa.
2. Odpoveď uzla: po prijatí RREQ cieľový alebo medziľahlý uzol s platnou cestou vygeneruje paket odpovede na požiadavku (RREP). RREP obsahuje poradové číslo cesty, počet skokov a ďalšie potrebné informácie. RREP je unicast, takže sa vráti späť do zdrojového uzla po ceste stanovenej RREQ.
3. Údržba cesty: AODV využíva mechanizmus poradového čísla na aktuálnosť a zabezpečenie ciest bez slučiek.
4. Posielanie správ o chybách v sieti: AODV detekuje a spracováva zlyhania spojenia pomocou Route Error (RERR) pakety. Keď uzol zistí nefunkčnú alebo neplatnú cestu, vysiela správu RERR, ktorá upozorní uzly na zneplatnenie príslušných ciest. Tento proaktívny prístup pomáha udržiavať presnosť a efektívnosť smerovacích informácií.
5. Ukladanie cesty do vyrovnávacej pamäte: AODV využíva tento mechanizmus na ukladanie nedávno použitých trás. Keď uzol dostane RREP, uloží cestu do vyrovnávacej pamäte pre budúce použitie. Ak uzol prijme dátový paket pre cieľ s platnou cestou vo svojej vyrovnávacej pamäti, môže paket okamžite poslať ďalej bez potreby zisťovania cesty.

AODV ponúka niekoľko výhod pre MANET, ako napríklad škálovateľnosť. AODV sa dobre prispôbuje veľkosti siete, pretože vytvára cesty iba v prípade potreby. Ďalšou výhodou AODV je efektívne využitie zdrojov. Vytvorením trás na požiadanie AODV optimalizuje využitie sieťových zdrojov, ako je šírka pásma a energia, minimalizácia zbytočných výmen riadiacich správ. Tak tiež AODV má relatívne nízke oneskorenie pri vytváraní cesty kvôli svojej reaktívnej povahe. Vďaka tomu je vhodný pre vysoko dynamické siete s častými zmenami topológie. [5]

Ale má však AODV aj určité obmedzenia, ako napríklad to že vyžaduje pravidelné správy o údržbe cesty na obnovenie ciest a detekciu zlyhaní spojenia, čo môže byť veľmi nákladné pre sieť. Ďalšou nevýhodou je aj škálovateľnosť vo veľkých sieťach, lebo s rastúcou veľkosťou siete sa môže zvyšovať počet služobných paketov posielaných uzlami. Použitie ukladania cesty do vyrovnávacej pamäte to však môže do istej miery zmierniť. Tak tiež AODV vo svojej podstate nepodporuje požiadavky na kvalitu služieb (QoS), ako sú obmedzenia šírky pásma alebo oneskorenie.

Kvôli týmto nevyhodám existuje veľké množstvo vylepšení AODV protokolu. AODV K-means je jedným z príkladov takého vylepšenia.

A. Vylepšenie AODV protokolu pomocou metóda K-means

Modifikácia protokolu AODV pomocou metódy K-means bola navrhnutá autormi v [7]. Ako už bolo spomenuté v prípade keď uzol chce komunikovať z uzlom cestu ku ktorému nepozná spusti proces vyhľadávania cesty. V takom prípade bude posielat pakety požiadavky všetkým svojim susedným uzlom. Na obmedzenie počtu poslaných paketov pri procese vyhľadávania cesty autori použili algoritmus K-means, ktorý rozdelí všetky susedne uzly do klastrov. Hlavnými parametrami pri kastrovaní sú vzdialenosť suseda do cieľa, počet chýb pri prenose a počet voľnej vyrovnávacej pamäti susedného uzla.

Na základe tejto modifikácie autori uvádzajú zníženie oneskorenia a počtu služobných paketov v sieti. [7]

B. Vylepšenie AODV protokolu pomocou GMM

Gaussovské klastrovanie, alebo Gaussovské zmesové modely (Gaussian mixture models, GMM), je populárny UI algoritmus používaný na klastrovanie údajov [8]. Cieľom Gaussovho klastrovania je priradiť dátové body klastrom na základe pravdepodobnosti príslušnosti ku každému klastru.

Algoritmus spočiatku vyžaduje, aby bol špecifikovaný počet klastrov a iterácií. Pre každý klastor sú priradené náhodné inicializovane stredne hodnoty a kovariančne matice Gaussovských distribúcií. Na základe týchto parametrov algoritmus vypočíta pravdepodobnosť príslušnosti každého dátového bodu do každého klastra. Vypočíta funkciu hustoty pravdepodobnosti dátového bodu pomocou stredných hodnôt a kovariančných matic Gaussovských distribúcií spojených s každým klastrom. Pravdepodobnosti sa počítajú pomocou Bayesovej vety a predpokladu nezávislých a identicky distribuovaných údajov.

Po vypočítaní pravdepodobnosti algoritmus aktualizuje parametre každého gaussovského rozdelenia (strednú hodnotu a kovarianciu). Kroky 2 a 3 sa opakujú iteratívne, kým sa nedosiahne konvergencia. Konvergencia nastáva, keď zmena parametrov (stredných hodnôt a kovariancií) medzi iteráciami klesne pod preddefinovanú hranicu. Po konvergencii algoritmus priradí každý

dátový bod klastru s najvyššou pravdepodobnosťou. Klaster s najvyššou pravdepodobnosťou predstavuje najpravdepodobnejší klaster, do ktorého dátový bod patrí. [8]

Gaussovské klastrovanie je široko používaným algoritmom v rôznych oblastiach, ako je rozpoznávanie vzorov, spracovanie obrazu a analýza údajov. Preto na vylepšenie protokolu AODV bol nami zvolený tento algoritmus. V navrhnutom vylepšenom algoritme klastrovanie tiež prebehlo vo fáze zisťovania cesty od zdrojového uzla do cieľového, analogické ako to je v protokole AODV K-means. Len že za parametre klastrovanie boli zvolené len súradnice uzlov a vzdialenosť od cieľového uzla.

IV. SIMULÁCIA A ANALÝZA VYSLEDKOV

Úlohou praktickej časti tohto článku je simulácia a porovnanie vylepšených pomocou metódami umelej inteligencie modifikovaných smerovacích protokolov AODV a samotného nezmeneného protokolu AODV. V našom príklade bol na simuláciu pohybu uzlov zvolený sieťový simulátor NS3 [9]. Parametre simulácie sú uvedené v Tabuľke 1. Pri simulácii sme menili počet uzlov na ploche od 50 uzlov do 100. Rýchlosť pohybu uzlov bola náhodne vybraná pre každý uzol v intervale od 0 m/s do 8 m/s.

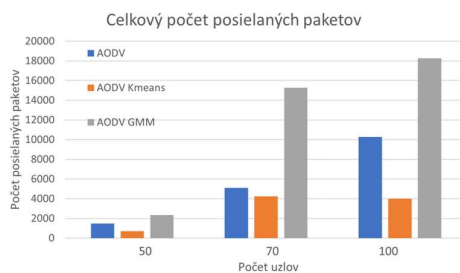
Tabuľka 1
 Parametre simulácie

Parameter	Hodnota
Simulačná oblasť	800 x 800 m ²
Čas simulácie	100 s
Počet uzlov v sieti	50, 70, 100
Minimálna rýchlosť pohybu uzlov	0 m/s
Maximálna rýchlosť pohybu uzlov	8 m/s
Komunikačný protokol	802.11n
Model mobility	Random waypoint

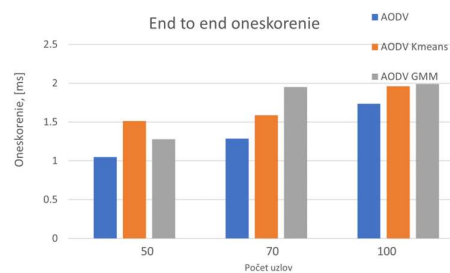
Na vyhodnocovanie algoritmov boli použité nasledujúce parametre [10]:

1. Celkový počet posielaných paketov. Tento parameter zahŕňa všetky pakety, ktoré boli posielane uzlami počas simulácie, aj služobné aj dátové.
2. End to end oneskorenie. End-to-end oneskorenie je definované ako čas potrebný na prenos dátového paketu cez sieť zo zdroja do cieľa.
3. Počet posielaných dátových paketov. Tento parameter zahŕňa len dátové pakety a nepočíta služobné pakety, ktoré uzly vymínajú medzi sebou na zabezpečenie prístupnosti siete.
4. Počet doručených dátových paketov. Tento parameter zahŕňa koľko dátových paketov bolo úspešne prenesené cez sieť od zdrojového uzla do cieľového.

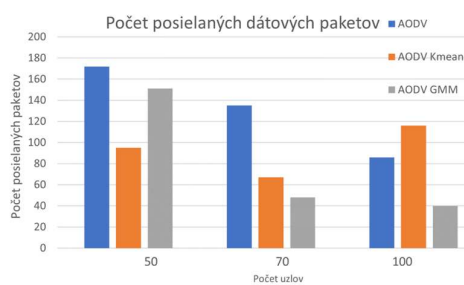
Na obrázkoch 3 až 6 môžeme vidieť výsledky simulácie po všetkých vyššie uvedených parametroch.



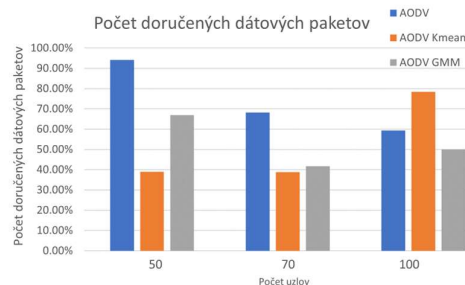
Obr. 3 Celkový počet posielaných paketov



Obr. 4 End to end oneskorenie



Obr. 5 Celkový počet posielaných paketov



Obr. 6 Počet doručených dátových paketov

AODV K-means v prípade keď sieť pozostáva s veľkého počtu uzlov (100 uzlov) ukazuje výrazný pokles celkového počtu posielaných paketov a výrazný rast poslaných dátových paketov a úspešne doručených dátových paketov v porovnaní s druhými dvoma protokolmi. To znamená že v prípade veľkého počtu uzlov pri použití štandardného protokolu AODV a protokolu AODV GMM sieť je viac víťazná služobnými paketmi a preto počet dátových paketov aj poslaných aj doručených klesá. End to end oneskorenie je približne pre všetky 3 algoritmy rovnaké pri počte uzlov 100. Tak tiež môžeme vidieť že AODV GMM ukazuje najvyššie hodnoty celkového počtu poslaných paketov, oneskorenia a najnižšie hodnoty posielaných a úspešne doručených dátových paketov. AODV protokol je najlepší v prípade keď v sieťe nie je veľký počet uzlov. Rýchlejšie sa objaví cesta a tým pádom prenesie vyšší počet dátových paketov. Algoritmus AODV GMM neukazuje dobre výsledky, to môže súvisieť s parametrami ktoré boli využité pri klastrovaní.

V. ZAVER

Cieľom tejto štúdií bolo preskúmanie protokolu AODV v kontexte MANET siete. Tak tiež zámerom bolo skúmanie výkonnostných aspektov tohto protokolu a porovnanie štandardného protokolu s jeho vylepšenými verziami pomocou algoritmov K-means a GMM (Gaussian Mixture Model).

Výsledky našej analýzy naznačujú, že implementácia K-means algoritmu v protokole AODV prináša zlepšenie v prípadoch veľkého počtu uzlov v sieťi. K-means algoritmus nám umožnil efektívne rozdeliť uzly do skupín na základe ich vlastnostnej, čo viedlo k lepšiemu smerovaniu a minimalizácii služobnej prevádzke pri komunikácii. GMM algoritmus, na druhej strane neposkytol zlepšenia, ale to môže súvisieť s parametrami na základe ktorých sa uzly bolo klasifikované do klastrov.

V budúcnosti by bolo zaujímavé ďalej skúmať a vyhodnocovať výkonnosť týchto algoritmov na reálnych MANET siatiach s rôznymi topológiami. Ďalšie vylepšenia a optimalizácie protokolu AODV s použitím K-means a GMM by mohli viesť k ešte lepším výsledkom a rozšíreniu možností pre efektívne a spoľahlivé komunikácie v mobilných ad hoc sieťach.

POĎAKOVANIE

Tento výskum bol financovaný agentúrou na podporu výskumu a vývoja SR, výskumný grant č. APVV-17-0208 a VEGA 1/0260/23.

REFERENCIE

- [1] V. K. Quy, N. T. Ban, V. H. Nam, D. M. Tuan, N. D. Han, "Survey of recent routing metrics and protocols for mobile ad-hoc networks," *Journal of Communications*, vol. 14, no. 2, pp. 110-120, 2019.
- [2] A. Srivastava, A. Mishra, B. Upadhyay and A. k. Yadav, "Survey and overview of Mobile Ad-Hoc Network routing protocols," 2014 International Conference on Advances in Engineering and Technology Research (ICAETR - 2014), pp. 1-6, 2014.
- [3] K. Ahmad, N. I. Udzir, G. Ch. Deka "Opportunistic Networks: Mobility Models, Protocols, Security, and Privacy," 2018.
- [4] A.H. Mohsin, "Optimize Routing Protocol Overheads in MANETs: Challenges and Solutions: A Review Paper", *Wireless Personal Communications*, vol. 126, no. 4, pp. 2871-2910, 2022.
- [5] RFC3561, C. Perkins, E. Belding-Royer, S. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing," 2003.
- [6] F. Al-Dhief, N. Sabri, M. Salim, S. Fouad, S. Aljunid, "MANET Routing Protocols Evaluation: AODV, DSR and DSDV Perspective", *MATEC Web of Conferences*, 2018.
- [7] Z. Shirazi, S. Mirabedini, "Dynamic K-Means Algorithm for Optimized Routing in Mobile Ad Hoc Networks," *International Journal of Computer Science & Engineering Survey*, №7, pp. 01-14, 2016.
- [8] J. Diaz-Rozo, C. Bielza, P. Larranaga, "Clustering of Data Streams With Dynamic Gaussian Mixture Models: An IoT Application in Industrial Processes," in *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3533-3547, 2018
- [9] NS-3 Network Simulator Manual, NS-3 Consortium, 2021.
- [10] N. Shah, H. El-Ocla and P. Shah, "Adaptive Routing Protocol in Mobile Ad-Hoc Networks Using Genetic Algorithm," in *IEEE Access*, vol. 10, pp. 132949-132964, 2022.

Analýza smerovacích protokolov na báze strojového učenia pre MANET siete

¹ Kristián ŠIMON, ²Natalia KURKINA, ³Ján PAPAĽ

^{1,2,3} Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹ kristian.simon@student.tuke.sk, ²natalia.kurkina@tuke.sk, ³jan.papaj@tuke.sk

Abstrakt — Siete MANET majú mnohé prospešné vlastnosti, no zároveň majú aj problémy, s ktorými sa v prípade sietí založených na tradičnej infraštruktúre nestretávame. Najväčšou výhodou týchto sietí je možnosť voľného pohybu uzlov, čo následne komplikuje proces smerovania. Bolo vytvorených mnoho algoritmov pracujúcich na rôznych princípoch, ktoré si dokážu poradiť s dynamickou topológiou sietí MANET a sú tak vhodné na vykonávanie smerovacích procesov. Aj tieto protokoly však majú svoje obmedzenia. Aby sme zvýšili efektívnosť smerovania, často používame algoritmy strojového učenia, preto v tomto článku venujeme svoju pozornosť analýze smerovacích protokolov pre MANET siete, ktoré sú založené na báze strojového učenia.

Keywords — MANET, smerovanie, strojové učenie, AODV, GPSR, AODV Kmeans, AntHocNet

I. ÚVOD

Mobilné ad hoc siete sú dnes jedným z najnovších typov sietí, o ktoré je obrovský záujem vďaka svojim výhodným vlastnostiam, ako je napríklad vynikajúca mobilita. Hoci siete MANET vyzerajú veľmi perspektívne, majú špecifické problémy, s ktorými sa v prípade sietí založených na tradičnej infraštruktúre nestretávame. Hlavnými problémami sú smerovanie, nepoškodený príchod balíkov na miesto určenia, minimalizácia meškania a bezpečnosť používateľov a ich dát pri zachovaní vysokej úrovne mobility. Na odstránenie vyššie uvedených problémov boli vytvorené rôzne protokoly. [1]

Na zvýšenie efektivity smerovacích protokolov používaných v MANET sieťach, často tiež sa využívajú aj algoritmy strojového učenia. Cieľom tejto práce je preskúmať smerovacie procesy v sieťach MANET, ktoré zabezpečujú rôzne smerovacie protokoly a algoritmy strojového učenia. V prvej časti práce sa zaoberáme a podrobne popisujeme charakteristiky, princíp fungovania a využitie MANET sietí. Ďalej preskúmame smerovacie protokoly, ktoré sa najčastejšie používajú v sieťach MANET, ako napríklad AODV. Posledná časť práce sa venuje analýze výsledkov zo simulácií a ich porovnaniu.

II. MOBILNÉ AD HOC SIETE

Mobilne ad hoc siete (MANET) sú siete založené na potrebe, čo znamená, že sa môžu vytvárať podľa potreby a aj ukončiť. Môžu vzniknúť takmer kdekoľvek v osadách a továrňach, ale aj vo vlaku, či uprostred lesa. Pre zariadenia tvoriace tieto siete je charakteristické, že sú rôznorodé, takže smartfón alebo tablet môže byť súčasťou siete MANET rovnako ako notebook. Dôležitým kritériom z pohľadu zariadení je podpora bezdrôtovej komunikácie, keďže tieto zariadenia sú navzájom prepojené týmto spôsobom, teda vytvárajú medzi sebou priame spojenie. Je dôležité spomenúť, že použitie bezdrôtovej komunikácie medzi mobilnými zariadeniami v sieti nevyžaduje vopred vybudovanú infraštruktúru, čo znamená, že siete MANET je možné vytvárať a používať aj bez káblvej infraštruktúry. [2]

Sieť MANET môže byť pripojená k internetu, napríklad prostredníctvom prístupového bodu. V tomto prípade prístupový bod funguje ako brána medzi sieťami. Komunikácia medzi zariadeniami, ktoré tvoria sieť, môže prebiehať pomocou rôznych bezdrôtových technológií (Wi-Fi, Bluetooth etc.). Nové zariadenia sa môžu k týmto sieťam ľubovoľne pripojiť a tie, ktoré sa v nich už nachádzajú, môžu odísť.

Uzly siete MANET sú mobilné, čo znamená, že môžu ľubovoľne meniť svoju geografickú polohu. To znamená, že topológia siete je dynamická a nie je možné predvídať zmeny v topológii.

V dôsledku dynamicky sa vyskytujúcich zmien v topológii je smerovanie zložitá a stratégie smerovania používané v tradičných káblových infraštruktúrach sa nedajú použiť. Okrem náhodného voľného pohybu uzlov, proces smerovania je ďalej komplikovaný obmedzeným vysielačím výkonom uzlov. To znamená, že uzly siete nie sú nevyhnutne schopné nadviazať priamu komunikáciu. Vo všeobecnosti je komunikácia v týchto sieťach multi-hopová, takže dátové pakety odoslané zdrojovým uzlom sa dostanú do cieľa pomocou medziľahlých uzlov.

III. SMEROVACIE PROTOKOLY V MANET SIEŤACH

Úlohou smerovacích protokolov je zmapovať cesty medzi odosielačimi a prijímačimi stranami zúčastňujúcimi sa na komunikáciu. Výsledkom toho je, že uzol, ktorý chce poslať dátové pakety, pozná po ukončení procesu smerovania cestu k cieľovej stanici, na ktorú môže poslať požadované pakety. Komunikácia v sieti MANET môže prebiehať jednoskokovým alebo viacskokovým spôsobom. Ako sa pakety dostanú do cieľa závisí od vzdialenosti medzi komunikujúcimi stranami. Proces, počas ktorého sa rozhodujeme, ktorý medziľahlý uzol budeme využívať je úlohou daného smerovacieho protokolu. V ad hoc sieťach rozlišujeme tri typy smerovacích stratégií. Sú to na požiadanie, známe aj ako reaktívne, tabuľkové alebo proaktívne a hybridné metódy smerovania. [3]

A. Proaktívne smerovanie

Proaktívne smerovacie protokoly sa často označujú ako smerovacie protokoly riadené tabuľkou. Pri tomto type stratégie smerovania má každý uzol v ad hoc sieti takzvanú smerovaciu tabuľku. Tieto tabuľky obsahujú rôzne informácie o všetkých uzloch v sieti, ako je poloha uzlov a cesty k uzlom. Týmto spôsobom, ak chcú spolu komunikovať dva uzly, nie je potrebné pred samotným prenosom dát vykonávať zisťovanie cesty, ako v prípade reaktívnych smerovacích stratégií, takže komunikácia medzi odosielačom a príjemcom môže byť rýchlejšia. Medzi proaktívne smerovacie protokoly patria Destination Sequenced Distance Vector (DSDV) protokol a Greedy Perimeter Stateless Routing (GPSR) protokol. [4], [5]

B. Reaktívne smerovanie

Protokoly reaktívneho smerovania sú najčastejšie zvolené smerovacie stratégie v sieťach MANET. Tieto typy protokolov sú navrhnuté tak, aby eliminovali alebo aspoň znížili hlavný problém proaktívnych smerovacích algoritmov, ktorým je sieťová zaťaženie spojená s aktualizáciou smerovacej tabuľky uzlov vo veľkých sieťach. V prípade, keď chcú dva uzla komunikovať medzi sebou odosielača strana skontroluje svoju tabuľku smerovania, aby zistila, či obsahuje cestu k prijímajúcej strane. Ak zdrojový uzol nemá platnú cestu do cieľa, spustí takzvaný proces zisťovania cesty. Na konci procesu, ak existuje aspoň jedna platná cesta medzi cieľovým a zdrojovým uzlom, bola objavená, takže zdrojový uzol môže začať komunikáciu. Hlavnou výhodou reaktívnych smerovacích protokolov je, že relatívne menej zaťažujú sieť, aj keď pozostáva z veľkého počtu uzlov. Na druhej strane nevýhodou je, že uzly pred začatím komunikácie nepoznajú cesty vedúce k cieľovej stanici, preto ich treba odhaliť pred odoslaním dátových paketov, čo má za následok oneskorenie komunikácie. Niektorými príkladmi najpopulárnejších reaktívnych smerovacích protokolov sú Ad hoc On-Demand Distance Vector (AODV) protokol a Dynamic source routing protocol (DSR). [4], [8]

IV. STROJOVÉ UČENIE PRE SMEROVANIE V MANET SIEŤACH

Strojové učenie je odvetvie počítačovej vedy a umelej inteligencie, ktorého cieľom je napodobniť proces ľudského učenia analýzou prostredia, aby sa stroje mohli učiť nezávisle. Proces strojového učenia sa dá rozdeliť na dve hlavné časti, ktoré sú tréningovanie a testovanie. V rámci tréningového procesu získava strojové učenie vzory pre vstupné dáta a požadované výstupné výsledky. Po analýze týchto informácií sa strojové učenie dostane do fázy učenia, kde dokáže filtrovať prepojenia zo vstupných údajov a zodpovedajúcich výstupných výsledkov a zovšeobecňovať ich. Týmto spôsobom, pomocou spojení získaných vo fáze učenia, dokáže spracovať úplne nový súbor údajov, analyzovať ho a produkovať požadované výstupné hodnoty, aj keď sa s novo prichádzajúcimi údajmi predtým nestretol. Vo všeobecnosti proces učenia nie je obmedzený. Takže počas prevádzky algoritmu strojového učenia môže niekoľkokrát prejsť procesom učenia s rôznymi tréningovými údajmi, čím sa zabezpečí vývoj presnosti algoritmu. Keďže existuje nespočetné množstvo algoritmov strojového učenia, v záujme transparentnosti ich môžeme zoskupiť podľa ich vlastností a fungovania. [11], [12]

A. AODV K-means smerovací algoritmus pre siete MANET

V [13] autori navrhli vylepšenie protokolu AODV pomocou algoritmu strojového učenia K-means. Algoritmus K-means usporiada susedné uzly daného uzla do klastrov, kam môžu byť pakety-požiadavky pri hľadaní cesty (RREQ) vysielané. Aby sa znížil počet vysielaných paketov RREQ, vysielajú sa iba do jedného klastra. Pakety RREQ sa vysielajú do klastra, ktorý je najbližšie k cieľovej stanici. Keď uzol prijme paket RREQ, skontroluje, či je to cieľ a či pozná cestu, ktorá k nemu vedie. Ak nie je splnená žiadna z podmienok, daný uzol pokračuje vo vysielaní paketu RREQ. Tento proces sa opakuje, kým paket RREQ nedosiahne uzol, ktorý pozná cestu k cieľu, alebo ho neprijme samotný cieľový uzol. V tomto prípade je proces vyhľadávania cesty pozastavený a začne sa proces posielania samotných dát. [13]

B. AntHocNet smerovací algoritmus pre siete MANET

Ďalším algoritmom, ktorý využíva algoritmy strojového učenia, je AntHocNet. AntHocNet rozdeľuje proces smerovania na tri hlavné časti - zisťovanie cesty, údržba a testovanie cesty a správa dátových paketov. Proces zisťovanie cesty je podobný procesu zisťovania cesty protokolu AODV. Počas tohto procesu odosielateľ odošle, takzvaný, forward ant paket do všetkých susedných uzlov. Tento paket sa šíri v rámci siete rovnakým spôsobom ako paket RREQ v prípade protokolu AODV. Potom, čo cieľ alebo uzol dostane paket forward ant, pošle, takzvaný, backward ant paket. Tento paket znovu sleduje svoju cestu, kým nedosiahne zdrojový uzol, podobne ako RREP pre protokol AODV. Ak backward ant nemôže sledovať cestu k zdroju, napríklad v prípade, že sa jeden z uzlov cesty vzdiali v dôsledku pohybu uzlov, čím sa preruší cesta, paket backward ant zahodí sa. Ak je nájdenie cesty neúspešné a žiadny backward ant nedosiahne zdroj po určitom čase, postup zisťovania cesty sa môže opakovať niekoľkokrát.

Údržba cesty sa prejaví po detekcii cesty a trvá počas trvania komunikácie medzi dvoma koncovými zariadeniami. Počas tohto procesu sa algoritmus pokúša aktualizovať a rozšíriť už získané informácie o smerovaní. Potom sa začína proces správy dátových paketov, v ktorom pomocou použitia algoritmu ACO (ant colony optimization) sa vyberá najkratšia a nepreťažená cesta od zdrojového uzla k cieľovému. [14], [15]

V. ANALÝZA VÝSLEDKOV SIMULÁCIÍ

Cieľom práce bolo preskúmať smerovacie procesy v sieťach MANET, ktoré zabezpečujú rôzne smerovacie protokoly a algoritmy strojového učenia. Podľa toho boli vybraté štyri algoritmy AODV, GPSR, aodv_Kmeans a AntHocNet. Smerovanie bolo skúmané pomocou simulátora NS3 [17].

Vzhľadom na dynamické zmeny topológie pre siete MANET bolo potrebné vykonať simuláciu niekoľkokrát s rovnakými nastaveniami, preto každá simulácia je opakovaná 20-krát a potom sú vypočítané priemerne hodnoty. Bolo vykonaných 300 simulácií s každým algoritmom pri 15 rôznych nastaveniach parametrov, čo znamená, že celkovo bolo vykonaných a analyzovaných 1200 simulácií. Výkonnosť smerovacích algoritmov v sieťach MANET bola hodnotená na základe niekoľkých kritérií – počet úspešne doručených dátových paketov a oneskorenia. Smerovací algoritmus je tým efektívnejší, čím viac dátových paketov dokáže doručiť na miesto určenia s malým oneskorením. Tabuľka 1 zobrazuje zoznam parametrov simulácií, ktoré boli používané.

Tabuľka 1
Parametre simulácií

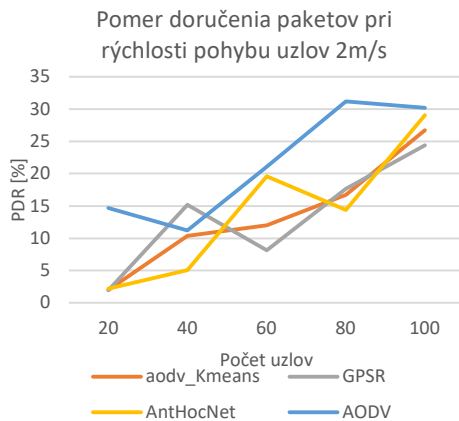
Protokol	AODV, GPSR, aodv_Kmeans, AntHocNet
Počet uzlov	20, 40, 60, 80, 100
Plocha	500m * 500m
Rýchlosť	2,5 m/s
Rádiový dosah	50 m
Čas simulácia	100 sekúnd
Veľkosť paketu	1024 byte
Počet vysielaných paketov	100

A. Pomer doručenia paketov

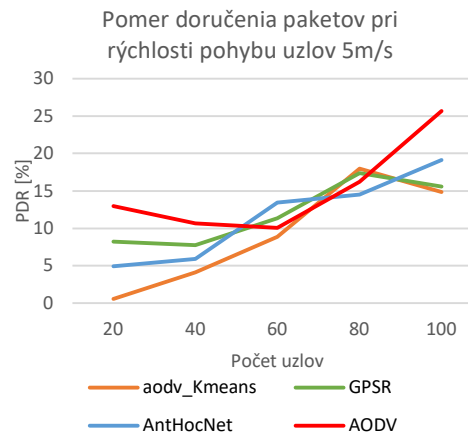
PDR alebo Packet Delivery Ratio udáva počet úspešne prijatých dátových paketov v percentách. PDR teda určuje pomer dátových paketov odoslaných odosielajúcim uzlom a prijatých prijímacím uzlom. Túto hodnotu môžeme vypočítať na základe nasledujúceho vzorca:

$$PDR(\%) = \frac{\text{Počet Prijatých Paketov}}{\text{Počet Odoslaných Paketov}} * 100$$

Na obrázkoch 1, 2 môžeme vidieť hodnoty PDR produkované GPSR, AODV, aodv_Kmeans a AntHocNet s rastúcim počtom uzlov a rýchlosťou 2 a 5 m/s.



Obrázok 1
Pomer doručenia paketov pri rýchlosti pohybu uzlov 2m/s



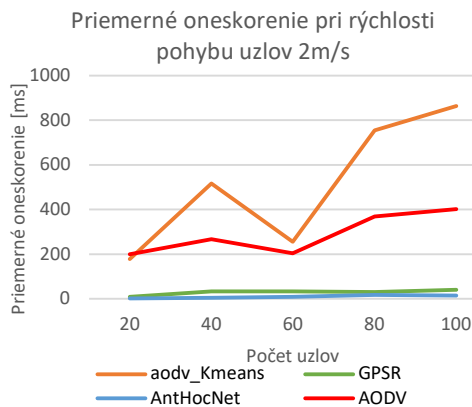
Obrázok 2
Pomer doručenia paketov pri rýchlosti pohybu uzlov 5m/s

B. Priemerné oneskorenie

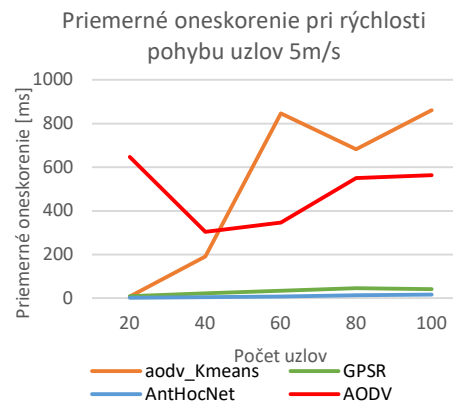
Okrem PDR je účinnosť smerovacieho algoritmu určená časom, ktorý uplynie medzi odoslaním a prijatím dátových paketov. Latencia medzi koncovými zariadeniami môže byť spôsobená mnohými faktormi. Z hľadiska efektívnosti smerovacieho algoritmu je výhodnejšie menšie oneskorenie, pretože odoslané dátové pakety prijme cieľová stanica skôr. Priemerné oneskorenie medzi koncovými zariadeniami možno vypočítať na základe nasledujúceho vzorca:

$$\text{Average End to End Delay}(ms) = \frac{\sum \text{End to End Delay}}{\text{Počet Prijatých Paketov}} * 1000$$

Na obrázkoch 3, 4 môžeme vidieť priemerné oneskorenie produkované GPSR, AODV, aodv_Kmeans a AntHocNet s rastúcim počtom uzlov a rýchlosťou.



Obrázok 3
Priemerné oneskorenie pri rýchlosti pohybu uzlov 2m/s



Obrázok 4
Priemerné oneskorenie pri rýchlosti pohybu uzlov 5m/s

VI. ZÁVER

Porovnaním týchto výsledkov sme dospeli k nasledujúcemu záveru. Medzi algoritmi, ktoré sme skúmali, aodv_Kmeans dosahuje najvyššie hodnoty PDR, podobné ako AntHocNet alebo GPSR pri malých rýchlostiach pohybu uzlov, ale pri zvyšovaní rýchlosti uzlov, dosahuje nižšiu hodnotu PDR. S veľkým počtom uzlov však dosiahol zhruba rovnaké hodnoty, ako AntHocNet alebo GPSR. Pri použití tohto algoritmu však bola latencia medzi koncovými zariadeniami oveľa vyššia, čo sa zvyšovalo so zvyšujúcim sa počtom uzlov. Algoritmy AntHocNet a GPSR mali podobné hodnoty PDR pri nízkych rýchlostiach, takže rozdiel medzi pomerom odoslaných a prijatých dátových paketov pri rýchlostiach 2 a 5 m/s nebol významný. Keď sa však uzly pohybovali rýchlosťou 10 m/s, algoritmus GPSR bol schopný dosiahnuť lepšiu hodnotu PDR aj pri nízkom počte uzlov, ktoré AntHocNet dokázal dosiahnuť len vtedy, keď počet uzlov bol zvýšený. Protokol AODV dokázal dosiahnuť lepšie hodnoty PDR ako iné algoritmy, ktoré sme skúmali, najmä pri použití nízkeho a vysokého počtu uzlov. End-to-end oneskorenie však bolo

dosť vysoké. Vo všetkých prípadoch bolo end-to-end oneskorenie najpriaznivejšie pri používaní AntHocNet. Algoritmus GPSR bol schopný doručiť dátové pakety na miesto určenia za cenu o niečo dlhšieho oneskorenia. Keď sa počet uzlov zvýšil, latencia GPSR aj AntHocNet sa zvýšila. Stručne povedané, aodv_Kmeans dokázal doručiť dátové pakety do cieľa s oveľa dlhším oneskorením a to vo väčšine scenárov, ktoré sme skúmali, dosiahol nižšie hodnoty PDR v porovnaní s algoritmi AODV, GPSR a AntHocNet. Hoci hodnoty PDR protokolu AODV boli lepšie, ako hodnoty iných algoritmov, vo väčšine scenárov, ktoré sme skúmali, ale oneskorenie medzi koncovými zariadeniami bolo relatívne vysoké. Algoritmy GPSR a AntHocNet v niektorých prípadoch dosiahli hodnoty PDR podobné ako AODV, ale vo väčšine prípadov bol pozorovaný opak. Avšak v prípade týchto dvoch algoritmov bolo oneskorenie medzi koncovými zariadeniami oveľa menšie. Stručne povedané, pre dosiahnutie čo najlepšieho PDR sa odporúča použitie protokolu AODV, ale v tomto prípade však musíme počítať s väčším oneskorením. V prípade, keď potrebujeme čo najnižšie oneskorenie, môžu byť ideálne algoritmy AntHocNet a GPSR. Pri nízkej rýchlosti bolo použitie AntHocNet priaznivejšie. GPSR dosiahol podobné hodnoty PDR pri nízkej rýchlosti, ale zistilo sa, že oneskorenie AntHocNet je lepšie. Pri vysokej rýchlosti je však aplikácia GPSR tou správnou voľbou, keďže aj pri nízkom počte uzlov dokázala dosiahnuť približne rovnakú hodnotu PDR ako AntHocNet.

REFERENCES

- [1] M. Conti, S. Giordano, "Mobile ad hoc networking: milestones, challenges, and new research directions," in *IEEE Communications Magazine*, vol. 52, no. 1, pp. 85-96, 2014.
- [2] F. Abdel-Fattah, K. A. Farhan, F. H. Al-Tarawneh, F. AlTamimi, "Security Challenges and Attacks in Dynamic Mobile Ad Hoc Networks MANETs," 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), pp. 28-33, 2019.
- [3] A. R. Rajeswari, "A mobile ad hoc network routing protocols: A comparative study." *Recent trends in communication networks*, pp. 1-24, 2020.
- [4] D. N. Patel, S. B. Patel, H. R. Kothadiya, P. D. Jethwa, R. H. Jhaveri, "A survey of reactive routing protocols in MANET." *International Conference on Information Communication and Embedded Systems*, pp. 1-6, 2014.
- [5] H. Kaur, V. Sahni, Dr. Manju Bala. "A survey of reactive, proactive and hybrid routing protocols in MANET: a review." *network*, pp. 498-500, 2013.
- [6] A. Silva, N. Reza and A. Oliveira. "Improvement and Performance Evaluation of GPSR-Based Routing Techniques for Vehicular Ad Hoc Networks." *IEEE Access*, pp. 21722-21733, 2019.
- [7] B. Karp, H.T. Kung, "Greedy Perimeter Stateless Routing for Wireless Networks." *Proceedings of the Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pp. 243-254, 2000.
- [8] P. Khatri, M. Rajput, A. Shastri, K. Solanki, "Performance study of ad-hoc reactive routing protocols." *Journal of Computer Science*, pp. 1159-1163, 2010.
- [9] H. Moudni, M. Er-rouidi, H. Mouncif, B. El Hadadi, "Performance analysis of AODV routing protocol in MANET under the influence of routing attacks." *2016 International Conference on Electrical and Information Technologies* pp. 536-542, 2016.
- [10] I. D. Chakeres, E. M. Belding-Royer. "AODV routing protocol implementation design." *24th International Conference on Distributed Computing Systems Workshops*, pp. 698-703, 2004.
- [11] El Naqa I., M.J. Murphy, "What Is Machine Learning?" *Machine Learning in Radiation Oncology: Theory and Applications*, pp. 3-11, 2013
- [12] A. Forster, "Machine Learning Techniques Applied to Wireless Ad-Hoc Networks: Guide and Survey." *2007 3rd International Conference on Intelligent Sensors, Sensor Networks and Information*, pp. 365-370, 2007.
- [13] K.P. Rao, T. S. Murugan, "Energy efficient clustering technique using k-means and AODV-ACO Routing with Secured AES Cryptography in MANET." *International Journal of Intelligent Engineering and Systems*, pp. 292-302, 2018.
- [14] Di Caro, Gianni a Ducatelle, Frederick a Gambardella, Luca Maria. „AntHocNet: an adaptive nature-inspired algorithm for routing in mobile ad hoc networks.“ *European transactions on telecommunications*, 2005: 443—455
- [15] N. Shah, H. El-Ocla, P. Shah, "Adaptive Routing Protocol in Mobile Ad-Hoc Networks Using Genetic Algorithm," in *IEEE Access*, vol. 10, pp. 132949-132964, 2022.
- [16] Ducatelle, Frederick. „Adaptive routing in ad hoc wireless multi-hop networks.“ 2007: 57-79.
- [17] NS-3 Network Simulator Manual, NS-3 Consortium, 2021.

Tvorba datasetov pre strojové učenie v MANET

¹Daniel ĎURČA, ²Natalia KURKINA, ³Ján PAPAĽ

^{1,2,3} Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹daniel.durca@student.tuke.sk, ²natalia.kurkina@tuke.sk, ³jan.papaj@tuke.sk

Abstract — V dnešnej dobe, kedy sa mobilné zariadenia stávajú stále viac rozšírenými a súvisiace technológie stále vylepšujú, mobilné ad-hoc siete sa stávajú dôležitým nástrojom pre komunikáciu a prenos dát. Cieľom tejto práce je preto analyzovať aktuálny stav datasetov pre strojové učenie využiteľných pre mobilné siete a analyzovať existujúce datasety pre MANET. Pre tento účel sme sa zamerali na publikované štúdie a konferencie, ktoré sa zaoberajú touto tematikou, zamerali sme sa taktiež na konkrétne publikované datasety a porovnali sme rôzne metódy tvorby datasetov.

KLúčové slová — tvorba datasetov, strojové učenie, mobile ad-hoc network, MANET, syntetické datasety, Python.

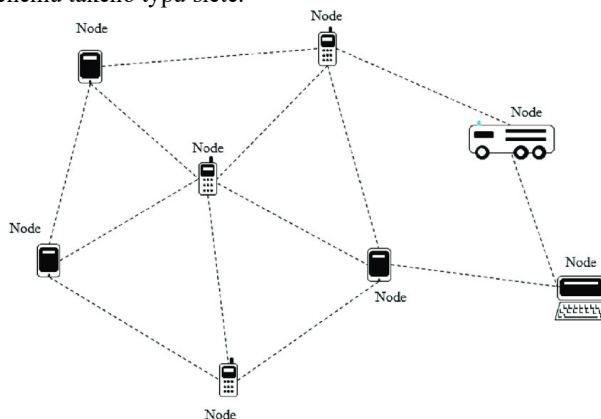
I. ÚVOD

Ad-hoc siete sú bezdrôtové siete, ktoré sa skladajú z uzlov, ktoré je možné vytvoriť na rôznych miestach bez potreby pevnej infraštruktúry. Pozostávajú z mobilných zariadení, ktoré disponujú bezdrôtovými technológiami a môžu medzi sebou komunikovať a vytvoriť sieť bez použitia káblov alebo iných fyzických prepojení. [1] Každý uzol v takejto topológii môže byť súčasne cieľovým alebo zdrojovým, a taktiež môže byť súčasťou prenosu dát medzi inými uzlami.

Každý uzol je v topológii označovaný identifikátormi a jeho pozícia sa zaznamenáva pomocou jeho neighbors – susedných uzlov. Susedné uzly sú uzly, ktoré sú priamo spojené s daným uzlom. Každý uzol tak môže byť susedným uzlom, ale aj susedom susedného uzla, čo umožňuje vytvárať viacvrstvové siete [1].

Ad-hoc siete sa členia do niekoľkých skupín, ktoré sa líšia podľa využitia a ich vlastností. Medzi najznámejšie z nich patria - Wireless Mesh Network (WMN), Wireless Sensor Network (WSN), Mobile Ad Hoc Network (MANET) [2].

Mobilné ad-hoc siete (MANET) sú siete, ktoré sa skladajú z mobilných uzlov, ktoré sa môžu pohybovať nezávisle a môžu sa vzájomne komunikovať bez použitia pevnej infraštruktúry. Siete MANET sú veľmi flexibilné a využívané v rôznych odvetviach, či už medicíne, zdravotníctve alebo v doprave [3]. Tieto siete sú založené na princípe samostatnosti uzlov a ich schopnosti komunikovať s inými uzlami v sieti bez nutnosti centrálnej koordinácie. Na obrázku 1 môžeme vidieť štandardnú schému takého typu siete.



Obr. 1 Príklad štandardnej schémy MANET siete [2]

Ďalším plusom je, že daný typ siete môže byť rýchlo nasadený v prípade výpadku alebo zrušený takmer okamžite v prípade potreby. Medzi nevýhody však patria časte prerušenia spojenia medzi zariadeniami a bezpečnostné problémy spojené so zabezpečením prenosu a obmedzená kapacita [4].

II. DATASETY PRE STROJOVÉ UČENIE V SIEŤACH MANET

V sieťach MANET strojové učenie zohráva dôležitú úlohu pri riešení rôznych problémov. Preto je dôležité mať k dispozícii kvalitné datasey, ktoré umožnia tréning a testovanie modelov strojového učenia.

A. MAWI Working Group Traffic Archive

MAWI dataset, patrí k jedným z používaných datasetov pre mobilne siete a obsahuje informácie o prevádzke internetových sietí. Jedná sa o reálny dataset, teda bol vytvorený pomocou zberu dát z rôznych internetových sietí [5].

Tento dataset obsahuje údaje o rôznych typoch prevádzky. Dáta sú zaznamenávané v reálnom čase na rôznych miestach a sú ukladané v podobe paketov, s informáciami o zdrojovej, cieľovej adrese, typu prevádzky a ďalšími.

MAWI dataset je dôležitým zdrojom informácií pre vedcov a výskumníkov v oblasti sieťovej bezpečnosti a analýzy prevádzky. Dáta z tohto datasetu sa používajú na vývoj nových nástrojov na detekciu a ochranu pred útokmi, ako aj na analýzu prevádzky a identifikáciu nových trendov v sieťovej prevádzke. Dáta v tomto datasete sú aktualizované denne a dostupné na v [6].

B. CIC DoS dataset 2017

CIC DoS dataset 2017 je súbor dát, ktorý obsahuje informácie o útokoch typu DoS na sieťové zariadenia [7]. CIC DoS dataset 2017 obsahuje viac ako 12 000 rôznych súborov dát, ktoré sa líšia veľkosťou a typom útoku. Príprava tohto datasetu pomocou zachytávania útokov sa zameriavala na univerzálne typy aplikovaných DoS útokov s pomalou rýchlosťou, ktoré sa často vyskytujú v dvoch variantoch: pomalé odosielanie a pomalé čítanie [7].

CIC DoS dataset 2017 je významným zdrojom informácií pre výskumníkov v oblasti sieťovej bezpečnosti a analýzy prevádzky. Dáta z tohto datasetu sa používajú na vývoj nových nástrojov na detekciu a ochranu pred útokmi typu DoS.

C. CRAWDAD/DARTMOUTH OUTDOOR Dataset

Princíp OUTDOOR datasetu od CRAWDAD DARTMOUTH spočíva v zbere dát o pohybe uzlov v mobilných ad-hoc sieťach v rôznych vonkajších prostrediach. Tento dataset obsahuje informácie o topológii siete, pohybe uzlov, rýchlosti uzlov, vzdialenosti a mnohých ďalších parametroch [8].

Počas experimentov boli uzly v sieti v pohybe a komunikovali medzi sebou pomocou rôznych smerovacích a prístupových protokolov. Výsledkom tohto procesu boli dáta, ktoré boli následne spracované a uložené v OUTDOOR datasete. Tieto dáta môžu byť použité na tréning a testovanie modelov strojového učenia, ktoré sa zameriavajú na optimalizáciu využitia zdrojov a minimalizáciu oneskorenia siete vo vonkajších prostrediach.

Medzi používanými protokolmi v tomto datasete boli - AODV (Ad-hoc On-demand Distance Vector) [4], APRL (Adaptive Probabilistic Routing Protocol with Location) [9], ODMRP (On-Demand Multicast Routing Protocol) [10], STARA (Self-organizing Time Division Multiple Access Resource Allocation) [11].

Využívanie rôznych protokolov v datasete umožňuje testovať a porovnávať výkon rôznych protokolov v reálnych vonkajších prostrediach. Tieto dáta môžu byť následne využité na vylepšenie smerovacích a prístupových protokolov.

D. MANIAC (The Mobile Ad Hoc Networks Interoperability And Cooperation) Challenge

V rámci MANIAC Challenge [12] je k dispozícii široká škála dátových súborov, ktoré obsahujú informácie o rôznych aspektoch mobilných ad hoc sietí. Tieto súbory zahŕňajú informácie o topológii siete, dátové prenosy a rôzne typy zlyhaní, ktoré sa môžu vyskytnúť v mobilných ad hoc sieťach [12].

V MANIAC Challenge sa kladie veľký dôraz na dôkladnú analýzu dát a využitie vedeckých metód pri riešení problémov v oblasti mobilných ad hoc sietí. Tento dataset bol vytvorený s cieľom podporiť výskum a inovácie v oblasti mobilných ad hoc sietí a umožniť výskumníkom a inžinierom testovať a vyhodnocovať svoje riešenia v reálnych podmienkach. Využitie dátových súborov z MANIAC Challenge môže pomôcť výskumníkom a inžinierom v oblasti mobilných ad hoc sietí pri vývoji nových technológií a zlepšení existujúcich riešení [13].

V súťaži MANIAC Challenge sa zúčastnili rôzne univerzity a výskumné inštitúcie z celého sveta. Jednou z týchto univerzít bola aj Technická univerzita v Košiciach [14]. Tímy z týchto inštitúcií mali možnosť využiť dataset z MANIAC Challenge pri vývoji svojich riešení a testovať ich v reálnych podmienkach.

III. METÓDY TVORBY DATASETOV

A. Metódy tvorby reálnych datasetov

V MANET (Mobile Ad hoc Network) sa tradične používajú rôzne metódy tvorby datasetov. Jednou z nich je meranie parametrov v reálnom čase, ktoré sa následne zaznamenávajú. Tieto parametre môžu byť napríklad výkon siete, priepustnosť, oneskorenie a podobne.

Metódy tvorby reálnych datasetov majú výhody aj nevýhody. Medzi výhody patrí, že tieto dáta sú získané v reálnych podmienkach a sú preto veľmi presné. Ďalšou výhodou je, že tieto dáta sú veľmi podrobné a obsahujú mnoho parametrov. To umožňuje, aby boli tieto dáta použité na veľa rôznych výskumných úloh.

Nevýhodou reálnych datasetov je, že sú veľmi obmedzené. Keďže sú tieto dáta zaznamenávané len v určitom konkrétnom čase. Najčastejšie parametre, ktoré sa zaznamenávajú v MANET, sú napríklad Packet Delivery Ratio (PDR) – pomer doručených paketov k počtu odoslaných, End to End Delay (E2ED) – oneskorenie odoslania a doručovania dát od zdrojového uzla k cieľovému a Selfish Value – miera dôveryhodnosti uzla, a mnoho ďalších [15].

Vzhľadom na obmedzenia metód tvorby datasetov v MANET, sa v poslednej dobe stále viac používajú simulácie. Simulácie tak naopak môžu pomôcť zaznamenať rôzne scenáre, ktoré sa môžu pri meraní zobrazíť len za určitých podmienok a môže byť zložitejšie na zachytenie tradičným spôsobom. V tom prípade je lepšie využiť metódy tvorby syntetických datasetov, ktoré dokážu simulovať siete v rôznych časoch, za rôznych podmienok, a taktiež na rôznych miestach, čo môže zvýšiť reprezentatívnosť dát.

B. Metódy tvorby syntetických datasetov

Existuje niekoľko spôsobov tvorby syntetických datasetov v MANET. Jedným z najčastejších spôsobov je použitie generatívnych modelov, ktoré umožňujú tvorbu datasetov s rôznymi topológiami siete.

Ďalšou možnosťou tvorby syntetických datasetov v MANET je použitie simulátorov, ako je napríklad NS3 alebo OMNeT++. NS3 je rozšíreným simulačným nástrojom, ktorý umožňuje tvorbu syntetických datasetov s rôznymi parametrami, vrátane topológie siete, pohybu uzlov, kanálových podmienok a podobne [16]. Tieto parametre môžu byť v týchto simulačných softvéroch ľahko nastavené a umožňujú tvorbu syntetických datasetov, ktoré zodpovedajú reálnym podmienkam v MANET.

Výhodou syntetických datasetov je, že umožňujú rýchle a efektívne testovanie nových algoritmov v MANET. Taktiež umožňujú opakované testovanie s rôznymi parametrami, čo je často ťažké dosiahnuť s reálnymi datasetmi. Navyše, syntetické datasetové metódy umožňujú tvorbu datasetov s rôznymi topológiami siete.

Nevýhodou syntetických datasetov je, že nemusia úplne zodpovedať reálnym podmienkam v MANET. Napríklad, môžu sa vyskytnúť chyby v modelovaní pohybu uzlov. Taktiež, syntetické datasetové metódy môžu byť obmedzené vzhľadom na rozsah parametrov, ktoré môžu byť zahrnuté v datasete.

C. Porovnanie syntetických a reálnych datasetov pre strojové učenie v MANET

V tabuľke Tabuľka 1 vidíme porovnanie výhod a nevýhod syntetických a reálnych dát, ktoré boli opísané v predošlých kapitolách, a následne zosumarizované v tejto tabuľke.

Tabuľka 1 - Porovnanie využitia reálnych dát a syntetických pre strojové učenie

	Reálne dáta	Syntetické dáta
Výhody	- Vyššia presnosť	- Rýchlosť vytvárania
	- Realistickosť	- Vyššia rozmanitosť
	- Autenticita	- Škálovateľnosť
Nevýhody	- Vyššie náklady	- Nižšia autenticita
	- Obťažnosť získavania	- Nižšia presnosť
	- Menšia rozmanitosť	- Nižšia kvalita
	- Menšia reprezentatívnosť	

Využitie syntetických dát v ML tak závisí na konkrétnom prípade použitia a jeho cieľoch. Ak sú k dispozícii reálne dáta a sú dostatočne kvalitné a v dostatočnom množstve, mali by byť použité primárne ako tréningové. Avšak v našom prípade, kde sa jedná o datasety pre strojové učenie v MANET, existujúce reálne datasety sú v obmedzenom množstve a ich zber je finančne a časovo náročný, z dôvodu rýchleho vývoja, dynamiky MANET. Syntetické dáta sú pre nás tak

cenným nástrojom a v ďalších častiach tejto práce sa budeme venovať tvorbe týchto datasetov.

IV. GENEROVANIE SYNTETICKÝCH DÁT PRE MANET

A. Generovanie náhodných dát

Táto metóda spočíva v tom, že pre každý atribút datasetu sa vygenerujú náhodné hodnoty. Tieto hodnoty môžu byť viazané na určitý rozsah, môžu byť závislé od iných atribútov a môžu obsahovať šum. Táto metóda je často používaná pri generovaní simulovaných dát pre účely testovania a vývoja. Výhodou tejto metódy je, že umožňuje generovanie veľkého množstva dát v krátkom čase a umožňuje testovať modely na rôznych scenároch.

Parametre pri tvorbe datasetov boli volené tak, aby dosahovali, čo najvyššiu realističnosť modelu pri ďalšom využití v strojovom učení. V tabuľke Tabuľka 2 sú popísané dátové typy parametrov s ich rozsahmi.

Tabuľka 2 - Tabuľka parametrov s dátovými typmi pre náhodne generovanie datasetov

Dátový typ	Parameter (názov premennej/poľa)	Hodnoty v premennej/poľi
int	Veľkosť siete (network_size)	100000
object	Typ uzla (node_types)	originating, intermediate, target
object	Dátový typ (data_types)	text, image, audio, video
object	Typ komunikácie (communication_types)	unicast, broadcast, multicast
float64	Strata paketov (packet_loss)	0.001, 0.005, 0.01, 0.02, 0.1
int64	Konektivita (connectivity)	Hodnota 0 alebo 1
float64	Prenosová rýchlosť (transfer_speed)	Hodnota od 2.4kbps – 20Gbps ((2.4/1000) – 20000)
float64	Pozícia uzla (node_positions_x, node_positions_y)	0 – 10000 v rozsahu network_size
int64	Identifikátor uzla (node_id)	unikátne číslo každého uzla v rozsahu network_size
int64	Využitý kanál (channel_usage)	Hodnota od 1-13
float64	Frekvencia prenosu (frequency)	30 – 5000 v rozsahu network_size
float64	Oneskorenie (delay)	Hodnota od 0.001 - 0.1
object	Protokol (protocols_used)	OLSR, DSDV, WRP, ZRP, AODV, DSR, TORA

Pomocou generovania náhodných dát boli vytvorené tri datasety. Pre porovnanie kvality datasetov jednotlivé datasety boli otestované na modeloch strojového učenia a porovnané navzájom ich presnosť a výpočtové časy. Ale pred tým je potrebné dáta transformovať a normalizovať. Na transformáciu dát bol použitý mechanizmus One-hot encoding [17]. Na normalizáciu dát bol použitý Min-Max mechanizmus [18]. Ďalej dataset bol rozdelený na testovaciu a tréningovú časť v pomere 80:20. Na porovnanie boli vybrané nasledujúce modely strojového učenia - k-NN (k-najbližších susedov), Náhodné lesy – Random Forest, SVM – Support Vector Machines. Na tabuľke 3 môžeme vidieť výsledky pre tri rôzne datasety a tri rôzne algoritmy strojového učenia. Z tabuľky môžeme vyčítať, že najvyššie hodnoty pri takmer všetkých algoritmoch má Dataset 1, takže ho môžeme pri tejto metóde generovania datasetov náhodnými hodnotami určiť ako najúspešnejší.

Tabuľka 3 - Porovnanie výpočtových časov a presnosti modelov strojového učenia - náhodne generované datasety

	Dataset 1	Dataset 2	Dataset 3
k-NN presnosť (%)	49.94	49.89	49.98
Výpočtový čas k-NN (s)	8.2	5.8	4.3
SVM presnosť (%)	50.22	50.29	50.00
Výpočtový čas SVM (s)	32.5	34.5	31.8
Náhodné stromy presnosť (%)	50.52	50.24	49.45
Výpočtový čas Náhodné stromy (s)	402	414.4	326.8

B. Generovanie syntetických dát pomocou simulácie pohybu uzlov

Generovanie kvalitných datasetov nám môže zabezpečiť napríklad simulácia pohybu uzlov v sieti. Tento spôsob generovania syntetických dát má mnoho výhod. Jednou z najväčších výhod je, že umožňuje vytvárať siete s rôznymi topológiou a štruktúrami, ktoré môžu byť použité na testovanie algoritmov a modelov. Taktiež umožňuje vytvárať siete s rôznymi charakteristikami, ktoré môžu byť použité na testovanie algoritmov a modelov. Jednou z najpoužívanejších knižníc na generovanie a manipuláciu s grafmi v programovacom jazyku Python je NetworkX [19]. Táto knižnica obsahuje množstvo funkcie na vytvorenie uzlov v sieti, ich spojení a ich vizualizáciu. Pomocou NetworkX bola vytvorená sieť s topológiou o veľkosti 100 uzlov pre každý protokol a 100 odlišných spojení pre každý uzol, čo povolilo vytvoriť okolo 70000 riadkový dataset pri 7 smerovacích protokoloch. K každému uzlu boli pridane váhy, dôležitosť alebo kvalitu spojenia medzi jednotlivými uzlami, kde 100 je najkvalitnejšie spojenie a 1 - najmenej.

Premenné, ktoré budú zachytené alebo vygenerované pre jednotlivé siete sú: *weight_node* - váha uzla, *distance* - vzdialenosť od cieľového uzla, *channel_usage* - číslo použitého kanálu od 1 po 13 a *speed* - rýchlosť pohybu uzla.

Pre správnu implikáciu algoritmov strojového učenia bola vykonaná numerická normalizácia hodnôt a dáta boli rozdelené na tréningové a testovacie v pomere 80:20. Z tabuľky 4 môžeme všimnúť že Dataset 2 dosahuje najvyššie hodnoty presnosti v takmer všetkých meraniach, takže ho môžeme určiť za kvalitnejší.

Tabuľka 4 - Porovnanie výpočtových časov a presnosti modelov ML - datasety vytvárané simuláciou siete

	Dataset 1	Dataset 2	Dataset 3
k-NN presnosť (%)	99.02	99.09	99.04
Výpočtový čas k-NN (s)	2.4	1.2	0.6
SVM presnosť (%)	98.96	99.01	98.91
Výpočtový čas SVM (s)	7.2	9.8	7.1
Náhodné stromy presnosť (%)	100	100	100
Výpočtový čas Náhodné stromy (s)	2.4	2.1	2.1

C. Analýza a porovnanie výsledkov tvorby syntetických datasetov

Pri 3 náhodných datasetoch sa hodnoty pri skoro všetkých meraniach držali okolo hodnoty +-50% ako to môžeme vidieť z tabuľky 3. Avšak pri datasetoch tvorených pomocou metódy simulácie siete, tieto hodnoty takmer dosahovali úspešnosť 100%, čo je prívětivé pri modeloch strojového učenia čo môžeme vidieť z tabuľky 4.

Tabuľka 5 - Porovnanie hodnôt modelov ML oboch metód generovania datasetov

	Dataset 1 – náhodné generovanie	Dataset 2 – generovanie pomocou simulácie siete
Presnosť k-NN (%)	49.94	99.09
Presnosť Náhodne lesy (%)	50.52	100
Presnosť SVM (%)	50.22	99.01

V Tabuľka 5 môžeme vidieť porovnanie výsledkov oboch z datasetov pri jednotlivých algoritmoch strojového učenia. Na základe uvedených hodnôt môžeme vidieť, že druhý dataset dosahuje výrazne lepšie výsledky ako prvý dataset. Presnosti všetkých algoritmov pri druhom datasete sú takmer 100%, avšak pri prvom datasete sa pohybujú okolo hodnoty 50%.

Z toho vyplýva, že druhý dataset je výrazne kvalitnejší a aj vhodnejší pre použitie v rôznych aplikáciách. Avšak, je dôležité zdôrazniť, že aj prvý dataset môže byť užitočný v niektorých prípadoch alebo špecifických aplikáciách.

V. ZÁVER

Tato práca bola venovaná téme tvorby datasetov pre strojové učenie v mobilných ad-hoc sieťach.

Na základe výsledkov porovnania niekoľkých syntetických vytvorených datasetov môžeme uviesť, že dataset, ktorý bol vytvorený pomocou simulácie pohybu uzlov v sieti, je kvalitnejší a vhodnejší pre použitie v rôznych aplikáciách. Avšak, prvý dataset môže byť užitočný v niektorých špecifických aplikáciách. V každom prípade, tieto výsledky ukazujú, že je dôležité vykonávať viacero meraní a porovnávať výsledky, aby sme mohli získať správne informácie o

kvalite dát a presnosti modelov.

Ďalšie kroky pri práci v danej oblasti výskumu budú sa zamerat' na rozšírenie existujúcich datasetov a zlepšenie ich kvality. Ďalším smerom by mohlo byť využitie týchto datasetov v reálnych aplikáciách, aby sme mohli overiť ich účinnosť a presnosť.

POĎAKOVANIE

Tento výskum bol financovaný agetúrou na podporu výskumu a vývoja SR, výskumný grant č. APVV-17-0208 a VEGA 1/0260/23.

REFERENCIE

- [1] M. Barbeau a E. Kranakis, Principles of ad-hoc networking, John Wiley & Sons, 2007.
- [2] V. Student a R. Dhir, A study of ad hoc network: A review, Int. J, 2013.
- [3] R. Md. N. A. M. Saudi, M. A. Arshad, A. G. Buja and A. F. A. Fadzil, "Mobile Ad-Hoc Network (MANET) Routing Protocols: A Performance Assessment," Springer, Singapore, 28 Marec 2019.
- [4] J. Loo, J. Lloret Mauri, J. Hamilton Ortiz a G. Williams, Mobile Ad Hoc Networks: Current Status and Future Trends, Boca Raton: Taylor & Francis, 2011.
- [5] O. I. AL-SANJARY, Comparison and detection analysis of network traffic datasets using K-means clustering algorithm, Journal of Information & Knowledge Management, 2020.
- [6] M. W. Group, „MAWI Working Group Traffic Archive,“ [Online]. Available: <https://mawi.wide.ad.jp/mawi/>.
- [7] C. I. f. Cybersecurity, „CIC DoS dataset (2017),“ 2017. [Online]. Available: <https://www.unb.ca/cic/datasets/dos-dataset.html>.
- [8] D. Kotz, C. Newport, N. Dubrovsky, A. Fiske, J. Liu, C. Masone, S. McGrath a Y. Yuan, „CRAWDAD DARTMOUTH/OUTDOOR Dataset,“ 17 Október 2003. [Online]. Available: <https://iecc-dataport.org/open-access/crawdaddartmouthoutdoor>.
- [9] B. Karp a H. T. Kung, Dynamic Neighbor Discovery and Loop-Free, Multi-Hop Routing for Wireless, Mobile Networks, Cambridge: Harvard University, 1998.
- [10] S.-J. Lee, W. Su a M. Gerla, On-Demand Multicast Routing Protocol in Multihop Wireless Mobile Networks, Los Angeles, CA: Kluwer Academic Publishers, 2002.
- [11] P. Gupta a P. R. Kumar, A System and Traffic Dependent Adaptive Routing Algorithm for Ad Hoc Networks, San Diego, CaliforniaUSA: IEEE, 1997.
- [12] H. Will, F. Juraschek, M. Gunes a J. Schiller, Experiences from the MANIAC Challenge, Berlin, Germany: Freie Universität Berlin and Humboldt University Berlin, 2011.
- [13] L. A. DaSilva, A. B. MacKenzie, M. S. Thompson a E. Q. Baumann, The MANIAC Challenge: Educational Experiences in Ad Hoc Networking, IEEE CS, 2009.
- [14] I. Klimek a I. Sidimák, MANIAC: Mobile Ad-Hoc Networks Interoperability and Cooperation: THE LIVE AND LET LIVE STRATEGY, Košice, Slovensko: Dept. of Computers and Informatics, FEI TU of Košice, Slovensko, 2008.
- [15] S. Kanthimathi a J. R. Prathuri, Classification of Misbehaving nodes in MANETS using Machine Learning Techniques, Bangalore, India: IEEE, 2020.
- [16] A. R. Khan, S. M. Bilal a M. Othman, A Performance Comparison of Open Source Network Simulators for Wireless Networks, Madrid, Spain: IEEE, 2012.
- [17] J. Brownlee, Data preparation for machine learning: data cleaning, feature selection, and data transforms in Python., Jason Brownlee, 2020, pp. 4-5.
- [18] V. R. Joseph a A. Vakayil, SPlit: An Optimal Method for Data Splitting, Taylor & Francis Group, LLC, 2021.
- [19] A. Hagberg, P. Swart a D. Schult, Exploring network structure, dynamics, and function using NetworkX, Los Alamos, NM (USA): Los Alamos National Lab.(LANL), 2008.

Návrh konštrukčnej a elektrickej časti robotického ramena

¹Peter GIROVSKÝ, ²Patrik MOLNÁR

^{1,2} Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach

¹peter.girovsky@tuke.sk, ²patrik.molnar@tuke.sk

Abstrakt — Tento príspevok sa zaoberá návrhom konštrukcie a riadenia robotického ramena. V práci je predstavený návrh konštrukcie robotického ramena, ktoré bolo vytlačené pomocou 3D tlače a návrh elektrickej časti pre napájanie a riadenie krokových motorov, ktoré boli použité pre pohon jednotlivých osí robotického ramena. Riadiaca časť je postavená na technológii jednočipových procesorov radu ATMEL.

Kľúčové slová — robotické rameno, aktuátor, elektronika, riadenie, krokový motor

I. ÚVOD

Robotika predstavuje vednú disciplínu, ktorá sa zaoberá konštrukciou, riadením a následnou implementáciou robotov. Robotiku môžeme rozdeliť na aplikovanú a priemyselnú robotiku. Aplikovaná robotika sa zaoberá vývojom, konštrukciou a implementáciou robotov a priemyselná robotika sa prevažne zaoberá priemyselnými výrobnými procesmi robotizovanej prevádzky ako celku a konštrukcii priemyselných robotov. [1]

Robotov delíme do niekoľkých kategórií na:

- Manipulačné – slúžia na manipuláciu s predmetmi
- Mobilné – pohyblivé roboty pomocou kolies, pásov a pod.
- Informačné a riadiace – využívajú sa na zber dát, meranie, spracúvanie údajov
- Komplexné – vznikajú spojením minimálne dvoch vyššie spomenutých robotov

Z hľadiska vývoja môžeme robotov rozdeliť do troch základných generácií:

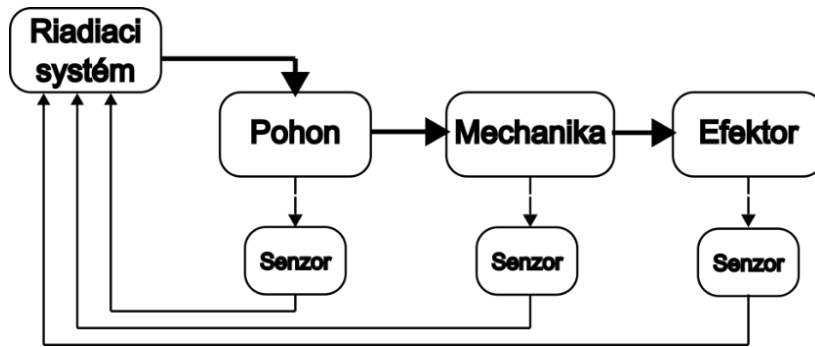
- 1. generácia – obsahuje vnútornú spätnú väzbu pevným riadiacim programom
- 2. generácia – je doplnená o interné a externé snímače, slúžiace na kontrolu procesov robota
- 3. generácia - do tejto generácie robotov patria roboty s vysokou mierou integrácie snímačov pre zber dát z okolitého prostredia, dokážu pracovať autonómne [2]

Vo všeobecnosti pod pojmom priemyselný robot rozumieme univerzálne využiteľný viacosý manipulátor, ktorého pohyb je voľne programovateľný.

Konštrukčné vlastnosti robotov sú dané ich kinematickou štruktúrou, teda spôsobom usporiadanie jeho kinematických dvojíc v reťazci. Najrozšírenejším konceptom je otvorený kinematický reťazec, ktorý obsahuje rotačné a translačné kinematické dvojice. Na základe tejto štruktúry môžeme rozdeliť roboty do štyroch základných skupín:

- Kartézske – TTT
- Cylindrické – RTT
- Sférické – RRT
- Angulárne – RRR

Riadenie robota zabezpečuje riadiaci systém generovaním signálov potrebných pre jednotlivé osi robota. Primárnou úlohou riadiaceho systému je generovanie signálov pre servopohony, preto pre zabezpečenie správneho riadenia je nevyhnutné, aby riadiaci systém dostával spätnú väzbu od snímačov.



Obr. 1 Principiálna bloková schéma riadenia.

Riadiaci systém robota zabezpečuje nasledujúce funkcie:

- Operačný systém ktorý podporuje multitasking
- Integrované prídavné funkcie
- Prerušenia
- Interpolácie trajektórie robota (osová, lineárna, kruhová)
- Riadenie osí robota
- Paralelná spolupráca robotov
- Spracovanie digitálnych a analógových vstupno/výstupných signálov

Riadenie pohybu priemyselných robotov môžeme rozdeliť na bodové riadenie a riadenie dráhy robota. Bodové riadenie sa využíva v prípade ak je potrebné dostať robota v pracovnom priestore do bodov, ktoré medzi sebou nemajú žiadnu súvislosť, teda robot prechádza sekvenciou diskretných bodov v priestore. Takýto druh riadenia sa predovšetkým využíva napr. pri bodovom zváraní. Dráhové riadenie sa využíva ak je potrebné riadiť robota v celej jeho dráhe, t.j. medzi jednotlivými bodmi dráhy robota existuje nejaká funkčná závislosť. [1, 3]

Programovanie trajektórie pohybu robota môžeme rozdeliť do dvoch základných skupín:

- On-line programovanie
- Off-line programovanie

On-line programovanie robota prebieha priamo na konkrétnom pracovisku, pričom operátor navádza robota do jednotlivých pracovných pozícií, ktoré sú následne ukladané do riadiaceho systému, a následne robot potom vykonáva tieto pohyby automaticky. Medzi on-line programovacie metódy patria:

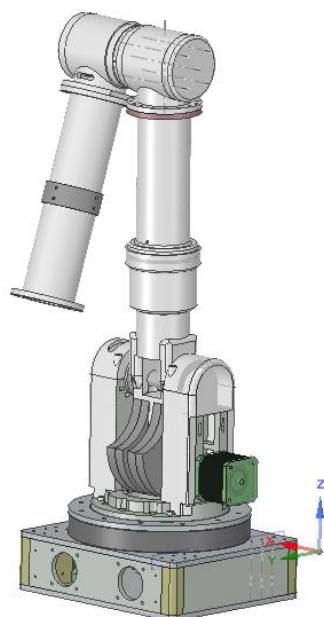
- Play-back – metóda vhodná na menej náročné pohyby s nižšou presnosťou, kde operátor prechádza robotom celý technologický proces
- Teach-pedant – metóda pri ktorej je pomocou ovládača (Pedant) navádzaný robot do jednotlivých presne definovaných polôh, ktoré sú zaznamenávané

Off-line programovanie neprebíha priamo na pracovisku, ale k tomuto spôsobu programovania robota je potrebný virtuálny model pracoviska a robota, na ktorom sa programátorom naprogramujú jednotlivé trajektórie, ktoré sú následne vykonávané na reálnom pracovisku.

Každá z týchto metód predstavuje svoje výhody a nevýhody, v ktorých sa zohľadňujú jednotlivé podmienky ako napr. dostupnosť technológie, bezpečnosť, možnosť prerušenia prevádzky počas programovania a pod. [1, 3]

II. NÁVRH ROBOTICKÉHO RAMENA

Pri návrhu robotického ramena bolo potrebné zohľadniť rôzne aspekty, ako napr. materiál na výrobu, mechaniku a elektrické obvody. Po analýze robotických systémov sme sa rozhodli pre vytvorenie angulárneho typu robota, nakoľko sa najčastejšie využíva v priemysle aj vo výskume. Návrh modelu robotického ramena je zobrazený na obr.2.



Obr. 2 Návrh modelu robotického ramena.

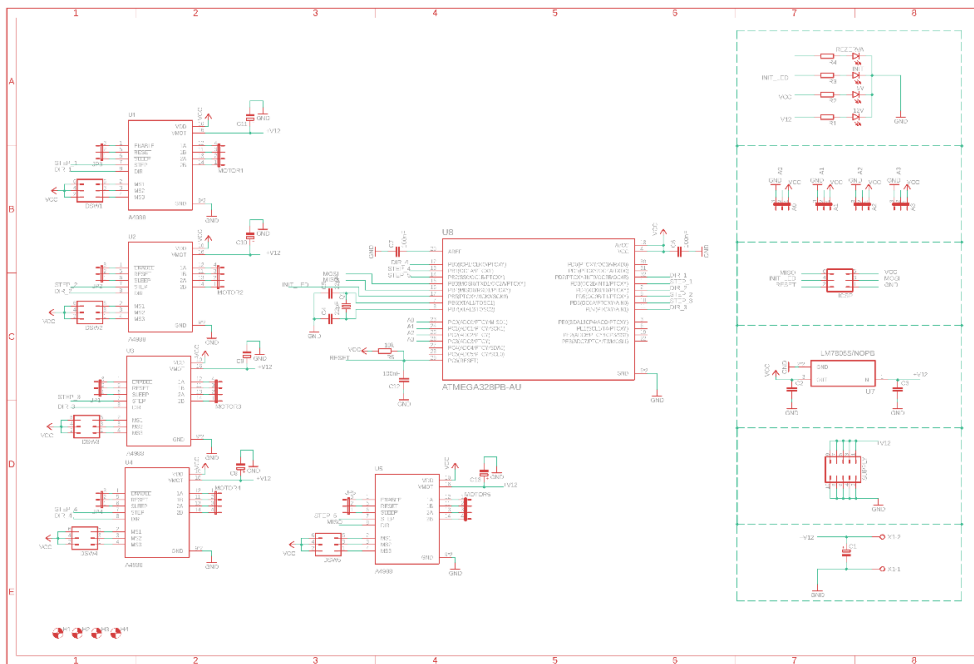
Navrhnutá konštrukcia robota pozostáva z podstavy, karuselu, ramena a koncového efektora. Celá podstava bola od začiatku návrhu vytvorená tak, aby sa v budúcnosti dala rozšíriť o lineárne ložiská pre posuv vo vodorovnej osi. Táto podstava bola zhotovená z ocelových platní. Karusel ramena ako aj zvyšné časti ramena boli vytvorené z materiálu PLA pomocou FDM 3D tlačiarne. Celým ramenom otáča sprevodovaný krokový motor uložený vo vnútornej časti podstavy. Rovnako na aj na pohon ostatných osí robotického ramena boli použité krokové motory, celkovo má navrhnuté robotické rameno 4 osí. Pre pohon všetkých osí boli použité krokové motory Creality 3D 42-40, ktorých parametre sú uvedené v tabuľke 1.

Tabuľka 1
Parametre krokového motora Creality 3D 42-40

Parameter	Hodnota
Nominálne napätie	4,83V
Nominálny prúd	1,5A
Otáčky	1-1000 ot./min
Nominálny moment	0,4Nm
Počet krokov na otáčku	200
Uhol kroku	1,8°

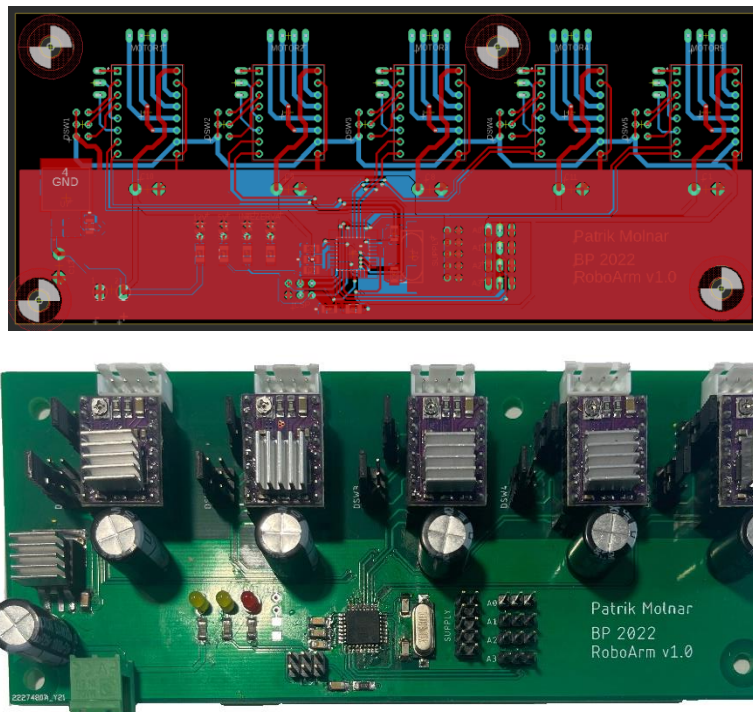
A. Riadiaca časť robota

Pri realizácii návrhu riadiacej časti robotického ramena sme využili program Autodesk EAGLE. Elektronika ramena pozostáva z napájacej, procesorovej a ovládacej časti. Keďže na pohon ramena sú využívané krokové motory, ktoré majú stálu spotrebu elektrickej energie, teda aj v statickom stave pre udržanie momentu, bolo nutné na napájanie využiť priemyselný spínaný 12V/20A zdroj. Nakoľko procesor potrebuje pre napájanie 5V/500mA bolo potrebné toto vstupné napájacie napätie upraviť pomocou stabilizátora napätia LM7805. Pre riadenie ramena bol využitý jednočipový mikroprocesor radu Atmel.



Obr. 3 Schéma zapojenia.

Riadiaca elektronika pre robotické rameno bola doplnená o ovládače pre použité krokové motory. Ovládače majú na doske plošného spoja vyvedené výstupy pre jumper spojky vďaka ktorým je možné nastaviť pre každý motor uhol jedného posunu na 1/4, 1/8 alebo 1/16 zo štandardného posunu o jeden krok. Ako už bolo spomenuté vyššie, krokové motory majú veľké prúdové zaťaženie, preto museli byť ovládače vybavené chladičom. Okrem ovládačov obsahuje doska aj analógové vstupy a výstupy pre prípadné pripojenie externých senzorov ako napr. senzor polohy alebo teploty. Navrhnutá elektronická časť je zobrazené na obr.4.



Obr. 4 Riadiaca doska robotického ramena.

Po skompletovaní jednotlivých dielov bola následne testovaná funkčnosť navrhnutého robotického ramena zobrazeného na obr. 5.



Obr. 5 Robotické rameno

III. ZÁVER

V tomto príspevku bol prezentovaný návrh konštrukcie a riadenia pre robotické rameno. Mechanická časť robotického ramena bola vytlačená z PLA materiálu na FDM 3D tlačiarne. Už pri výplni 20% malo rameno dostatočnú pevnosť a nakoľko PLA materiál odoláva teplotám približne do 150°C, nemalo by dochádzať k deformáciám ramena vplyvom teploty od motorov.

Na pohon robotického ramena boli použité krokové motory, ktoré majú veľký statický moment, ale napriek tomu dynamický moment motora nebol dostatočný a pri dlhšom chode dochádzalo prehrievaniu motorov. Z tohto dôvodu bude potrebné doplniť prevodovky pre pohony a doplniť chladenie elektrickej časti robota.

POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe Zmluvy č. APVV-19-0210.

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23.

LITERATÚRA

- [1] VIRALAB, 2011. Příručka Automatizační a robotická technika, SK 93100530, 2011, 64 s., 5 vydanie
- [2] ŠOLC, F. ŽALUD, L., 2001. Robotika. Brno : 2002, FEKT Vysoké učení technické Brno
- [3] JURÍŠICA, L. HUBINSKÝ, P. KORDOŠ, J., 2005. Robotika. Brno : 2005, Slovenská technická univerzita v Bratislave
- [4] ABB Robotics. [Online]. Dostupné na internete: <https://new.abb.com/products/robotics>
- [5] Allegro DMOS Microstepping Driver datasheet. [Online]. Dostupné na internete: https://components101.com/sites/default/files/component_datasheet/A4988%20Stepper%20Motor%20Driver%20Module.pdf
- [6] Atmel ATmega 328P datasheet. [Online]. Dostupné na internete: https://ww1.microchip.com/downloads/en/DeviceDoc/Atmel-7810-Automotive-Microcontrollers-ATmega328P_Datasheet.pdf
- [7] robotika.sk. [Online]. Dostupné na internete: <http://www.robotika.sk/mains.php>
- [8] In-Circuit Serial Programming. [Online]. Dostupné na internete: <https://ww1.microchip.com/downloads/en/DeviceDoc/30277d.pdf>

Aplikácia mobilného robota MiroSot v úlohe vyhýbania sa prekážkam

¹Silvester STRIGÁČ, ²Tomáš TKÁČIK, ³Ján JADLOVSKÝ

^{1,2,3} Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹silvester.strigac@student.tuke.sk, ²tomas.tkacik@tuke.sk, ³jan.jadlovsky@tuke.sk

Abstrakt — Tento článok sa zaoberá aplikáciou mobilného robota MiroSot v úlohe vyhýbania sa prekážkam na hernej ploche robotického futbalu. Súčasťou riešenej úlohy je lokalizácia mobilného robota a prekážok, ktorá sa vykonáva na základe obrazu z kamery umiestnenej nad futbalovým ihriskom. Z výsledkov lokalizácie mobilného robota a prekážok sa s využitím algoritmu A* určí najkratšia cesta do cieľovej pozície, ktorá sa vyhladí metódou kubickej Hermitovej interpolácie pre získanie referenčnej trajektórie. Mobilný robot následne sleduje referenčnú trajektóriu s využitím riadenia so spätnou väzbou navrhnutého na základe modelu mobilného robota. Výsledky simulácií a experimentálneho overenia aplikácie sú prezentované vo forme grafov vrátane ich zhodnotenia.

Keľúčové slová — Algoritmus A*, mobilný robot, model mobilného robota, navigačné riadenie, spracovanie obrazu

I. ÚVOD

Mobilná robotika je disciplína, ktorá sa vo všeobecnosti snaží dosiahnuť autonómny pohyb robota bez zásahu človeka. V rámci tohto článku je použitá robotická platforma MiroSot, ktorá bola vyvinutá na pracovisku mobilnej robotiky v Centre moderných metód riadenia a priemyselnej informatiky (CMMRaPI) na KKUI FEI TU [1]. Robotická platforma bola pôvodne skonštruovaná pre aplikáciu robotického futbalu [2]. Jej všestranná implementácia však umožňuje využitie v rôznych úlohách mobilnej robotiky [3], ako napríklad návrh stratégií v robotickom futbale [4], diagnostike [2], či hľadani najkratšej cesty v bludisku s pomocou neurónových sietí [5]. V rámci tohto článku popíšeme použitie tejto platformy v aplikácii vyhýbania sa prekážkam.

Základom aplikácie popisovanej v článku je mobilný robot MiroSot. Jedná sa o štandardizovaný model mobilného robota s pôdorysom hliníkovej konštrukcie o rozmeroch 7.5 cm x 7.5 cm a o celkovej hmotnosti 450 g [2]. Mobilný robot je vybavený dvomi nezávislými DC motormi, ktoré sú cez prevodovku spojené s pogumovanými hliníkovými kolesami. Spoločne tak poskytujú maximálne teoretické zrýchlenie $5 \text{ m}\cdot\text{s}^{-2}$ a maximálnu teoretickú rýchlosť $2.5 \text{ m}\cdot\text{s}^{-1}$. Mobilný robot je riadený 32-bit jednočipovým mikropočítačom rady STM32F103 a s riadiacim počítačom komunikuje pomocou Bluetooth rozhrania. Absolútna pozícia robota je určovaná pomocou kamery umiestnenej na strope, ktorá sníma hernú plochu zhora [6].

Program na riadenie mobilného robota je vybudovaný v jazyku C++ na robotickej platforme ROS2. Jedná sa o open-source platformu, ktorá poskytuje softvérové nástroje pre vytváranie aplikácií v mobilnej robotike. Platforma ROS2 poskytuje komunikačné mechanizmy na báze publisher-subscriber modelu, čím umožňuje rozdeliť aplikáciu do decentralizovaných modulov.

Článok je členený do piatich častí. V úvode je predstavená robotická platforma MiroSot. Druhá časť článku sa zaoberá lokalizáciou mobilného robota a prekážok v prostredí. Tretia časť článku sa venuje plánovaniu a realizácii pohybu mobilného robota. To zahŕňa nájdenie najkratšej trasy z počítačovej pozície do koncovej, vyhladenie trajektórie a realizáciu pohybu pri zohľadnení kinematických vlastností mobilného robota. Štvrtá časť článku popisuje realizované experimenty na simulačnej úrovni a na úrovni reálneho mobilného robota. V piatej a zároveň záverečnej časti článku zhodnocujeme výsledky a predostrieme možnosti ďalšieho rozvoja aplikácie.

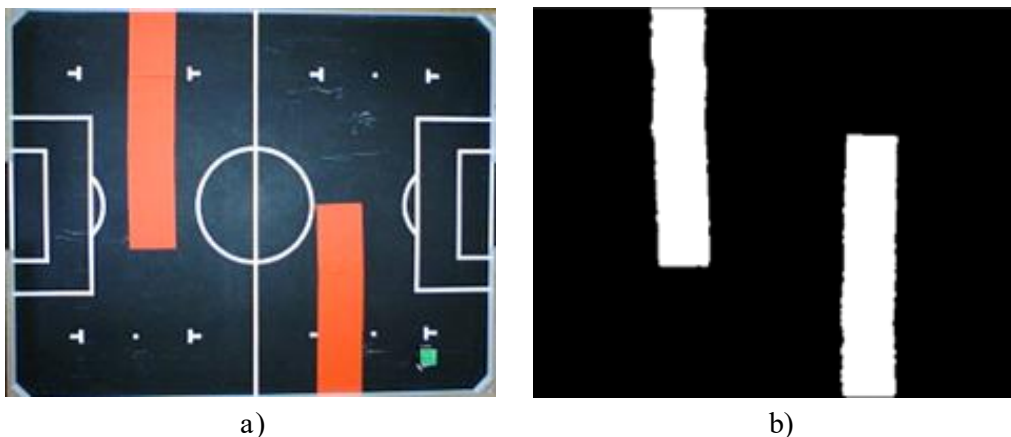
II. LOKALIZÁCIA MOBILNÉHO ROBOTA A PREKÁŽOK POMOCOU METÓD POČÍTAČOVÉHO VIDENIA

Prvým krokom v rámci aplikácie vyhýbania sa prekážkam je lokalizácia mobilného robota a prekážok v prostredí. Lokalizáciu vykonáme spracovaním obrazu z kamery, ktorá je umiestnená nad hernou plochou robotického futbalu [3]. Mobilný robot je z hornej strany

označený zeleným štvorcom a prekážky sú reprezentované červeným papierom položeným na hernej ploche. Výrazne odlišné farby mobilného robota, prekážok a pozadia (hernej plochy) nám umožnia jednoduchšie lokalizovať mobilného robota a prekážky [7].

Na Obr. 1a je zobrazený nespracovaný obraz z kamery. Prvým krokom spracovania obrazu je oddeliť pixely pozadia od pixelov predmetov záujmu. Nakoľko v aplikácii sme zvolili výrazné farby pre označenie mobilného robota a prekážok, môžeme využiť farebnú segmentáciu obrazu. Aby bola farebná segmentácia robustná voči zmene intenzity osvetlenia, vykonávame farebnú segmentáciu v HSV farebnom modeli. HSV farebný model využíva na reprezentáciu výslednej farby tri zložky (odtieň, saturáciu a hodnotu jasu), pomocou ktorých je možné priamo segmentovať objekty na základe ich farby. RGB farebný model sa v úlohe farebnej segmentácie nevyužíva, nakoľko odtieň farby je závislý od všetkých troch kanálov naraz [7]. Výstupom farebnej segmentácie je binárny obraz, kde pixely s logickou hodnotou 1 spadajú do definovaného farebného intervalu a ostatné pixely majú priradenú logickú hodnotu 0. Farebný interval pre zelenú farbu, resp. farbu mobilného robota sme zvolili v HSV farebnom modeli v intervale od [25,52,75] do [90,100,255]. Farebný interval pre červenú farbu zase od [0,65,45] do [10,100,255].

Obraz z kamery je v praxi častokrát zašumený, čo spôsobí aj zašumenie výsledného binárneho obrazu [8]. Tento šum by mohol spôsobiť nesprávnu lokalizáciu mobilného robota alebo prekážok. Z toho dôvodu sme použili morfológické operácie, ktorých úlohou je zmierniť vplyv nedokonalnej farebnej segmentácie hlavne vplyvom šumu. Morfológické operácie aplikujeme pomocou štruktúrného elementu v tvare plus (+). Najprv viacnásobne aplikujeme operáciu erózie, aby sme odstránili zrnitý šum a tenké čiary, ktoré môžu spôsobiť detekciu falošných objektov. Nežiaducim efektom tejto operácie je zmenšenie objektov záujmu. Z toho dôvodu následne vykonáme rovnaký počet krát operáciu dilatácie, čím sa snažíme navrátiť sledovaným objektom ich pôvodný tvar a veľkosť [8]. Vo výsledku tak získame binárny obraz, kde pixely s logickou hodnotou 1 reprezentujú mobilného robota v prípade segmentácie zelenej farby a prekážky v prípade segmentácie červenej farby. Na Obr. 1b je zobrazený spracovaný binárny obraz po aplikácii farebnej segmentácie a morfológických operácií v úlohe lokalizácie prekážok.



Obr. 1 Lokalizácia prekážok – a) pôvodný obraz z kamery; b) binárny obraz zachytávajúci prekážky [9].

Posledným krokom lokalizácie mobilného robota a prekážok je prevod z lokálneho súradnicového systému do globálneho, čím možno uzavrieť úlohu lokalizácie. Po úspešnej lokalizácii mobilného robota a prekážok je nevyhnutné nájsť cestu z aktuálnej pozície mobilného robota do cieľovej pozície, čomu sa budeme venovať v ďalšej časti.

III. PLÁNOVANIE A REALIZÁCIA POHYBU MOBILNÉHO ROBOTA

V tejto časti článku sa venujeme plánovaniu a samotnej realizácii pohybu mobilného robota. Pohyb mobilného robota je realizovaný z jeho aktuálnej pozície, ktorá môže byť ľubovoľná pri zabezpečení orientácie robota smerom hore (vzhľadom na Obr. 1a), do cieľovej pozície, ktorá je fixne stanovená. V úlohe plánovania pohybu používame algoritmus A* na nájdenie najkratšej cesty, ktorú následne vyhladáme, a pohyb robota realizujeme v spätnoväzobnej riadiacej štruktúre s dopredným zosilnením na základe kinematického modelu mobilného robota.

A. Hľadanie najkratšej cesty mobilného robota do cieľa

Prvým ktorom v plánovaní pohybu mobilného robota je nájdenie najkratšej cesty z aktuálnej pozície robota do cieľa pri zohľadnení prekážok. Na túto úlohu sa používajú algoritmy na hľadanie najkratšej cesty, ako napr. Dijkstrov algoritmus, A* a pod. V rámci navrhovanej

aplikácie sme zvolili algoritmus A* pre jeho jednoduchosť a efektívitu implementácie.

Segmentovaný obraz s prekážkami (viď. Obr. 1b) sme najprv diskretizovali do 2D mapy uzlov, pričom veľkosť jedného uzla zodpovedala veľkosti mobilného robota (7.5 cm x 7.5 cm). Uzly, ktoré boli aspoň čiastočne pokryté prekážkou, sme označili za nepriechodné. Štartovaciu pozíciu mobilného robota sme určili na uzol, ktorý obsahoval ťažisko mobilného robota z úlohy lokalizácie. Rovnako sme určili aj cieľový uzol na základe pozície cieľa [9]. Takto zadefinovanú úlohu možno riešiť pomocou algoritmu A*.

Algoritmus A* priradí každému uzlu mapy hodnotu $g(n)$ a $h(n)$. Hodnota $g(n)$ reprezentuje cenu dosiahnutia n -tého uzla z počiatočného uzla a hodnota $h(n)$ predstavuje hodnotu heuristickej ceny dosiahnutia cieľa z n -tého uzla. Hodnota heuristickej ceny pre dosiahnutie cieľa $h(n)$ sa častokrát volí ako najkratšia vzdialenosť medzi n -tým uzlom a cieľom, resp. ako Euklidovská vzdialenosť n -tého uzla od cieľového. Súčet oboch hodnôt $f(n) = g(n) + h(n)$ vyjadruje odhad celkovej ceny cesty pri použití n -tého uzla, pričom algoritmus A* sa snaží túto hodnotu minimalizovať [10].

Algoritmus A* operuje nad dvomi množinami uzlov, tzv. otvorené a uzavreté. Otvorené uzly možno považovať za nepreskúvané. Algoritmus v každej iterácii vyberie jeden uzol z množiny otvorených uzlov s najnižšou hodnotou $f(n)$, aktualizuje hodnotu $g(n)$ tohto uzla, a presunie ho do množiny uzavretých uzlov. Tento postup sa opakuje, kým nebol dosiahnutý cieľový uzol, alebo kým neboli vyčerpané všetky otvorené uzly (neexistuje cesta medzi počiatočným a cieľovým uzlom). Takto navrhnutý algoritmus považujeme za oportunistický, nakoľko volí cestu najvyššieho zisku [10].

Nájdená cesta algoritmom A* sa skladá z diskretných bodov, resp. uzlov, ktoré nemožno použiť priamo v mobilnej robotike. Z toho dôvodu je potrebné nájdenú cestu interpolovať a vytvoriť tak referenčnú trajektóriu.

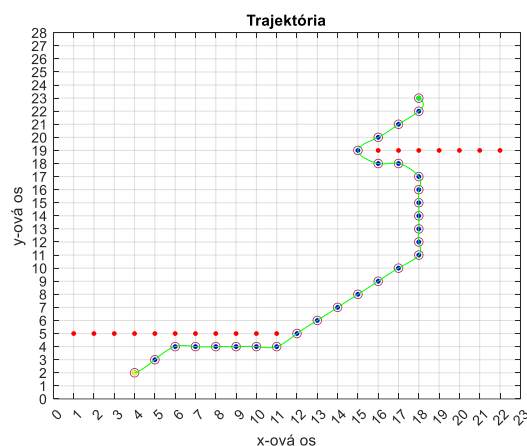
B. Plánovanie trajektórie pohybu mobilného robota

Na získanie referenčnej trajektórie je nevyhnutné interpolovať uzly, ktoré sme získali pri hľadaní najkratšej cesty. Interpoláciou získame spojitú referenčnú trajektóriu, ktorú bude mobilný robot sledovať. Na interpoláciu referenčnej trajektórie sme zvolili metódu kubickej Hermitovej spline (CHS) interpolácie, ktorá patrí do skupiny metód lokálnej interpolácie [11].

Interpoláčna metóda CHS využíva 4 po sebe idúce uzly (U_0 až U_3). Trajektória sa interpoluje medzi dvomi prostrednými (U_1 a U_2) uzlami, pričom krajné uzly (U_0 a U_3) ovplyvňujú zakrivenie výslednej trajektórie. Výpočet ľubovoľného bodu $p(t)$ medzi prostrednými uzlami (U_1 a U_2) je realizovaný nasledovne:

$$p(t) = h_{00}(t)U_1 + h_{10}(t)(U_1 - U_0) + h_{01}(t)U_2 + h_{11}(t)(U_3 - U_2), \quad (1)$$

kde t je nezávislá premenná v rozsahu 0 až 1 udávajúca relatívnu pozíciu výsledného bodu $p(t)$ a $h_{00}(t)$, $h_{10}(t)$, $h_{01}(t)$, $h_{11}(t)$ sú Hermitovské základné funkcie [11].



Obr. 2 Nájdená optimálna cesta z počiatočnej pozície [4, 2] do cieľovej pozície [18, 23], [9].

Na Obr. 2 sú uzly nájdenej cesty algoritmom A* znázornené modrými bodkami a získaná krivka metódou CHS interpolácie znázornená zelenou farbou. Vypočítaná krivka je referenčná trajektória, ktorú bude mobilný robot sledovať [9]. Sledovaním referenčnej trajektórie dosiahneme cieľ aplikácie, a to vyhýbanie sa prekážkam pri ceste do cieľovej pozície. Pre potreby sledovania referenčnej trajektórie je potrebné zohľadniť kinematické a dynamické vlastnosti mobilného robota.

C. Model a realizácia pohybu mobilného robota

Na realizáciu pohybu mobilného robota po referenčnej trajektórii je nevyhnutné poznať dynamický a kinematický model mobilného robota [12]. V prípade mobilného robota MiroSot bol dynamický model odvodený v [13] a algoritmus riadenia zohľadňujúci dynamický model je implementovaný vo firmvéri mobilného robota. Dynamický model zahŕňa riadenie rýchlosti kolies diferenciálneho podvozku na referenčnú rýchlosť, ktorá sa do mobilného robota posieľa cez robotickú platformu ROS2 [3]. V rámci tejto časti sa zameriame na kinematický model mobilného robota MiroSot.

Kinematický model sa skladá z kinematického modelu orientovaného bodu a kinematického modelu podvozku mobilného robota [13]. Orientovaný bod je aproximáciou mobilného robota, v ktorej ignorujeme fyzické rozmery robota, a je umiestnený v strede osi otáčania kolies. V prípade mobilného robota MiroSot považujeme stred otáčania kolies aj za ťažisko robota. Pohyb orientovaného bodu je odvodený od uhlovej ω a lineárnej v rýchlosti mobilného robota [12]. Rýchlosť pohybu a orientáciu možno vyjadriť v maticovom tvare:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\varphi} \end{bmatrix} = \begin{bmatrix} \cos \varphi & 0 \\ \sin \varphi & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix}, \quad (2)$$

kde vektor $[\dot{x}, \dot{y}, \dot{\varphi}]^T$ vyjadruje zmenu polohy a orientácie orientovaného bodu v čase a vektor $[v, \omega]^T$ je vektor vstupných rýchlostí (lineárnej a uhlovej) do kinematického modelu.

Kinematický model diferenciálneho podvozku mobilného robota je vyjadrený pre rýchlosti oboch kolies robota (ω_r a ω_l). Tieto rýchlosti sú transformované na lineárnu v a uhlovú ω rýchlosť orientovaného bodu. Transformácia z lineárnej v a uhlovej ω rýchlosti orientovaného bodu na uhlové rýchlosti pravého ω_r a ľavého ω_l kolesa diferenciálneho podvozku mobilného robota je definovaná nasledovne [12]:

$$\begin{bmatrix} \omega_r \\ \omega_l \end{bmatrix} = \begin{bmatrix} \frac{1}{R} & \frac{L}{2R} \\ \frac{1}{R} & -\frac{L}{2R} \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix}, \quad (3)$$

kde R je polomer kolies a L je vzdialenosť medzi kolesami.

Na sledovanie referenčnej trajektórie mobilným robotom využijeme spätnoväzobnú riadiacu štruktúru s dopredným zosilnením (štruktúra FFC). Cieľom dopredného zosilnenia je sledovať referenčnú trajektóriu a cieľom spätnej väzby je kompenzovať prípadné odchýlky od tejto trajektórie. Referenčná trajektória musí zohľadniť aj celkový čas T , za ktorý ju mobilný robot prejde. Následne trajektóriu rozdelíme do N bodov, pričom platí $T = N \cdot T_s$, kde T_s je konštantná perióda vzorkovania. Tým sme rozdelili celú dráhu pohybu robota do N bodov, ktoré budú predstavovať čiastočný cieľ robota v k -tej iterácii. V každom bode určíme doprednú lineárnu $v_{ff}(k)$ a uhlovú $\omega_{ff}(k)$ rýchlosť orientovaného bodu [13]:

$$v_{ff}(k) = \frac{\Delta \rho_{ref}}{\Delta t} = \frac{\sqrt{[x_{ref}(k) - x_{ref}(k-1)]^2 + [y_{ref}(k) - y_{ref}(k-1)]^2}}{t(k) - t(k-1)}, \quad (4)$$

$$\varphi_{ff}(k) = \tan^{-1} \frac{\Delta y_{ref}}{\Delta x_{ref}} = \tan^{-1} \frac{y_{ref}(k) - y_{ref}(k-1)}{x_{ref}(k) - x_{ref}(k-1)}, \quad (5)$$

$$\omega_{ff}(k) = \frac{\Delta \varphi_{ff}}{\Delta t} = \frac{\varphi_{ff}(k) - \varphi_{ff}(k-1)}{t(k) - t(k-1)}. \quad (6)$$

Takto sme definovali sekvenciu krokov, ktoré musí robot vykonať, aby sa dostal zo svojej aktuálnej pozície do cieľovej pozície referenčnej trajektórie. To však nezohľadňuje poruchy reálneho sveta, ako sú napríklad trenie, nepresnosti v dynamickom modeli, chyby aktuátorov a pod. Aby mobilný robot reálne dosiahol cieľ riadenia, je nevyhnutné navrhnuť spätnú väzbu, ktorá tieto poruchy kompenzuje. Cieľom spätnej väzby je minimalizovať pozičnú chybu e_p orientovaného bodu, ktorá je definovaná nasledovne:

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{ref} - x \\ y_{ref} - y \\ \varphi_{ref} - \varphi \end{bmatrix}. \quad (7)$$

Za predpokladu malých odchýlok možno využiť zákon riadenia Δv_{fb} v tvare:

$$\Delta v_{fb} = K_{fb} \Delta e_p = \begin{bmatrix} k_1 & 0 & 0 \\ 0 & \text{sgn}(v_{ff}) \cdot k_2 & k_3 \end{bmatrix} \begin{bmatrix} \Delta e_1 \\ \Delta e_2 \\ \Delta e_3 \end{bmatrix}, \quad (8)$$

kde k_1, k_2 a k_3 sú koeficienty matice spätnoväzobného zosilnenia K_{fb} , ktorých návrh je uvedený v [13]. Výsledný zákon riadenia je v tvare:

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \begin{bmatrix} v_{ff} \cos(e_3) - v_{fb} \\ \omega_{ff} - \omega_{fb} \end{bmatrix}. \quad (9)$$

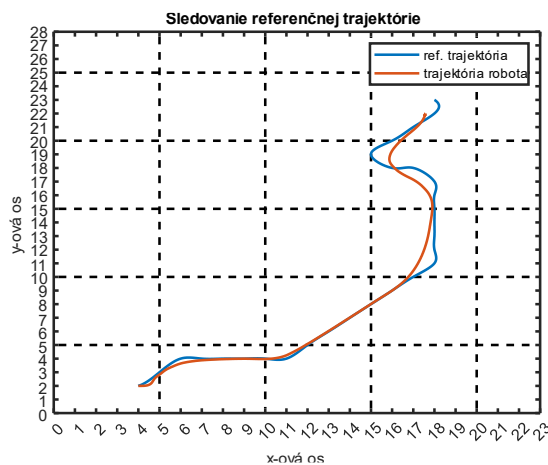
Navrhnutý zákon riadenia je implementovaný v riadiacej slučke v rámci robotického platformy ROS2. V nasledujúcej časti článku sa budeme venovať overeniu navrhnutej aplikácie vyhýbania sa prekážkam.

IV. EXPERIMENTÁLNE OVERENIE POHYBU ROBOTA PRI SLEDOVANÍ REFERENČNEJ TRAJEKTÓRIE

Výsledná aplikácia zahŕňa detekciu objektov (mobilného robota a prekážok), plánovanie trajektórie (nájdanie najkratšej cesty a jej interpolácia) a navigačné riadenie za účelom sledovania referenčnej trajektórie. Cieľom aplikácie je dostať mobilného robota do cieľovej pozície tak, aby sa vyhol prekážkam.

A. Simulácia navigačného riadenia za účelom vyhýbania sa prekážkam

V prvom kroku experimentálneho overenia navrhnutej aplikácie sme implementovali riadenie na simulačnej úrovni. Mobilný robot v rámci simulačného overenia začínal s počiatocnými súradnicami $[x_0; y_0] = [4; 2]$ a natočením $\varphi_0 = \frac{\pi}{4}$. Cieľová pozícia, na ktorú sa má dostať mobilný robot, je $[x_N; y_N] = [18; 23]$. Výstup simulačného overenia aplikácie je zobrazený na Obr. 3.

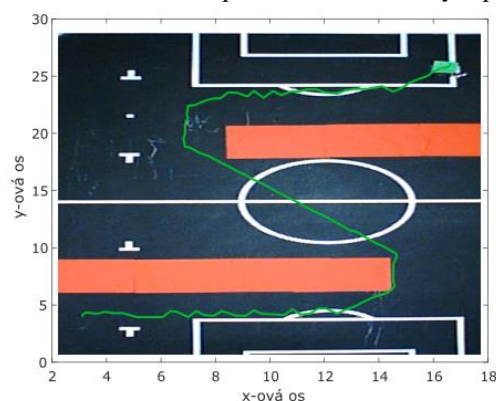


Obr. 3 Sledovanie referenčnej trajektórie navigačným riadením mobilného robota [9].

Výsledky simulácie zobrazené na Obr. 3 naznačujú nedostatky navrhnutého navigačného riadenia, ktoré sú zreteľné najmä pri náhlych zmenách referenčnej trajektórie. Aj napriek týmto nedostatkom je nevyhnutné overiť navrhnuté riadenie aj na reálnom mobilnom robotovi.

B. Navigačné riadenie reálneho modelu mobilného robota MiroSot za účelom vyhýbania sa prekážkam

Pri overovaní aplikácie na reálnom mobilnom robotovi je počiatocná pozícia mobilného robota $[x_0; y_0] = [3.11; 4.11]$ s natočením $\varphi_0 = 0$ a cieľová pozícia $[x_N; y_N] = [17; 26]$. Prekážky v rámci aplikácie zobrazené na Obr. 4 a sú reprezentované červeným papierom.



Obr. 4 Sledovanie trajektórie reálnym mobilným robotom MiroSot v aplikácii vyhýbania sa prekážkam [9].

Na Obr. 4 je zobrazené experimentálne overenie navrhnutej aplikácie, kde zelená čiara reprezentuje trajektóriu mobilného robota detegovanú kamerou. Mobilný robot MiroSot sa dostal z počiatočnej do cieľovej pozície za čas 10.51 s. Cieľová pozícia na, ktorú sa mobilný robot MiroSot dostal, je $[x_N; y_N] = [16.65, 26.12]$. Táto pozícia sa líši od požadovanej cieľovej pozície, čo je spôsobené ukončovacou podmienkou, ktorá požaduje presnosť cieľovej pozície na úrovni $[\Delta x_N; \Delta y_N] = [\pm 1, \pm 1]$ na predchádzanie nežiaducich oscilácií v cieľi. Výchyľky voči referenčnej trajektórii na Obr. 4 môžu byť spôsobené nedokonalým určením pozície robota, čo môže byť ešte viac zvýraznené spätnou väzbou v navigačnom riadení.

Vo výsledku možno prehlásiť, že aplikáciu vyhýbania sa prekážkam sa podarilo overiť na simulačnej úrovni, a aj pomocou reálneho mobilného robota MiroSot.

V. ZÁVER

Tento článok prezentuje návrh a výsledky aplikácie pre vyhýbanie sa prekážkam v navigačnom riadení s využitím mobilného robota MiroSot. Úloha lokalizácie mobilného robota a prekážok je realizovaná s využitím metód počítačového videnia, čoho výstupom je binárny obraz zachytávajúci sledované objekty. Binárny obraz zachytávajúci prekážky je následne diskretizovaný do 2D mapy uzlov, ktorá sa využíva pri hľadaní cesty z počiatočnej pozície do cieľovej. Na nájdenie cesty je využitý algoritmus A*, ktorého výstup je postupnosť uzlov. Táto postupnosť je následne interpolovaná s využitím kubickej Hermitovej interpolácie pre získanie referenčnej trajektórie. Na sledovanie referenčnej trajektórie mobilným robotom je implementované riadenie v štruktúre FFC. Výsledná aplikácia je overená simulačne a následne s pomocou reálneho mobilného robota MiroSot s uspokojivými výsledkami. Aplikácia reálneho mobilného robota MiroSot je realizovaná pomocou robotickej platformy ROS2.

V budúcnosti možno vylepšiť aplikáciu presnejšou lokalizáciou mobilného robota, ktorá by zohľadňovala aj jeho orientáciu. Zároveň možno využiť metódy optimalizácie referenčnej trajektórie, ktorá je aktuálne optimalizovaná pre osemsmerný pohyb v mriežke.

REFERENCIE

- [1] CMMRaPI: Laboratórium výrobných liniek a rozpoznávania obrazov V145. [cit. 19.5.2023] Dostupné na internete: <http://kyb.fe.i.tuke.sk/laboratoria/miest/V147.php>.
- [2] KOPČÍK, M.: Vytvorenie metodiky pre diagnostiku senzorických systémov v rámci distribuovaného systému riadenia so zameraním na mobilnú robotiku : Dizertačná práca. Košice: FEI TU, 2017. 132 s. (školiťel: doc. Ing. Ján Jadlovský, CSc.)
- [3] JADLOVSKÝ, J., JADLOVSKÁ, S., ČERKALA, J., KOPČÍK, M., ČABALA, J., ORAVEC, M., VARGA, M., VOŠČEK, D., TKÁČIK, M., BŘEZINA, A.: Návrh metodiky pre modelovanie, riadenie, simuláciu a nedeštruktívnu diagnostiku mobilných robotov. In: Strojárstvo (11/2017): MEDIA/ST, Žilina. ISSN 1335-2938.
- [4] ŠTEFANIŠIN, J.: Návrh a realizácia stratégie mobilného robotického futbalu kategórie MiroSot v rámci systému ROS : Diplomová práca. Košice: FEI TU, 2018. 72 s. (školiťel: doc. Ing. Ján Jadlovský, CSc.)
- [5] KONTUĽ, J.: Optimalizácia riadiaceho systému pre pohyb mobilného robota v bludisku na báze systému ROS s využitím neuronových sietí a učenia s posilňovaním : Diplomová práca. Košice: FEI TU, 2020. 105 s. (školiťel: doc. Ing. Ján Jadlovský, CSc.)
- [6] CMMRaPI: Robotický futbalista MiroSot. [cit. 19.5.2023] Dostupné na internete: <http://kyb.fe.i.tuke.sk/laboratoria/modely/msp.php>.
- [7] BUNDZEL, M., ZOLOTOVÁ, I.: Počítačové videnie v praxi. Košice: elfa, 2013. 91 s. ISBN 978-80-8086225-1.
- [8] GOYAL, M.: Morphological image processing. In: IJCSST. Roč. 2, č. 4 (2011). s. 161-165. ISSN 0976-8491.
- [9] STRIGÁČ, S.: Návrh aplikácie pre riešenie komplexných úloh mobilnej robotiky s využitím moderných simulačných nástrojov MATLAB a ROS : Diplomová práca. Košice: FEI TU, 2022. 66 s. (školiťel: doc. Ing. Ján Jadlovský, CSc.)
- [10] DUCHOŇ, F., BABINEC, A., KAJAN, M., BEŇO, P., FLOREK, M., FICO, T., JURÍŠICA, L.: Path Planning with Modified a Star Algorithm for a Mobile Robot. In: Procedia Engineering. Roč. 96 (2014), s. 59-69. ISSN 1877-7058.
- [11] LI, J., CHEN, S.: The Cubic α -Catmull-Rom Spline. In: Mathematical and Computational Applications. Roč. 21, č. 3 (2016), s.1-14. ISSN 2297-8747.
- [12] ČERKALA, J., JADLOVSKÁ, A., KLEIN, T.: Modeling and Control of Mobile Robot with Differential Chassis, In: Electrical Engineering and Informatics 6 : proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. - Košice : TU, 2015 S. 656-661. ISBN 978-80-5532178-3.
- [13] ČERKALA, J., JADLOVSKÁ, A.: Matematické modelovanie a riadenie kolesových mobilných robotov. Košice: elfa, 2019. 274 s. ISBN 978-80-8086-273-2.

Možnosti simulácie strát výkonových polovodičových súčiastok v prostredí MATLAB/Simulink

Marek PÁSTOR, Jaroslava ŽILKOVÁ, Juraj JEVČÁK

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

marek.pastor@tuke.sk, jaroslava.zilkova@tuke.sk, juraj.jevcak@student.tuke.sk

Abstrakt — Článok sa zaoberá možnosťami simulácie vodivostných a spínacích strát výkonových polovodičových súčiastok v prostredí MATLAB/Simulink. Článok porovnáva výsledky simulácie pre zvolený integrovaný výkonový polovodičový modul of firmy Infineon so simuláciou v programe PLECS Standalone.

Keywords — MATLAB/Simulink, menič, model, straty

I. ÚVOD

Počítačová simulácia a využívanie počítačových modelov pri návrhu elektronických systémov je dnes štandardom [1]-[3]. Z pohľadu zamerania sa na účinnosť meničov je potrebné predikovať účinnosť výkonového polovodičového meniča ešte vo fáze jeho návrhu a simulácie. Samotná topológia meniča nie je schopná akumulovať a premieňať energiu. Straty v meniči tak vznikajú hlavne v dôsledku spínania výkonových polovodičových súčiastok ako spínacie straty a vedenia prúdu ako vodivostné straty.

V oblasti výkonovej elektroniky je k dispozícii široká ponuka programov pre simulovanie vlastností výkonových polovodičových meničov. Štandardom v tejto oblasti je program OrCAD/PSpice od firmy Cadence, ktorý však nie je vhodný pre meniče s pokročilými algoritmi riadenia. Pre návrh meniča spolu s moderným riadením je výhodné použiť programy ako MATLAB/Simulink, Plecs, Plexim apod. Tieto programy umožňujú okrem simulácie samotného výkonového obvodu aj aplikovanie pokročilých metód riadenia [4].

II. VÝKONOVÉ POLOVODIČOVÉ MODULY

A. Použitie

Ak má výkonový polovodičový menič štandardnú štruktúru v podobe jednofázového alebo trojfázového mostíka tak pre zjednodušenie konštrukcie meniča a jeho chladenia sú často pre konštrukciu meniča využívané bezpotenciálové výkonové polovodičové moduly (Obr. 1).



a) výkonový polovodičový modul malého výkonu

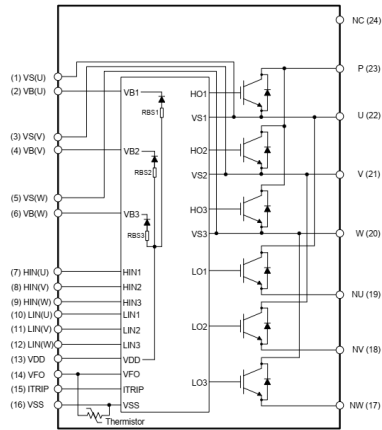


b) výkonový polovodičový modul stredného výkonu

Obr. 1 Bezpotenciálové výkonové polovodičové moduly [5] [6]

Okrem výkonových polovodičových súčiastok, ktorými sú najčastejšie tranzistory MOSFET alebo IGBT môže výkonový polovodičový modul obsahovať aj budiace obvody a prípadne tepelnú a nadprúdovú ochranu (Obr. 2). To zvyšuje spoľahlivosť zapojenia výkonového

polovodičového meniča a umožňuje dosahovať vysoké výkonové hustoty. Integrovanie výkonovej a budiacej časti meniča do jedného modulu výrazne zjednodušuje návrh najrozšírenejších topológií výkonových polovodičových meničov. Integrovanie výkonovej časti meniča však komplikuje meranie na meniči, kde napr. pri module na Obr. 2 nie je možné samostatne merať kolektorové prúdy horných tranzistorov. Pre určenie vodivostných a spínacích strát meniča je tak veľmi vhodné použiť simuláciu a celkové straty modulu verifikovať len tepelným meraním.



Obr. 2 Typická štruktúra výkonového polovodičového modulu s integrovaným budičom [5]

B. Straty výkonového polovodičového modulu

Straty výkonového polovodičového modulu sú prevažne tvorené stratami výkonových polovodičových súčiastok. Tie sa delia na straty vodivostné a straty spínacie. Vodivostné straty sú závislé od veľkosti prúdu pretekajúceho súčiastkou a spínacie straty závisia od veľkosti prúdu, napätia a spínacej frekvencie. Straty samozrejme závisia aj od parametrov výkonovej súčiastky ako je saturačné napätie, odpor kanálu v zopnutom stave, spínacie časy a teplota. Pre modelovanie bude použitý výkonový polovodičový modul IKCM15H60GA ktorý obsahuje tranzistory IGBT a antiparalelné diódy a jeho vnútorná štruktúra je znázornená na Obr. 2.

Vodivostné straty bipolárnych súčiastok

Na vodivosti bipolárnych súčiastok sa podieľajú ako majoritné, tak aj minoritné nosiče náboja. Ich výstupná charakteristika tak má charakteristické prahové napätie a po jeho prekonaní je aproximovaná priamkou ktorej smernicou je odpor v zapnutom stave. Výstupné charakteristiky tranzistora IGBT sú opísané rovnicou priamky [7]:

$$u_{CE} = U_{CE0} + r_T i_C \quad (1)$$

kde:

U_{CE0} – prahové napätie (kolektorové napätie pri nulovom kolektorovom prúde),

r_T – odpor v zapnutom stave.

Vodivostné straty tranzistora IGBT P_V je možné definovať na základe strednej (AV) a efektívnej (RMS) hodnoty kolektorového prúdu [7]:

$$P_{VT} = U_{CE0} I_{AV} + r_T I_{RMS}^2 \quad (2)$$

Statická charakteristika diódy je opísaná rovnicou priamky [7]:

$$u_D = U_{D0} + r_D i_D \quad (3)$$

U_{D0} – prahové napätie (napätie v priepustnom pri nulovom anódovom prúde),

r_D – diferenciálny odpor v zapnutom stave.

Vodivostné straty výkonovej diódy P_{VD} je možné definovať na základe strednej (AV) a efektívnej (RMS) hodnoty anódového prúdu:

$$P_{VD} = U_{D0}I_{AV} + r_D I_{RMS}^2 \quad (4)$$

Spínacie straty bipolárnych súčiastok

Spínacie straty tranzistora IGBT sú tvorené zapínacími a vypínacími stratami. Pri zanedbaní spätného zotavenia diód je možné zapínacie straty tranzistora IGBT P_{on} možné určiť na základe katalógovej hodnoty zapínacieho času tranzistora t_{on} [7]:

$$P_{on} = f_s \frac{U_{IC}}{2} (t_{ri} + t_{fv}) = f_s \frac{U_{IC}}{2} t_{on} \quad (5)$$

Podobne vypínacie straty tranzistora IGBT P_{off} je možné určiť na základe katalógovej hodnoty vypínacieho času tranzistora t_{off} [7]:

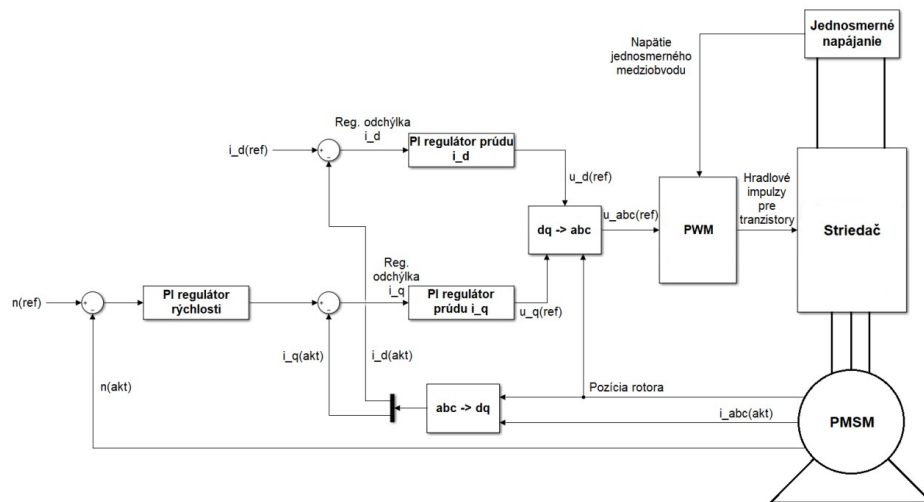
$$P_{off} = f_s \frac{U_{IC}}{2} (t_{rv} + t_{fi}) = f_s \frac{U_{IC}}{2} t_{off} \quad (6)$$

Spínacie straty výkonovej diódy sú tvorené stratami vypínacími spôsobenými spätným zotavením diódy pri jej vypínaní [7]:

$$P_{offD} = f_s \frac{1}{4} Q_{rr} U_D \quad (7)$$

III. SYNCHRÓNNY MOTOR S PERMANENTNÝMI MAGNETMI

Pre zabezpečenie reálnych pracovných podmienok výkonového polovodičového modulu bol tento použitý pre napájanie malého synchronného motora s permanentnými magnetmi. Jeho otáčky boli riadené vektorovým riadením, ktorého štruktúra je znázornená na Obr. 3.



Obr. 3 Bloková schéma riadenia synchronného motora s permanentnými magnetmi [8]

Základné parametre motora použité pri simulácii sú uvedené v Tabuľke 1.

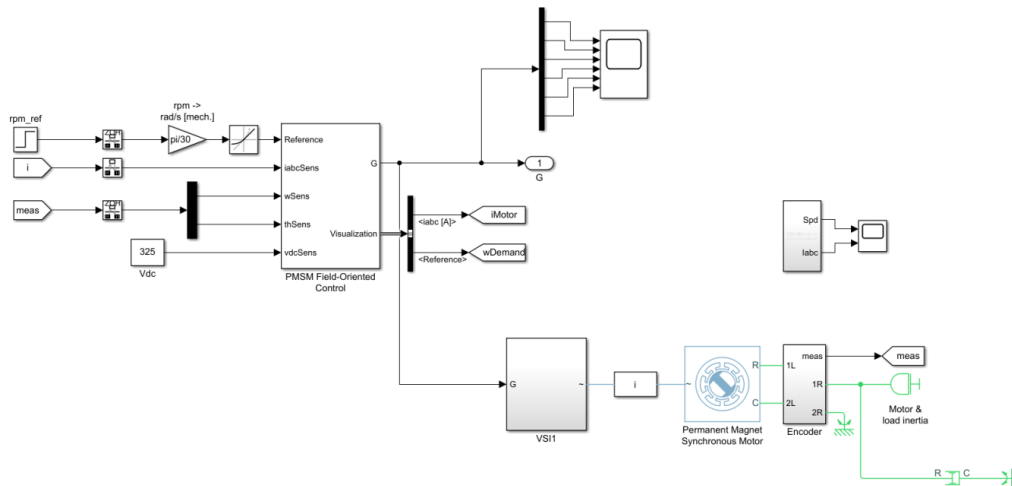
Tabuľka 1
 Vlastnosti pohonu so synchronným motorom s permanentnými magnetmi

Symbol	Veličina	Hodnota
U	napätie medziobvodu	325 V
P_N	nominálny výkon motora	850 W
M_N	nominálny moment motora	2,1 Nm
$2p$	počet párov pólov	4
L_d	indukčnosť motora v osi d	15,5 mH
L_q	indukčnosť motora v osi q	21,5 mH
f_s	spínacia frekvencia	14 kHz

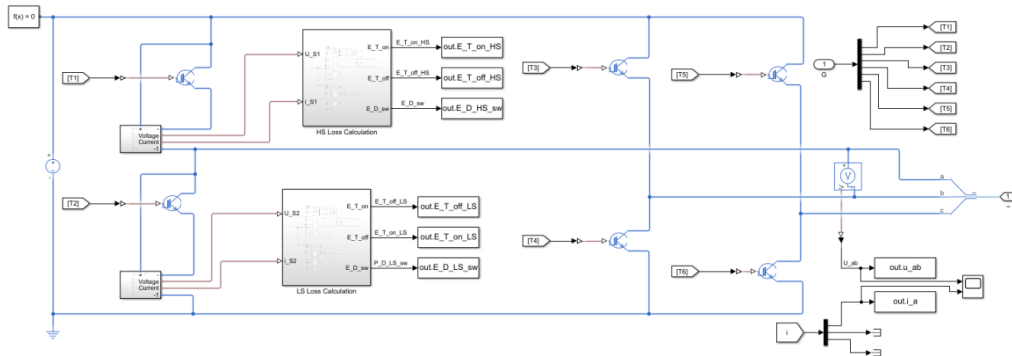
IV. SIMULÁCIA V PROSTREDÍ MATLAB/SIMULINK

A. Programové schémy

Základná programová schéma je znázornená na Obr. 4. Pozostáva z bloku vektorového riadenia synchronného motora s permanentnými magnetmi (PMSM Field-Oriented Control) výstupom ktorého je vektor hradlových impulzov (G) privedený do bloku napäťového striedača (VSI1). Výstup striedača je pripojený k motoru, ktorého otáčky sa snímajú pomocou enkódera. Podsystem napäťového striedača je znázornený na Obr. 5. Skladá sa z troch vetiev tvorených horným a spodným tranzistorom IGBT s antiparalelnou diódou. Keďže počas ustálenej činnosti motora sú priebehy všetkých horných a spodných súčiastok v rámci mostíka identické, tak straty sú merané len na jednom hornom (HS) a jednom spodnom (LS) tranzistore.



Obr. 4 Programová schéma riadenia synchronného motora s permanentnými magnetmi v prostredí MATLAB/Simulink [8]



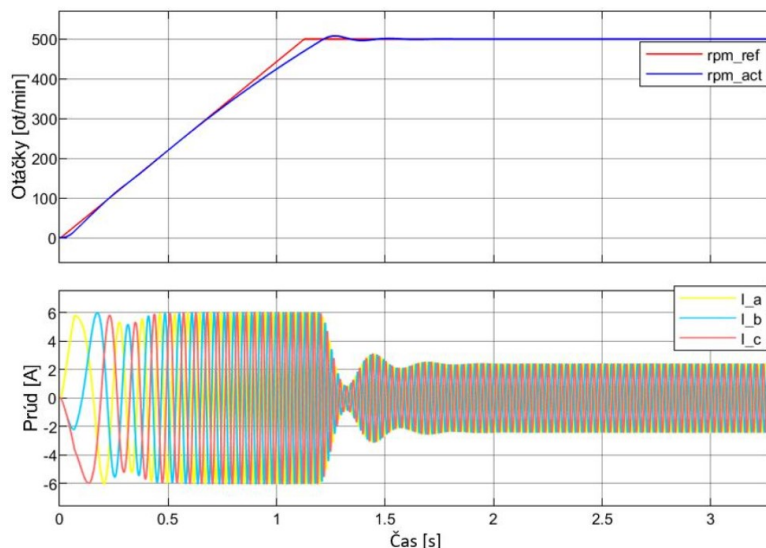
Obr. 5 Programová schéma podsystemu striedača [8]

Pre určenie strát tranzistorov a diód bola zvolená metóda vyhľadávacej tabuľky. Do nej boli zadané parametre z katalógového listu integrovaného výkonového modulu. Pre spínacie straty pri zapnutí a vypnutí tranzistorov IGBT boli na základe (5) a (6) zvolené dvojrozmerné vyhľadávacie tabuľky, ktorých vstupom boli kolektorové napätie a kolektorový prúd tranzistorov. Vodivostné straty boli určené na základe (2) z jednorozmernej tabuľky, ktorej vstupom bol kolektorový prúd IGBT tranzistorov. Pre diódy boli simulované vypínacie straty na základe (7) a vodivostné straty na základe (4).

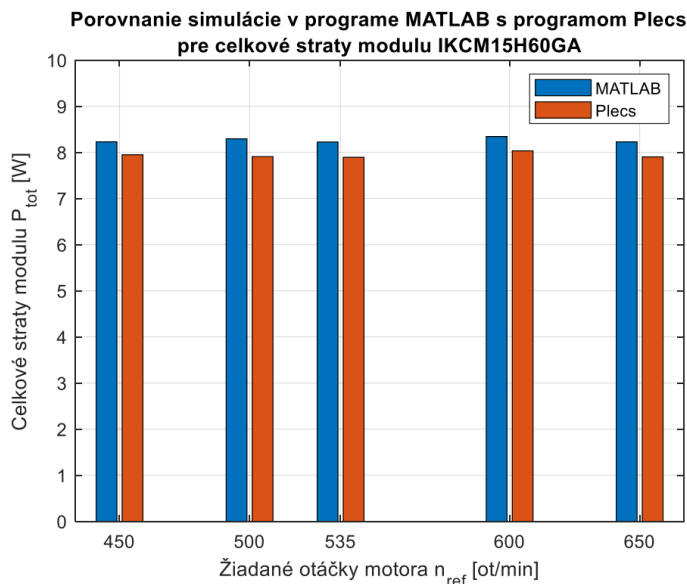
V. VÝSLEDKY SIMULÁCIE

Simulácia motora napájaného z integrovaného modulu IKCM15H60GA bola vykonaná pre rôzne rýchlosti otáčania sa motora v rozsahu 450 až 650 ot./min. Výsledky simulácie v prostredí MATLAB/Simulink boli porovnané so simuláciou v programe PLECS Standalone pre rovnakú konfiguráciu simulovaného systému.

Na Obr. 6 je znázornený simulovaný rozbeh motora na 500 ot./min a snímané prúdy motora použité pre odhad strát integrovaného výkonového modulu IKCM15H60GA v ustálenom stave otáčok motora.



Obr. 6 Priebeh otáčok a prúdov synchronného motora s permanentnými magnetmi napájaného z modulu IKCM15H60GA pre referenčnú hodnotu otáčok 500 ot/min [8]



Obr. 7 Porovnanie celkových strát modulu simulovaných v programe MATLAB/Simulink a Plecs Standalone pre rôzne otáčky motora [8]

Z porovnania celkových výkonových strát integrovaného modulu znázornených na Obr. 7 je vidieť, že oba programy predikujú v podstate rovnaké celkové straty. Keďže motor nebol pri simulácii zaťažovaný, tak celkové straty modulu závisia od otáčok len minimálne. Simulácia umožňuje odhad spínacích a vodivostných strát samostatne pre horný a dolný tranzistor ako aj pre ich antiparalelné diódy (celkovo 8 hodnôt strát). Simulované výsledky neboli verifikované meraním.

VI. ZÁVER

Článok sa zaoberá možnosťami simulácie celkových výkonových strát integrovaného výkonového modulu pri napájaní synchronného motora s permanentnými magnetmi v prostredí MATLAB/Simulink. Toto prostredie je často zvolené pre simuláciu riadiacich štruktúr elektrických pohonov. Zakomponovanie výpočtu strát výkonovej časti meniča umožňuje predikovať účinnosť systému. To môže byť prínosom napr. pri vývoji bielej techniky, kde sa na základe takejto simulácie dá predikovať zaradenie spotrebiča do energetickej triedy. Výsledky simulácie je ešte potrebné overiť meraním, prípadne simuláciou v programe, ktorý využíva PSpice modely tranzistorov a diód, a nie len predikuje straty na základe katalógového listu súčiastok.

POĎAKOVANIE

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23.

REFERENCIE

- [1] de Paula, W.J., Tavares, G.H.M., Soares, G.M., Almeida, P.S. and Braga, H.A.C. (2020), Switching losses prediction methods oriented to power MOSFETs – a review. *IET Power Electronics*, 13: 2960-2970. <https://doi.org/10.1049/iet-pel.2019.1003>
- [2] Syed, S.A.; Khalid, H.A.; Farooq, H. Analytical and Simulation Comparison of Losses in Non-Isolated DC/DC Converter Using Si and SiC Switches for PV Application. *Eng. Proc.* 2022, 20, 11. <https://doi.org/10.3390/engproc2022020011>
- [3] Hafezi, H.; Faranda, R. A New Approach for Power Losses Evaluation of IGBT/Diode Module. *Electronics* 2021, 10, 280. <https://doi.org/10.3390/electronics10030280>
- [4] M. Leso, J. Zilkova, P. Girovsky, "Development of Simple Fuzzy Logic Controller for DC-DC Converter", *Proc. IEEE 18th International Power Electronics and Motion Control Conference (PEMC)*, 2018, pp. 86–93
- [5] Infineon, katalógový list IKCM15H60GA: <https://www.infineon.com/cms/en/product/power/intelligent-power-modules-ipm/ikcm15h60ga>
- [6] Infineon, katalógový list F3I400R10W3S7F: https://www.infineon.com/cms/en/product/power/igbt/igbt-modules/f3i400r10w3s7f_b11/
- [7] D. Graovac, M. Purschel, IGBT Power Losses Calculation Using the Data-Sheet Parameters, Infineon, p. 17, 2009
- [8] J. Jevčák, Simulovanie výkonových polovodičových meničov s využitím toolboxu Simscape, diplomová práca, Technická univerzita v Košiciach, p. 87, 2023

Vizualizácia priemyselného robota

Jaroslava ŽILKOVÁ, Peter GIROVSKÝ

Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,
Technická Univerzita v Košiciach, Slovensko

jaroslava.zilkova@tuke.sk, peter.girovsky@tuke.sk

Abstrakt — Predložený príspevok sa zaoberá vizualizáciou priemyselných robotov v programe MATLAB. Vytvorené grafické používateľské rozhranie (GUI) umožňuje používateľovi jednoduchšiu prácu prostredníctvom interaktívnych grafických prvkov. GUI dovoľuje vizualizovať pohyby robota a slúži pre intuitívne a používateľsky nenáročné zobrazenia pozícií a trajektórií robota na základe riešení priamej a inverznej kinematickej úlohy.

Kľúčové slová — priemyselný robot, program MATLAB, robot ABB IRB 120, vizualizácia

I. ÚVOD

Na vysoké školy technického zamerania prichádzajú absolventi rôznych typov stredných škôl, pričom ich vedomosti zvyčajne nie sú na rovnakej úrovni. Pri riešení jednotlivých úloh je v rámci výučby veľmi dôležité, aby študent porozumel samotnému zadaniu, vedel stanoviť postupy riešenia, ktoré mu zabezpečia dosiahnutie cieľa, ku ktorému sa má dopracovať. Jednou z možností ako podporiť klasické osvedčené metódy výučby je rozšírenie o nástroje, ktoré ľahšie sprístupnia študentom skúmanú problematiku a podporia ich motiváciu pre štúdium jednotlivých predmetov. Počítačové modelovanie a využitie 3D modelov pri simuláciách reálnych situácií je jedným z takýchto silných nástrojov, ktorý sa v súčasnosti uplatňuje aj vo väčšine priemyselných odvetví. Počítačové modelovanie umožňuje nielen priblížiť reálny systém študentom, ale aj vizualizovať jeho činnosť s možnosťou uváženia rôznych reálnych situácií a pôsobenia rôznych vonkajších vplyvov.

Oblasť robotiky je výrazne dynamicky rozvíjajúcou sa oblasťou, pričom nové oblasti aplikácií sa rozširujú a využitie robotov z klasických priemyselných oblastí rýchle preniká do oblastí mimopriemyselných činností. Priemyselné roboty sú tradične najviac nasadzované v automobilovom, elektrotechnickom a strojárskom priemysle, pričom sa aplikácia robotov posúva aj do nevýrobných procesov, ako sú servisné činnosti a rozširuje sa ich použitie aj v oblasti poľnohospodárstva, stavebníctva, spracovaní potravín, upratovaní, bankovníctva, rehabilitácie i pomoci imobilným a telesne postihnutým ľuďom a v ďalších. V oblasti aplikácií robotov sa v súčasnosti venuje veľká pozornosť rozvoju hlavne mobilných, servisných robotov a tzv. osobných robotov či už pre zábavu alebo ako pomocníkov a spoločníkov.

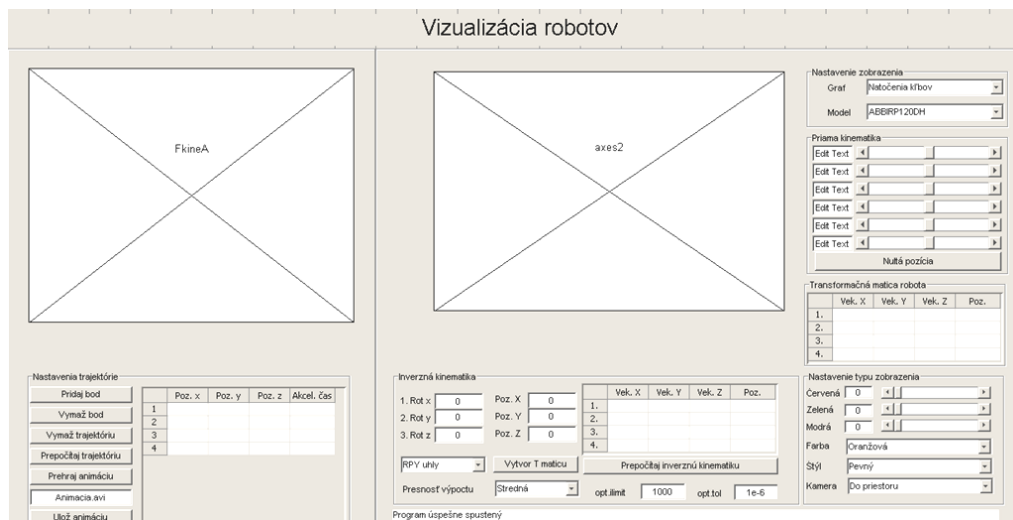
V určitých profesijných oblastiach neatraktivnosť a náročnosť pracovných činností rozširuje priestor pre nasadzovanie hlavne priemyselných robotov. Sú to už vyššie spomenuté tradičné oblasti použitia priemyselných robotov, kde sa stali už neoddeliteľnou súčasťou automatizovaných výrobných procesov. Pre absolventov vysokých škôl technických smerov je preto nevyhnutné, aby sa vedeli aktívne zapojiť do riešenia súčasných a nových technických úloh, ktorých náročnosť s rozvojom robotiky stále narastá. Pre podporu vzdelania v oblasti priemyselných robotov bolo, na základe [1-8], vytvorené grafické užívateľské rozhranie (GUI), ktoré študentom umožní nielen prácu s 3D modelom priemyselného robota ale aj zobrazenie aktuálnych stavov simulovaného robota, umožní rôzne spôsoby zobrazenia a zavedenie zmien. GUI dovoľuje vizualizovať pohyby robota a slúži pre intuitívne a používateľsky nenáročné zobrazenia pozícií a trajektórií robota na základe riešení priamej a inverznej kinematickej úlohy.

II. TVORBA GUI

Hlavným cieľom GUI je uľahčenie práce pre používateľov, ktorí nemusia poznať názvy a argumenty jednotlivých funkcií. Pre návrh GUI bol použitý editor programu MATLAB, GUIDE, ktorý umožňuje navrhovať grafické používateľské rozhrania [1].

A. Návrh GUI

Pred samotným návrhom rozhrania je dôležité vhodne rozmiestniť jednotlivé prvky na ploche editora. Prítom je vhodné prvky s podobnými funkciami združovať do jednotlivých panelov tak, aby ovládanie GUI bolo intuitívne, pričom názvy jednotlivých prvkov by mali výstižne popisovať, čo dané prvky zabezpečujú. Na Obr. 1 je zobrazený návrh rozmiestnenia grafických prvkov na ploche editora.

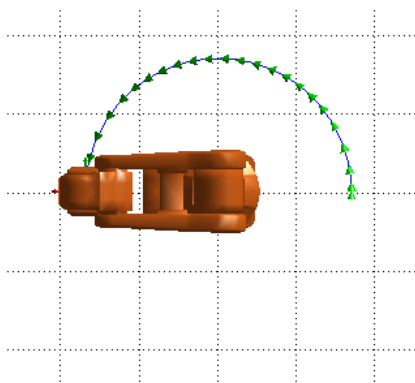


Obr. 1 Návrh rozmiestnenia grafických prvkov na ploche editora.

Po rozmiestnení a pomenovaní jednotlivých prvkov je dôležité každému prvku nastaviť parameter Tag, ktorý určuje názov prvku. Je nutné, aby názov prvku bol jednoslovný z dôvodu, že pri ukladaní GUI program MATLAB vygeneruje súbor, ktorý bude obsahovať funkcie z názvami totožnými s parametrom Tag. Vytvorené GUI obsahuje panely, v ktorých sú umiestnené prvky vytvorené na základe [1-8], určené na:

- Priame ovládanie priemyselného robota (Panel priama kinematika)
- Nepriame určenie pozície a orientácie priemyselného robota (Panel inverzná kinematika)
- Tvorbu trajektórie robota a prehrávanie animácie pohybu robota po trajektórii (Panel nastavenia trajektórie)
- Zmenu typu zobrazenia priemyselného robota (Panel nastavenia typu zobrazenia)
- Zmenu typu priemyselného robota a zmenu zobrazovania trajektórie do grafov (Panel nastavenie zobrazenia)

GUI okrem uvedených panelov obsahuje dve grafické okná.

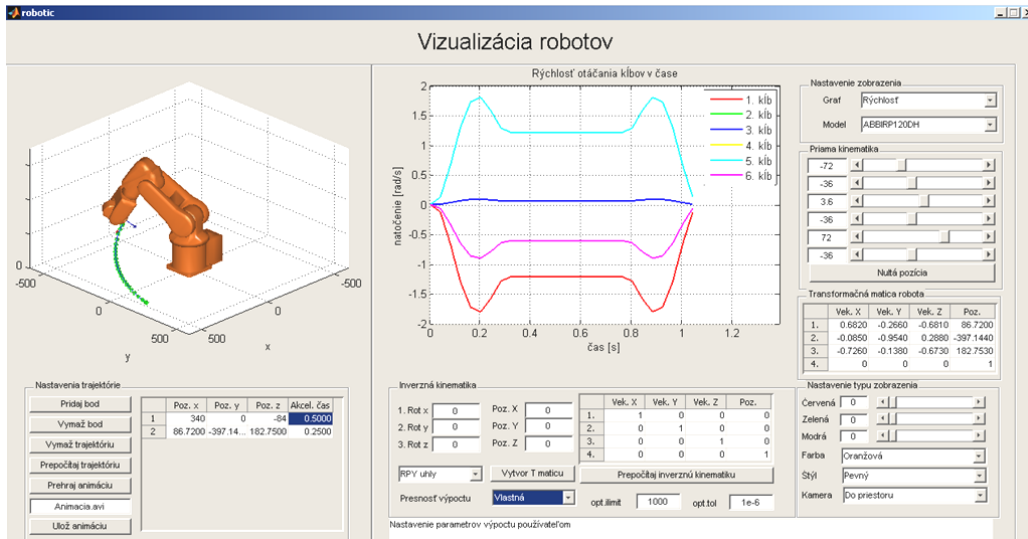


Obr. 2 Symbolické zobrazenie Robota ABB-IRB 120 so znázornením trajektórie na ploche editora.

Do pravého grafického okna sa vykresľuje graf trajektórie v závislosti od času. Do ľavého grafického okna (Obr. 2, Obr. 3) sa vykresľuje priemyselný robot zo symbolickým znázornením pozície a orientácie koncového člena robota a trajektória znázornená čiarou zo šípkami, ktoré vyznačujú orientáciu trajektórie.

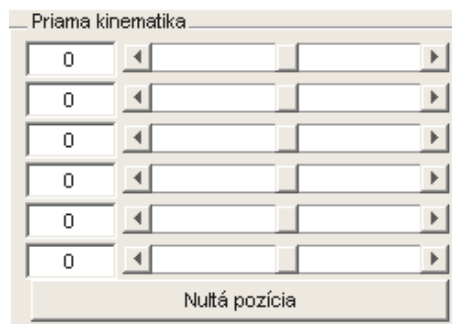
B. Popis vytvoreného GUI

Na Obr.3 je uvedené realizované rozmiestnenie jednotlivých grafických panelov podľa predchádzajúceho návrhu, ktorý je uvedený na Obr. 2. V tejto časti budú stručne charakterizované jednotlivé panely GUI.



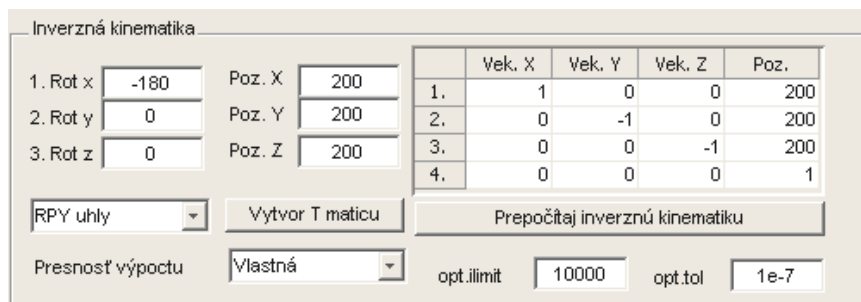
Obr. 3 Realizované rozmiestnenie grafických prvkov na ploche editora.

Panel Priama kinematika obsahuje 6 skupín posuvníkov a editačných textov, ktoré sú zobrazené na Obr.4. Tieto skupiny slúžia na nastavovanie spojových uhlov priemyselného robota. V editačnom texte sa zadávajú spojové uhly v stupňoch. Všetky posuvníky a editačné texty pre nastavovanie spojových uhlov sú navzájom prepojené a pri zmene hodnoty jedného objektu sa automaticky daná hodnota prenastaví na druhý objekt. Zmenou hodnoty objektu sa automaticky prekreslí aj vizualizačný model robota a transformačná matica koncového člena robota sa zapíše do tabuľky. Panel tiež obsahuje tlačidlo Nulťá pozícia, ktoré po stlačení vynuluje spojové uhly robota.



Obr. 4 Panel priama kinematika na ploche editora.

Panel Inverzná kinematika, navrhnutý pomocou [1-8], slúži na vytvorenie požadovanej transformačnej matice a prepočítanie spojových uhlov priemyselného robota, je zobrazený na Obr.5.



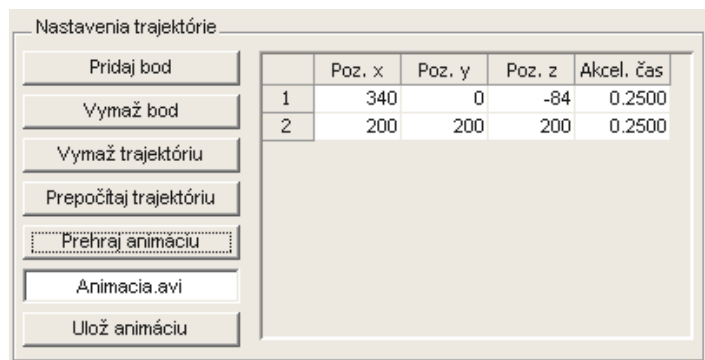
Obr. 5 Panel Inverzná kinematika na ploche editora.

Na vytvorenie transformačnej matice obsahuje panel v ľavej hornej strane šesť editačných textov. Editacie texty označené popisom Rot. slúžia na vytvorenie rotačnej časti matice. Rotačná matica sa počíta trojuholnými funkciami eul a RPY podľa [1-4]. Výber príslušnej funkcie sa vykonáva v menu. Editacie texty označené Poz. slúžia na zadávanie posuvnej časti transformačnej matice. Tlačidlom Vytvor T maticu sa následne vytvorí transformačná matica, ktorá sa zapíše do tabuľky. Tlačidlo Prepočítaj inverznú kinematiku slúži na prepočítanie inverznej kinematickej úlohy priemyselného robota. Presnosť výpočtu sa nastavuje v menu, v závislosti od presnosti výpočtu závisí doba výpočtu inverznej kinematickej úlohy priemyselného robota. V prípade nastavenia vlastnej presnosti výpočtu je možné zadávať parametre výpočtu prostredníctvom zmeny parametrov výpočtu (opt.limit a opt.tol).

Panel Nastavenie trajektórie, zobrazený na Obr.6, predstavuje nástroj na vytvorenie trajektórie a tvorbu videa. Tlačidlo Pridaj bod slúži na pridanie spojových uhlov pre aktuálnu pozíciu robota a zapíše pozíciu v karteziánskom priestore do tabuľky.

Tabuľka obsahuje stĺpec akceleračný čas, ktorý je možný v tabuľke ľubovoľne meniť. Tlačidlo Vymaž bod slúži na vymazanie posledného pridaného bodu do trajektórie. Tlačidlo Vymaž trajektóriu slúži na zmazanie celej trajektórie. Tlačidlo Prepočítaj trajektóriu prepočítava trajektóriu. Túto trajektóriu vykreslí do grafického okna k vizualizačnému modelu robota. Tiež vykreslí graf trajektórie do praveho grafického okna, ako je to zobrazené na Obr.3.

Tlačidlo Prehraj animáciu slúži na prehrávanie animácie pohybu robota po vygenerovanej trajektórii. Tlačidlo Ulož trajektóriu slúži na vytvorenie a uloženie animácie do pracovného adresára.



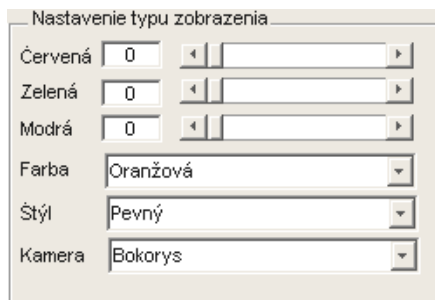
Obr. 6 Panel nastavenia trajektórie na ploche editora.

Panel Nastavenie typu zobrazenia, ktorý je zobrazený na Obr.7 slúži na zmenu nastavenia vizuálneho zobrazenia robota. V menu Farba sa nastavuje farba priemyselného robota, okrem predvolených farieb je možné nastavovať vlastnú farbu priemyselného robota prostredníctvom posuvníkov. V menu Štýl sa nastavuje spôsob zobrazenia robota. Na výber sú tri možnosti:

- Pevný štýl zobrazenia
- Sieťový štýl zobrazenia
- Priehľadný štýl zobrazenia

V menu Kamera sa nastavuje spôsob snímania scény kamerou:

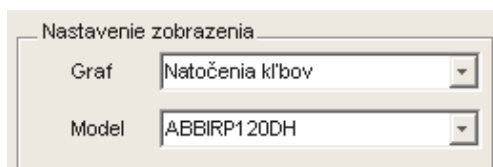
- Pohľad do priestoru
- Pohľad zhora
- Pohľad z boku
- Pohľad spredu



Obr. 7 Panel nastavenia typu zobrazenia na ploche editora.

Panel Nastavenie zobrazenia zobrazený na Obr.8 sa skladá z dvoch objektov typu menu. V prvom menu, ktoré je nazvané Graf sa nastavuje požadované vykreslenie grafu trajektórie priemyselného robota, ktoré umožňuje zobraziť jednotlivé grafy zobrazenia trajektórie:

- Závislosť natočenia spojových uhlov robota od času
- Závislosť rýchlosti spojových uhlov robota od času
- Závislosť pozície konečného článku robota od času



Obr. 8 Panel nastavenia zobrazenia.

V menu Model sú na výber dva vizualizačné modely priemyselného robota. Prednastavený model robota je menej detailný. Model ABBIRP120HPFK predstavuje vizualizačný model robota importovaný priamo od výrobcu.

POĎAKOVANIE

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23.

LITERATÚRA

- [1] Mathworks MATLAB: R2012a Documentation [online], <http://www.mathworks.com/help/techdoc/>.
- [2] P. Corke: *Robotics Toolbox for MATLAB*, Release 9. <http://www.petercorke.com/RTB/robot.pdf>.
- [3] P. Corke: *Robotics, Vision and Control*, Springer-Verlag Berlin Heidelberg, 2011. 570 s.
- [4] Reza N. Jazar: *Theory of Applied Robotics*, Springer New York Dordrecht Heidelberg London, 2007, 893 s.
- [5] MATLAB E-manual, [online]. http://www.mathworks.com/help/matlab/learn_matlab/product-description.html
- [6] MATLAB E-manual, [online]. Dostupné na internete :http://www.mathworks.com/help/matlab/matlab_oop/why-use-object-oriented-design.html
- [7] MATLAB E-manual, [online]. http://www.mathworks.com/help/matlab/matlab_prog/use-a-matlab-timer-object.html
- [8] MATLAB E-manual, [online]. http://www.mathworks.com/help/matlab/creating_plots/accessing-object-handles.html
- [9] A. Karger, M. Kargerová: *Základy robotiky a prostorové kinematiky*, ČVUT Praha, 2000, 265s. .
- [10] M. Lešo.,: *Vizualizácia v robotike*, Bakalárska práca, FEI TU v Košiciach, 2012.
- [11] V. Vranič: *Objektovo-orientované programovanie*, STU v Bratislave, 2008.

Rozpoznávanie používateľov v mobilnej aplikácii pomocou snímačov odtlačkov prstov

¹Renát HALUŠKA, ²Erik ŠVENK, ³Matúš PLEVA

^{1,2,3} Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹renat.haluska@tuke.sk, ²erik.svenk.student@tuke.sk, ³matus.pleva@tuke.sk

Abstrakt — S rozšíreným používaním smartfónov a rastúcou potrebou bezpečných a personalizovaných používateľských skúseností sa rozpoznávanie používateľov stalo kľúčovým aspektom vývoja mobilných aplikácií. Snímače odtlačkov prstov, ktoré sú v súčasnosti bežne integrované do mobilných zariadení, ponúkajú pohodlnú a spoľahlivú biometrickú metódu autentifikácie. V tomto článku skúmame využitie snímačov odtlačkov prstov na rozpoznávanie používateľov v mobilných aplikáciách. Rozoberáme základnú technológiu snímačov odtlačkov prstov, ich integráciu do mobilných zariadení a rôzne algoritmy a techniky používané na rozpoznávanie používateľov. Uvádžeme aj prípadovú štúdiu, v ktorej upravujeme mobilnú aplikáciu využívajúcu snímače odtlačkov prstov na autentifikáciu používateľov.

Kľúčové slová — biometria, biometrické systémy, odtlačky prstov, snímače odtlačkov prstov

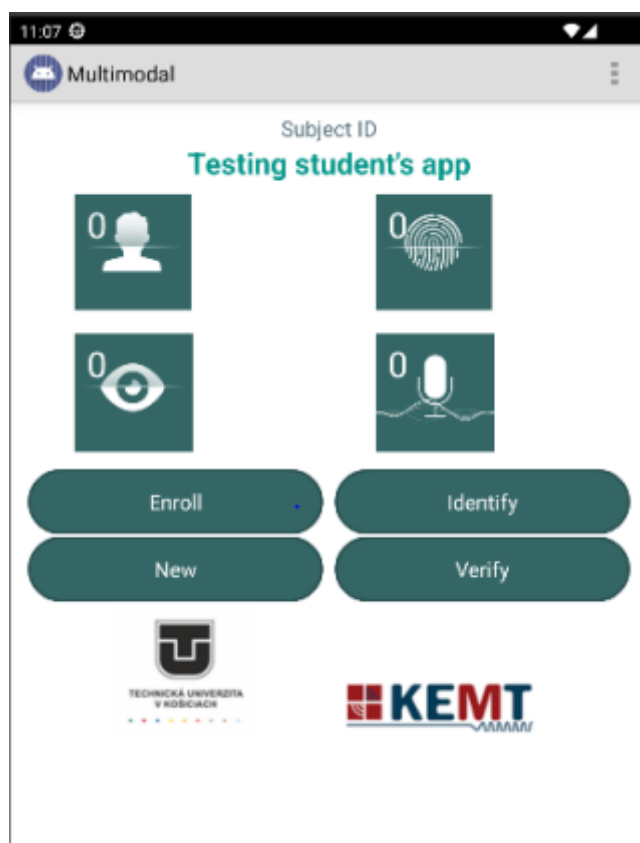
I. ÚVOD

Úlohy biometrickej identifikácie sa posudzujú na dva účely, ako je identifikácia a overovanie. V identifikačných systémoch sa identita osoby určuje z databázy známych záznamov [1]. Osoba si nerobí nárok na identitu a biometrické údaje sa vyhľadávajú podľa celej databázy šablón, čo si vyžaduje porovnávanie jedna k mnohým. V systémoch overovania sa biometrická šablóna osoby porovnáva len so šablónou uloženou v databáze identity, ktorú používateľ vyhlasuje za svoju vlastnú [2]. Vyžaduje sa len porovnanie jeden s jedným. Typický systém biometrickej identifikácie funguje v dvoch fázach, t. j. vo fáze registrácie a vo fáze identifikácie.

Počas fázy registrácie sa digitálna reprezentácia biometrických údajov získava skenovaním biometrických charakteristík špeciálnym biometrickým skenerom [3]. Digitálna reprezentácia sa ďalej spracúva s cieľom získať expresívnu reprezentáciu nazývanú šablóna. Počas identifikačnej fázy sa opakuje ten istý proces na vytvorenie šablóny a potom sa porovnáva so šablónami v databáze pomocou „matcher“ na určenie identity [4]. Ideálna biometria by mala mať vlastnosti univerzálnosti, jedinečnosti, stability, zberateľnosti a akceptovateľnosti [5]. Všeobecný biometrický systém obsahuje štyri moduly, ako sú snímače, modul hodnotenia kvality a extrakcie príznakov, modul porovnávania a rozhodovania a modul databázy.

II. MOBILNÁ APLIKÁCIA A POROVNANIE EXTERNÝCH SNÍMAČOV ODTLAČKOV PRSTOV

Aplikácia má viacero využití v biometrickej bezpečnosti, pretože dokáže spracovať viacero biometrických prvkov, a to odtlačky prstov, tvár, dúhovku a hlas, je tiež potrebné reprezentovať časť aplikácie, ktorú budeme používať. Naše možnosti navigácie v aplikácii sú znázornené na Obr. 1. Používateľ má na hlavnej obrazovke aplikácie na výber 5 možností. Niektoré z možností, konkrétne Enroll (Zapísať), Identify (Identifikovať) a Verify (Overiť), sú však schopné fungovať až po načítaní vzorky, ktorú by sme mohli buď uložiť do databázy, alebo identifikovať či overiť. Ďalšie tlačidlo, ktoré sa nachádza na hlavnej obrazovke, je tlačidlo "New" (Nový), ktoré vymaže načítané vzorky a zadanú identifikáciu (ID) a umožní nám načítať ďalšieho používateľa. Na hlavnej obrazovke sa nachádzajú aj možnosti merania troch ďalších biometrických údajov. Na vytvorenie novej upravenej verzie aplikácie sa používa nástroj Gradle. Vďaka nástroju Gradle môžeme vytvoriť aplikáciu oveľa jednoduchšie a rýchlejšie ako bez neho.

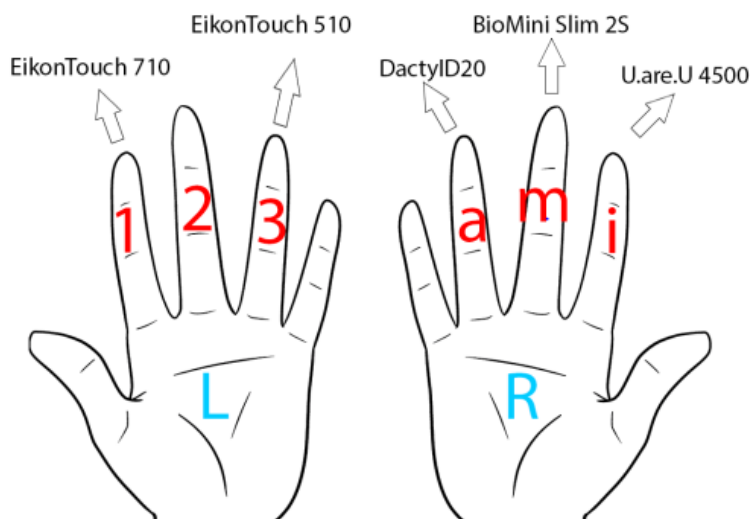


Obr. 1 Príklad mobilnej aplikácie

Na testovanie máme k dispozícii 5 snímačov odtlačkov prstov:

- BioMini Slim 2S,
- GreenBit DactyID20,
- DigitalPersona U.are.U 4500,
- EikonTouch 510,
- EikonTouch 710.

Z týchto snímačov sú prvé 3 optické snímače a zvyšné 2 sú kapacitné snímače. Aby bolo porovnanie jednoduchšie a prehľadnejšie, je potrebné vykonať niekoľko krokov, ktoré pomôžu pri meraní a vyhodnocovaní výsledkov. Ako prvý krok pri meraní s viacerými snímačmi na viacerých prstoch je potrebné určiť skratky pre prsty, aby sme mohli rozlíšiť, konkrétny prst. Keďže máme na meranie päť snímačov, vyberieme aj päť prstov, z ktorých odoberieme ich odtlačky (Obr. 2).



Obr. 2 Označenie prstov pre jednotlivé snímače

III. ZHROMAŽĎOVANIE TESTOVANÝCH HODNÔT

Zber testovacích hodnôt, je rozdelený na dve základné časti, a to:

- Zápis odtlačkov prstov do databázy.
- Porovnanie odtlačkov prstov so vzorkami v databáze.

Pri porovnávaní sa môžu porovnávať len 2 alebo viac prvkov. V tomto kroku dosiahneme, že v databáze máme prvý prvok, v našom prípade vzorku. Daný postup možno dosiahnuť tak, že vezmeme jeden zo snímačov a pre konkrétny prst, ktorý je spojený s daným externým snímačom odtlačkov prstov, ho nahráme do databázy.

Ďalším krokom, ktorý možno považovať za najdôležitejší, je samotné porovnanie odtlačkov prstov. Keďže základné odtlačky prstov už máme zapísané v našej databáze, všetko je pripravené na tento ďalší krok. Postup v tomto kroku bude takmer rovnaký ako v predchádzajúcom kroku, ale vzorku odtlačkov prstov nebudeme ukladať do databázy, ale budeme ju porovnávať so vzorkou, ktorá už bola do databázy nahraná v predchádzajúcom kroku. Po odobratí odtlačku prsta ho môžeme porovnať v mobilnej aplikácii pomocou tlačidla „Identifikovať“. Aplikácia nám zobrazí správu, v ktorej sú pre nás podstatné informácie o tom, s kým sme sa zhodli, či je identifikátor náš a hodnota skóre zhody, ktorá je pre naše merania najdôležitejšia. Po testovaní zapíšeme každú nameranú hodnotu do správnej bunky v tabuľke, ako je znázornené na Tab. 1.

Tabuľka 1
Tabuľka s nameranými hodnotami skóre podobnosti uloženého a zosnímaného odtlačku prstu

Snímače	Rm	Ra	Ri	L3	L1
BioMini Slim 2S	364	283	376	296	260
GreenBit DactylID20	238	364	378	302	302
U.are.U 4500	232	190	324	271	234
EikonTouch 510	288	287	361	283	277
EikonTouch 710	279	258	275	341	312

Výsledné hodnoty skóre zhody sú pre našu prácu najdôležitejšie, pretože vďaka nim môžeme zistiť spoľahlivosť a kvalitu snímania. Žlté polia znamenajú, že daný prst bol daným snímačom zapísaný do databázy, takže môžeme povedať, že pre danú testovanú osobu pri testovaní tým istým snímačom trikrát z piatich možných prípadov mal tento snímač najvyššie skóre zhody. Tabuľka obsahuje celkovo 25 nameraných hodnôt, pretože máme 5 externých snímačov odtlačkov prstov a pomocou každého externého snímača odtlačkov prstov sme vykonali 5 meraní.

Hodnoty v tejto tabuľke môžu byť napr:

- Minimálna hodnota = 190
- Maximálna hodnota = 364
- Priemerná hodnota = 295

Aby sme mohli z meraní vyvodit' závery, potrebujeme medzi sebou porovnávať len konkrétne stĺpce, t. j. namerané hodnoty z jedného konkrétneho prsta s použitím piatich rôznych externých snímačov odtlačkov prstov.

Zameriame sa na hodnoty v stĺpcoch, ktoré sa budú výrazne líšiť od ostatných hodnôt pre daný odtlačok prsta, tieto hodnoty sa nazývajú odľahlé hodnoty [6]. Tieto hodnoty v biometrii predstavujú údaje, ktoré sa výrazne odlišujú od ostatných údajov v danom súbore meraní. Tieto odľahlé hodnoty môžu byť spôsobené chybou merania, nesprávnym zaznamenaním údajov alebo inými faktormi, ktoré ovplyvňujú merania [7]. Ako príklad môžeme ilustrovať už spomínanú tabuľku a ak sa zameriame len na 2. stĺpec, uvidíme, že jedna z hodnôt je v porovnaní s ostatnými pomerne nízka a môžeme ju nazvať odľahlou hodnotou. Daná hodnota bola nameraná snímačom *U.are.U 4500* a k výslednému vyhodnoteniu pre daný snímač pripočítame +1 hodnotu odľahlej (outliers) hodnoty. Senzor s najnižšou hodnotou odľahlých (outliers) hodnôt by mal byť najspoľahlivejší.

IV. VYHODNOTENIE VYKONANÝCH TESTOV

Cieľom porovnania externých snímačov odtlačkov prstov môže byť zistenie, ako rôzne faktory, ako napríklad použitá technológia snímania, kvalita snímača, spoľahlivosť a presnosť snímania a rozpoznávania odtlačkov prstov, ovplyvňujú odtlačok prsta [8]. V tomto článku sa zameriavame predovšetkým na to aby sme zistili, ktorý snímač je najlepší, a tiež aby sme

porovnali dve technológie snímania odtlačkov prstov, ktorými sú optické a kapacitné snímanie odtlačkov prstov.

A. Meranie s najnižším priemerným skóre zhody

Jednou zo zaujímavých hodnôt je najnižšie priemerné skóre zhody. Priemerné skóre zhody celej tabuľky sa vypočítalo spočítaním všetkých hodnôt a vydelením počtom meraní. Najnižšia priemerná hodnota celej tabuľky bola 197,64 a v celej tabuľke boli tri odľahlé hodnoty. Takéto malé priemerné skóre zhody znamená, že daný testovaný subjekt má na prstoch relatívne málo značiek (Tab. 2).

Tabuľka 2
 Tabuľka s najnižším priemerným skóre zhody (197,64)

Snímače	Rm	Ra	Ri	L3	L1
BioMini Slim 2S	52	189	308	131	320
GreenBit DactylID20	97	192	330	177	274
U.are.U 4500	164	128	268	201	285
EikonTouch 510	77	87	172	184	350
EikonTouch 710	175	230	130	216	204

B. Meranie s najvyšším priemerným skóre zhody

Ďalšou zaujímavou hodnotou je opačná hodnota ako predchádzajúca, a teda najvyššia priemerná hodnota tabuľky zhodných výsledkov. Priemerná hodnota skóre zhody celej tabuľky bola vypočítaná súčtom všetkých hodnôt a vydelením počtom meraní.

Najvyššia priemerná hodnota celej tabuľky bola 415,68 a v celej tabuľke boli 4 odľahlé hodnoty. Keďže v tabuľke boli až 4 odľahlé hodnoty, ktoré sú výrazne nižšie ako ostatné hodnoty, skutočná priemerná hodnota by mohla byť ešte vyššia ako nami vypočítaná hodnota. Takáto vysoká priemerná hodnota skóre zhody znamená, že testovaná osoba má na prstoch veľa markantov a prsty osoby je ťažšie napodobniť a sfalšovať (Tab. 3).

Tabuľka 3
 Tabuľka s najvyšším priemerným skóre zhody (415,68)

Snímače	Rm	Ra	Ri	L3	L1
BioMini Slim 2S	412	481	326	435	448
GreenBit DactylID20	444	426	311	538	542
U.are.U 4500	366	450	535	435	284
EikonTouch 510	445	358	251	338	488
EikonTouch 710	445	464	113	421	636

C. Meranie s najvyšším priemerným skóre zhody

Ďalšou dôležitou hodnotou je najvyšší počet odľahlých hodnôt. Hodnoty odľahlých hodnôt sú v tabuľke zobrazené ako oranžové bunky na pozadí na Tab. 4. Zároveň však vidíme, že dve z piatich odľahlých hodnôt sa nachádzajú na pozícii snímača, pomocou ktorého bola vzorka zapísaná do databázy. V dôsledku toho môžeme povedať, že z 25 meraní sa za odľahlé hodnoty považuje 5 hodnôt, čo predstavuje 20 %. Tieto hodnoty použijeme na záverečné vyhodnotenie meraní, kde vyhodnotíme nepresnosť, spoľahlivosť a ďalšie vlastnosti, ktoré môžeme potvrdiť vďaka týmto odľahlým hodnotám. Tieto hodnoty by sa pravdepodobne dali odstrániť opakovaným meraním, ale v tomto článku sa zaoberáme týmito hodnotami v prvom meraní.

Tabuľka 4
Tabuľka s najvyšším počtom odľahlých hodnôt (5/25)

Snímače	Rm	Ra	Ri	L3	L1
BioMini Slim 2S	256	409	499	329	475
GreenBit DactyID20	384	360	487	288	390
U.are.U 4500	333	382	394	260	482
EikonTouch 510	327	271	439	214	272
EikonTouch 710	285	390	449	202	396

V. ZÁVER

Cieľom tohto článku bolo vykonať analýzu požiadaviek na implementáciu vybraného systému identifikácie a/alebo verifikácie používateľa z odtlačkov prstov na platforme Android a navrhnúť koncepciu jeho vylepšení s dôrazom na používateľskú skúsenosť, dostupné hardvérové riešenia a skúmanie ľudského faktora pri používaní systému.

Pri testovaní týchto snímačov sme použili metódu, pri ktorej sme merali hodnoty skóre zhody pre porovnanie odtlačkov prstov a tieto hodnoty sme zaznamenávali do tabuľky. Vo výslednej tabuľke sme hľadali odľahlé hodnoty, ktoré sme označili a nakoniec sčítali pre každý jednotlivý snímač, čo nás viedlo k výsledkom, že optické snímače sú lepšie ako kapacitné snímače a že najlepšimi snímačmi v tomto type merania sú *BioMini Slim 2S* a *DactyID20*.

POĎAKOVANIE

Výskum v tomto príspevku bol čiastočne podporený Vedeckou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskou akadémiou vied v rámci projektu VEGA 2/0165/21, čiastočne z projektu ITMS 313011W988 financovaného z operačného programu ERDF, a čiastočne v rámci projektov podporených Agentúrou na podporu výskumu a vývoja APVV-SK-TW-21-0002, APVV-22-0261 a APVV-22-0414.

LITERATÚRA

- [1] A. K. Jain, A. Ross and S. Prabhakar, "An introduction to biometric recognition," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 1, pp. 4-20, Jan. 2004, doi: 10.1109/TCSVT.2003.818349.
- [2] F. Lehmann, D. Buschek, "Heartbeats in the wild: a field study exploring ECG biometrics in everyday life." In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-14). doi: 10.1145/3313831.3376536
- [3] L. O'Gorman, "Comparing passwords, tokens, and biometrics for user authentication," in Proceedings of the IEEE, vol. 91, no. 12, pp. 2021-2040, Dec. 2003, doi: 10.1109/JPROC.2003.819611.
- [4] A. Jain, L. Hong, & S. Pankanti, S. (2000). Biometric identification. Communications of the ACM, 43(2), 90-98.
- [5] S. Prabhakar, S. Pankanti and A. K. Jain, "Biometric recognition: security and privacy concerns," in IEEE Security & Privacy, vol. 1, no. 2, pp. 33-42, March-April 2003, doi: 10.1109/MSECP.2003.1193209.
- [6] K. Challita, H. Farhat and K. Khaldi, "Biometric Authentication for Intrusion Detection Systems," 2010 First International Conference on Integrated Intelligent Computing, Bangalore, India, 2010, pp. 195-199, doi: 10.1109/ICIIC.2010.15.
- [7] B. Miller, "Vital signs of identity [biometrics]," in IEEE Spectrum, vol. 31, no. 2, pp. 22-30, Feb. 1994, doi: 10.1109/6.259484.
- [8] J. Wayman, A. Jain, D. Maltoni, D. Maio, "An introduction to biometric authentication systems." Biometric systems: Technology, Design and performance evaluation, Springer, London. pp. 1-20. doi:10.1007/1-84628-064-8_1

Mobilná aplikácia pre biometrické bezpečnostné systémy

¹Renát HALUŠKA, ²Dominik MOLČANYI, ³Matúš PLEVA

^{1,2,3} Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹renat.haluska@tuke.sk, ²dominik.molcanyi.student@tuke.sk, ³matus.pleva@tuke.sk

Abstrakt — Tento článok opisuje rôzne biometrické metódy používané v biometrických systémoch a ich vlastnosti. Opisuje metódu dynamiky stlačenia klávesov v biometrii a informuje čitateľa o súčasných trendoch využívajúcich túto metódu a o implementácii dynamiky stlačenia klávesov v rámci webovej aplikácie, ktorá sa používa ako súčasť výuky v predmete biometrické systémy bezpečnosti.

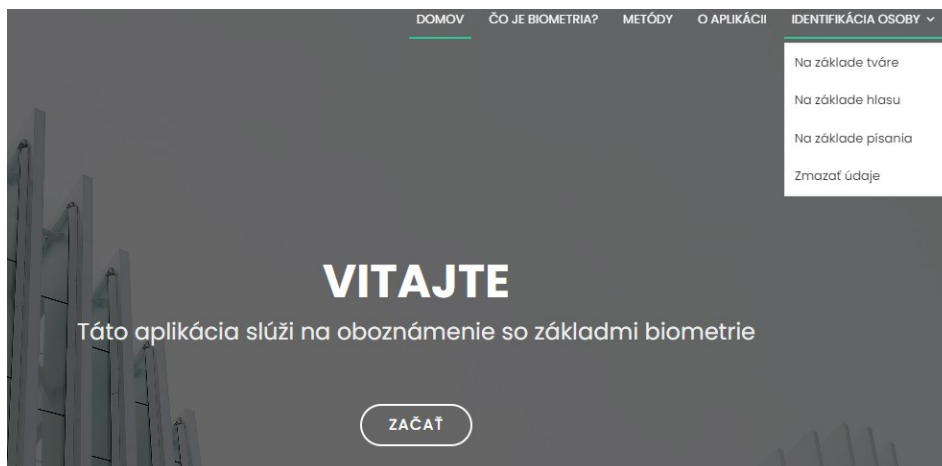
Kľúčové slová — Autentifikácia, Biometria, Biometrické systémy, Dynamika stlačenia klávesov, Identifikácia, Výkonnostné metriky, Webová aplikácia

I. ÚVOD

Mobilná ukázková aplikácia je softvérová aplikácia, ktorá beží na webovom serveri Apache a je vytvorená pomocou HTML, CSS, JavaScript a PHP [1]. Používatelia môžu k tejto aplikácii pristupovať prostredníctvom webového prehliadača s pripojením na školský server. Táto mobilná demonštračná aplikácia sa používa na interaktívnu výučbu medzi učiteľmi a študentmi a je zameraná na vzdelávanie v predmete Biometrické bezpečnostné systémy [2,3]. Vzhľadom na veľké množstvo informácií, ktoré denne získavame, je dôležité tieto informácie nielen prijímať, ale im aj rozumieť [4]. Preto táto webová vzdelávacia aplikácia umožňuje študentom oboznámiť sa s biometrickými metódami, ako je rozpoznávanie tváre, rozpoznávanie hlasu alebo najnovšie rozpoznávanie pomocou dynamiky stlačenia klávesov Implementácia poslednej biometrickej metódy, dynamiky stlačenia klávesov, je úlohou tejto práce a nasledujúcich častí [5].

II. WEBOVÁ APLIKÁCIA

Webová stránka je umiestnená na školskom serveri s webovým serverom Apache. Na Obr. 1 je zobrazená úvodná stránka, keď používateľ vstúpi na stránku. Používateľovi sa zobrazí uvítacia obrazovka. Keď používateľ klikne na tlačidlo "Začať", na stránke sa zobrazia teoretické informácie o biometrii. Na uvítacej obrazovke je možné si všimnúť aj ponuku možností. Ak chceme použiť dynamiku stlačenia klávesov (Keystroke Dynamics - KD) [6], stačí prejsť myšou na časť "Identifikácia osoby" a potom kliknúť na možnosť "Na základe písania".



Obr. 1 Domovská stránka webovej aplikácie

A. Rozpoznávanie pisateľa

Používateľa možno rozpoznať na základe jeho písania na klávesnici, preto sa použije biometrická metóda dynamiky stlačenia klávesov (KD). Táto funkcia sa použije, ak používateľ vyberie možnosť s názvom "Identifikácia osoby na základe písania", ktorá sa nachádza v ponuke. Vzhľad domovskej stránky po výbere uvedenej možnosti je vidieť na Obr. 2.



Obr. 2 Domovská stránka pre rozpoznávanie písania

Na úvodnej stránke sa používateľovi zobrazí slovo, ktoré sa použije ako vzor na rozpoznávanie, v tomto prípade je to "password", a textové pole s počítadlom vpravo. Do textového poľa používateľ zadá určené slovo a po správnom zadaní slova a stlačení klávesu *Enter* sa textové pole vymaže a počítadlo sa zvýši o jednotku. Počítadlo zobrazuje (Obr. 3) počet správne napísaných slov. Po zadaní každého znaku do textového poľa si používateľ môže v tabuľke pod úvodnou stránkou zobrazit' relatívny čas v milisekundách pre každý znak. Každý správne napísaný znak, ktorý zodpovedá zadanému slovu, je označený zelenou farbou. V opačnom prípade je označený bielou farbou.

n	[key] Down _n	[key] Up _n	Up _n - Down _n	Down _{n+1} - Up _n
1	[p] 1449.500	[p] 1574.500	125.000	561.100
2	[a] 2135.600	[a] 2314.700	179.100	1161.800
3	[s] 3476.500	[s] 3558.600	82.100	282.900
4	[s] 3841.500	[s] 3937.600	96.100	1261.000
5	[w] 5198.600	[w] 5326.500	127.900	1986.000
6	[r] 7312.500	[r] 7468.600	156.100	55.900
7	[e] 7524.500	[e] 7632.700	108.200	91.700
8	[d] 7724.400	[d] 7836.600	112.200	0.000

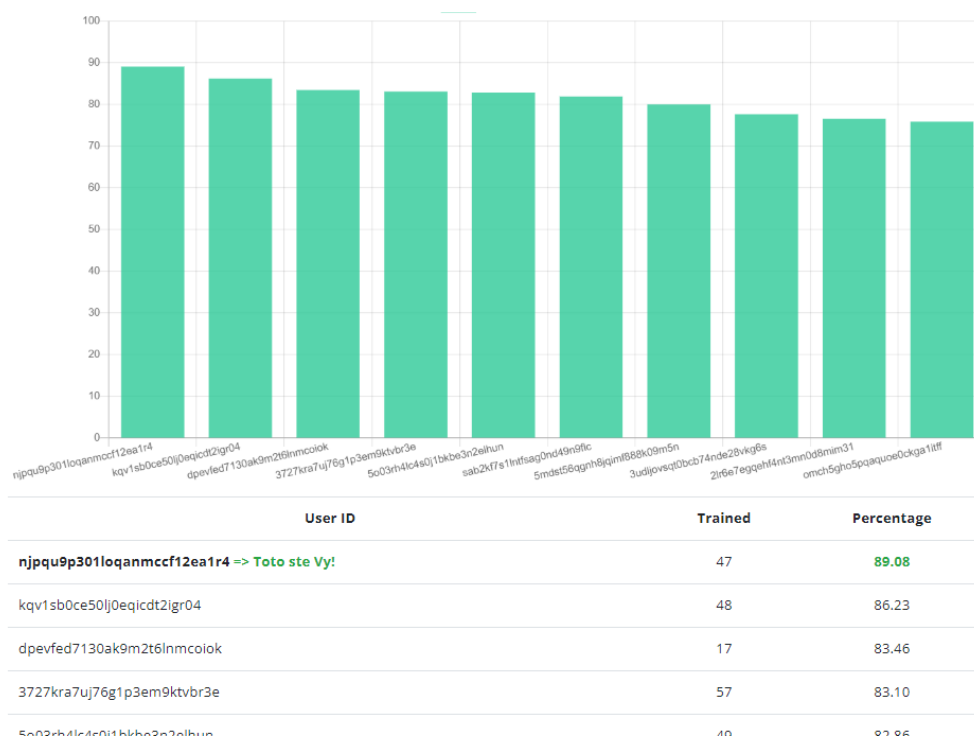
Obr. 3 Tabuľka s časovými údajmi

Tabuľka na obrázku 3 má 5 stĺpcov:

- Stĺpec č. 1 - určuje poradie znaku
- Stĺpec č. 2 - udáva, o ktorý znak išlo a jeho relatívny čas stlačenia v milisekundách
- Stĺpec č. 3 - udáva, o aký znak išlo, a jeho relatívny čas stlačenia v milisekundách
- Stĺpec č. 4 - udáva hodnotu Dwell time
- Stĺpec č. 5 - uvádza hodnotu času Latency

Ak používateľ aspoň raz zadá slovo "password", môže sa použiť funkcia rozpoznávania. Ak používateľ zadá slovo viac ako desaťkrát, zobrazí sa tlačidlo uložiť do databázy. Obe možnosti sú opísané nižšie. Pri vstupe na webovú stránku sa pre každého používateľa vygeneruje náhodný identifikátor, ktorý sa uloží do súborov cookie. Pri zmene počítača alebo vymazaní súborov cookie sa identifikátor zmení.

Ak používateľ aspoň raz zadá slovo "password" a klikne na tlačidlo "Rozpoznanie", na pozadí sa spustí funkcia rozpoznávania, ktorá zhromažďuje všetky "zelené", t. j. správne zadané znaky, ich časové značky a porovnáva ich so všetkými údajmi v databáze. Keď sa funkcia dokončí a rozpoznávanie je úspešné, používateľovi sa zobrazí graf a tabuľka výsledkov, ako je vidieť na Obr. 4.

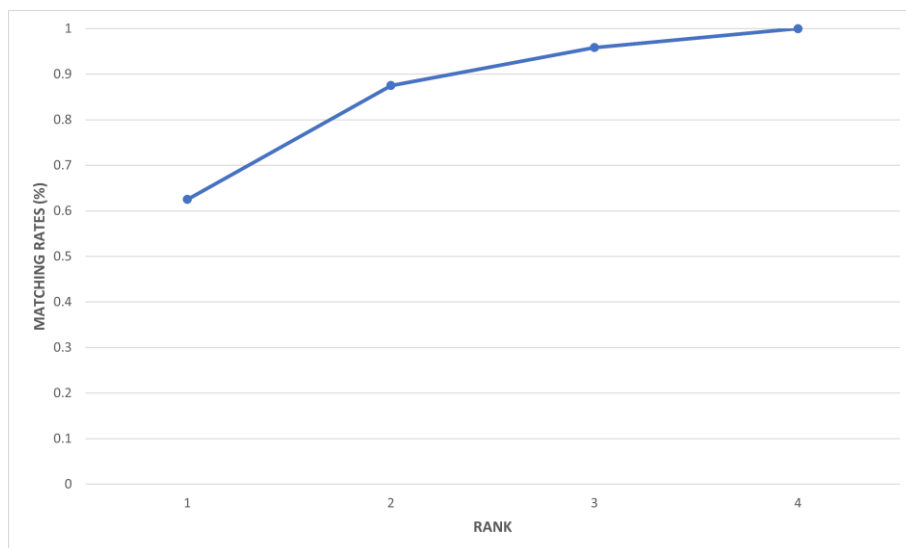


Obr. 4 Príklad opisu výsledkov

Na Obr. 4 možno vidieť graf, kde os y zobrazuje percento zhody a os x zobrazuje ID používateľov, ktorí boli uložení v databáze a s ktorými je najvyššia zhoda, maximálne do desiatich najlepších. Pod grafom sa nachádza tabuľka, z ktorej môžeme dodatočne vyčítať hodnotu *Trained*, ktorá ukazuje, koľko slov používateľ zadal do databázy, a teda z koľkých slov bol zostavený jeho model.

III. HODNOTENIE

Mobilná demonštračná aplikácia bola testovaná na 16 rôznych subjektoch vo vekovej skupine 22 až 49 rokov. Každý subjekt testoval aplikáciu na jednom a tom istom zariadení - notebooku Dell Latitude 7320 so zabudovanou klávesnicou. Každý subjekt dostal pokyn písať čo najprirodzenejšie a zároveň napísať minimálne desať slov. Celkovo bolo zozbieraných 24 identifikačných údajov (dosiahnutá pravdepodobnosť riešenia), pretože niektorí to vyskúšali viackrát z nasledujúcich dôvodov. Subjekt po napísaní 10 slov nebol nájdený na prvom mieste, a tak zadal a uložil ďalších 10 slov, kde už bola zhoda. Prípadne ľudia vyskúšali aplikáciu viackrát, pretože na prvom mieste bola nájdená zhoda a chceli to skúsiť znova. Na vyhodnotenie úspešnosti identifikácie bola vytvorená CMC krivka (Obr. 5).



Obr. 5 CMC krivka

Uvedená krivka CMC zobrazuje kumulatívnu presnosť identifikácie podľa poradia (do K-teho miesta) medzi najpodobnejšími modelmi (Top-K, Rank-K) s rôznymi pravdepodobnosťami identifikácie. Na základe uvedených hodnôt je vidieť, že systém dokáže rozpoznať používateľa s najlepším výsledkom v porovnaní so 17 modelmi v databáze s pravdepodobnosťou 62,5 %. Tento výsledok naznačuje, že systém je schopný rozpoznať jednotlivú osobu na základe jej jedinečnej dynamiky stlačenia klávesov s pomerne vysokou presnosťou. Pravdepodobnosť identifikácie mohla byť vyššia, ak by subjekty uložili do databázy viac údajov alebo napísali zamýšľané slovo viackrát.

Tabuľka 1 obsahuje tri stĺpce. V prvom stĺpci je uvedené meno osoby s vekom, prípadne prvé písmeno priezviska, ak sa mená opakujú. V druhom stĺpci je uvedené miesto vo výslednej tabuľke a v poslednom stĺpci je uvedený počet slov napísaných pred identifikáciou.

Tabuľka 1
Vybrané hodnoty získané z testovania

Name, age	Location (identification partition)	Number of stored words
Dávid, 24	TOP 1	66
Zuzana, 22	TOP 1	57
Gabriel, 24	TOP 1	53
Tomáš, 24	TOP 1	49
Erika, 23	TOP 1	48
Dominik M., 24	TOP 1	47
Erik, 24	TOP 1	46
Dominik H., 24	TOP 1	36
Lubomír, 24	TOP 1	33
Miroslav, 24	TOP 1	27
Samuel, 26	TOP 1	17
Dominik D., 24	TOP 2	15
Alexandra, 48	TOP 2	14
Rastislav, 49	TOP 2	13
Tomáš, 36	TOP 2	10
Matej, 39	TOP 3	12
Anna, 31	TOP 4	14

Na základe Tabuľky 1 môžeme tvrdiť, že ak subjekt zadá viac ako 17 slov, aplikácia je schopná rozpoznať subjekt na 1. priečke. Aby sme túto pravdepodobnosť zvýšili, musíme napísať a uložiť do databázy viac slov. Ak subjekt napísal do databázy menej ako 17 slov, systém ho už nie je schopný identifikovať na priečke 1. Z Obr. 5 je vidieť, že 60,25 % subjektov bolo identifikovaných správne, takže systém dokázal identifikovať 15 subjektov z 24 testovaných na priečke číslo 1 (TOP 1). Tento výsledok môže byť spôsobený rôznymi faktormi, ako sú chyby pri zaznamenávaní údajov, veľmi podobná dynamika stlačenia klávesov medzi rôznymi subjektmi alebo subjekty mali na porovnanie zadaný malý počet údajov.

IV. ZÁVER

Navrhované riešenie založené na vhodných technológiách vzhľadom na dostupné hardvérové riešenia bolo overené v experimentálnej prevádzke na vzorke minimálne 15 používateľov. Z výsledkov vyplýva, že biometrický systém dosiahol veľmi dobré výsledky na krivke CMC s hodnotami pre Rank 1 62,5%, pre Rank 2 88% a pre Rank 3 96%. Z našej prieskumnej štúdie vyplýva, že polovica respondentov pozná metódu biometrickej dynamiky stlačenia klávesov a väčšina z nich považuje túto metódu za bezpečné riešenie. Získané výsledky sú veľmi povzbudivé a ukazujú, že biometrické systémy založené na dynamike stlačenia klávesov majú potenciál byť v kombinácii s iným biometrickým znakom v multimodálnom systéme bezpečným a efektívnym riešením na identifikáciu používateľov v rôznych odvetviach, ako je zdravotníctvo, verejná správa a priemysel.

V budúcnosti by bolo zaujímavé otestovať dynamiku stlačenia klávesov na väčšej vzorke respondentov, s väčším súborom údajov, kde by sa údaje zbierali v dlhšom časovom intervale. Bolo by tiež zaujímavé zahrnúť rôzne scenáre a podmienky, v ktorých by takéto testovanie mohlo priniesť lepšie a presnejšie výsledky. Aplikácia by sa mohla z pohľadu bezpečnosti osobných dát vylepšiť tak, aby sa údaje z databázy vymazali alebo aby sa na zlepšenie výsledkov identifikácie

použilo strojové učenie.

POĎAKOVANIE

Výskum v tomto príspevku bol čiastočne podporený Vedeckou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskou akadémiou vied v rámci projektu VEGA 2/0165/21, čiastočne z projektu ITMS 313011W988 financovaného z operačného programu ERDF, a čiastočne v rámci projektov podporených Agentúrou na podporu výskumu a vývoja APVV-SK-TW-21-0002, APVV-22-0261 a APVV-22-0414.

LITERATÚRA

- [1] N. Litayem, Dhupia, B., & Rubab, S. (2015). Review of cross-platforms for mobile learning application development. *International Journal of Advanced Computer Science and Applications*, 6(1).
- [2] A.K. Jain, Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, 14(1), 4-20.
- [3] M. Raj, & Mahajan, T. (2021). Study of Opinions Related to the Use of Bio-Enabled Attendance System for Improving Punctuality Among Teachers. *MIER Journal of Educational Studies Trends and Practices*, 41-52.
- [4] T.H. Davenport, & Beck, J. C. (2001). The attention economy. *Ubiquity*, 2001(May), 1-es.
- [5] P.S. Teh, Teoh, A. B. J., & Yue, S. (2013). A survey of keystroke dynamics biometrics. *The Scientific World Journal*, 2013.
- [6] M. Bauer, (2014). Using Metrics and Keystroke Dynamics in Games to measure Player Experience.

Bunkové siete na báze bezpilotných lietajúcich prostriedkov

¹Peter JANDA, ¹Renát HALUŠKA, ¹Stanislav MARCHEVSKÝ,
²Slavomír PILLÁR

¹ Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika
² TELESERVIS spol. s r.o., J. Hagaru 9, Bratislava - mestská časť Rača 831 51

¹peter.janda@student.tuke.sk, ¹renat.haluska@tuke.sk, ¹stanislav.marchevsky@tuke.sk,
²slavomir.pillar@gmail.com

Abstrakt — Článok obsahuje analýzu vlastností bunkových sietí a možnosti využitia bezpilotných lietajúcich prostriedkov (UAV). V článku je vypracovaný prehľad lacných programovateľných dronov ako prvku dronovej siete. Vybrané UAV sú vhodné pre teoretickú ale aj praktickú výučbu návrhu dronových sietí.

Kľúčové slová — ad-hoc siete, bezpilotné lietajúce prostriedky, mobilné siete, siete UAV

I. ÚVOD

Bunková sieť je rádiová sieť distribuovaná po zemi cez bunky, kde každá bunka obsahuje vysielač/prijímač s pevným umiestnením známym ako základňová stanica (z angl. Ground Control Station, GCS), má základňovú stanicu analogickú s 802.11 AP, ktorá pomáha mobilným používateľom pripojiť sa k sieti. Medzi mobilnou a základňovou stanicou existuje vzdušné rozhranie fyzického protokolu a protokolu spojovej vrstvy. Tieto bunky spoločne poskytujú rádiové pokrytie väčších geografických oblastí. Používateľské zariadenie (z angl. User Equipment, UE), ako sú mobilné telefóny, drony, sú preto schopné komunikovať, aj keď sa zariadenie počas prenosu pohybuje medzi rôznymi bunkami [1].

II. BUNKOVÉ SIETE

Bunkové siete ponúkajú rozsiahlu, vysokorychlostnú a bezpečnú bezdrôtovú konektivitu v rámci licencovaného spektra, čo môže výrazne zlepšiť kontrolu a bezpečnosť operácií malých UAV a umožniť prevádzku za hranicu viditeľnosti (z angl. Beyond the Visual Line Of Sight, BVLOS). Bunková architektúra má niekoľko výhod:

- pokrytie môže byť rozšírené na veľké oblasti viacerými GCS, ktoré poskytujú prirodzenú redundanciu takým spôsobom, že ak je jedno spojenie slabé, iné spojenie môže fungovať lepšie,
- bunkové siete sú v súčasnosti schopné podporovať vysokú priepustnosť,
- zariadenie koncového užívateľa (UE) je malé, spotrebuje nízku energiu a ťaží z hromadnej výroby, ktorá môže zvýšiť jeho využitie v praktických implementáciách UAV,
- znižuje rušenie inými signálmi [2], [3], [4].

A. Bunkové siete na báze UAV

Operátori mobilných sietí (MNO) navrhujú svoje siete pre pokrytie a optimalizáciu capacity a preto sú bunkové siete optimalizované pre pozemnú širokopásmovú komunikáciu, kde sú BS (z angl. Base Station, BS) antény naklonené smerom dole, aby sa znížilo rušenie iných buniek. Naklápanie je možné mechanicky aj elektricky, ale zvyčajne sa využíva na pokrytie a optimalizáciu capacity.

Existujúce UAS (z angl. Unmanned Aircraft Systems, UAS) sa však spoliehajú hlavne na jednoduchú komunikáciu z bodu do bodu (z angl. point-to-point) komunikáciu cez nelicencované

pásma (napr. ISM 2,4 GHz), ktoré má nízku dátovú rýchlosť, je nespoľahlivé, nezabezpečené, náchylné na rušenie, je ťažké ho legitímne monitorovať a spravovať a môže fungovať len vo veľmi obmedzenom rozsahu. Keď, že sa očakáva, že počet UAV a ich aplikácií bude v nadchádzajúcich rokoch ďalej rásť, je nevyhnutné vyvinúť nové bezdrôtové technológie, ktoré umožnia výrazne zlepšiť komunikáciu UAV smerom k zemi.

Bunkovo prepojené UAV je sľubná technológia na dosiahnutie vyššie uvedeného cieľa, pričom UAV pre rôzne aplikácie sú integrované do existujúcich a budúcich bunkových sietí ako nové letecké používateľské zariadenia (UE). V porovnaní s tradičnou komunikáciou zem-UAV cez spojenia bod-bod má UAV s bunkovým spojením niekoľko atraktívnych výhod [5], [6].

B. UAV Siete

Bezpilotné lietajúce prostriedky sú komplexné letecké platformy určené na bezdrôtové snímanie a ovládanie. Používanie UAV, najmä kvadrotorov (z angl. quadrotor), preniklo do všetkých aspektov ľudského života. UAV môžu byť efektívne využité na skúmanie mestských oblastí vrátane budov, infraštruktúry, prírodných oblastí a zdravia životného prostredia. Výzvy, ako napríklad zlé poveternostné podmienky, obmedzená alebo úplna absencia sieťového pripojenia, obmedzený vizuálny dosah, nedostupnosť GPS signálu a iné podobné problémy sú typické pre používanie UAV, najmä väčšej skupiny (roj) UAV, pri úlohách typu vyhľadávanie a sledovanie [7].

Porovnanie rôznych bunkových sietí

Tabuľka 1
 Porovnanie rôznych bunkových sietí

	MANET	VANET	UAV sieť
Popis	Využívajú bezdrôtové spojenie na pripojenie sa do rôznych sietí	Ad-Hoc siete, v ktorých sú vozidlá mobilnými uzlami. Komunikácia prebieha medzi vozidlami a medzi vozidlami a mimocestných jednotiek	Vzdušné uzly založené na infraštruktúre alebo Ad-Hoc. Komunikácia prebieha medzi UAV a kontrolnou stanicou
Mobilita	Pomalá. Typická rýchlosť je 2m/s. Náhodný pohyb. Meniaci sa hustota	Rýchla. Typická rýchlosť je 20-30 m/s na diaľniciach, 6-10 m/s v mestských oblastiach. Prediktívny, limitovaný cestnou infraštruktúrou a pravidlami ciest.	Rýchlosti od 0 až po 100 m/s. Pohyb môže byť 2 rozmerný alebo 3 rozmerný, záleží od typu misie.
Topológia	Náhodná, Ad-Hoc	Hviezda s cestnou infraštruktúrou a Ad-Hoc medzi vozidlami	Hviezda s kontrolnou stanicou, Ad-Hoc/Mesh medzi UAV.
Zmeny v topológii	Dynamická – uzly sa nepredvídateľne pripájajú a odpájajú. Sieť náchylná na rozdeľovanie	Viac dynamická ako MANET. Lineárny pohyb. Rozdeľovanie siete.	Stojaca, pomalá alebo rýchla.
Typ využívanej energie	Väčšina uzlov je napájaná pomocou batérií, takže energia sa musí šetriť	Zariadenia môžu byť napájané svojou batériou.	Malé UAV sú obmedzené energiou. Batérie ovplyvňujú hmotnosť a čas lietania
Typické využitie vo verejnosti	Distribúcia informácií	Informácie o ceste a počasí, núdzové výstrahy	Záchranne operácie Poľnohospodárky prieskum Pozorovanie rôznych oblastí

C. Výhody UAV sietí

Vďaka takmer všade prítomnej dostupnosti celulárnych sietí na celom svete umožňujú bunkovo pripojené UAV pozemnému pilotovi diaľkovo ovládať a ovládať (z angl. Command and Control, C&C) UAV v podstate neobmedzeným operačným dosahom. Poskytuje tiež efektívne riešenie na udržiavanie bezdrôtového pripojenia medzi UAV a rôznymi inými zainteresovanými stranami, ako sú koncoví používatelia a riadiaci letovej prevádzky, bez ohľadu na ich

umiestnenie. Napríklad využitím bunkovej siete môžu byť živé videá priamo odosielané z UAV vzdialenému publiku po celom svete.

S pokročilými celulárnymi technológiami a mechanizmami autentifikácie majú celulárne pripojené UAV potenciál dosiahnuť výrazné zlepšenie výkonu oproti jednoduchej priamej komunikácii medzi zemou a UAV, pokiaľ ide o spoľahlivosť, bezpečnosť a komunikačnú priepustnosť.

Tradičná UAV navigácia sa spolieha najmä na satelitnú sieť, akou je napríklad systém GPS (z angl. Global Positioning System), ktorý je však náchylný na rušenie satelitných signálov, napríklad v dôsledku blokovania vysokými budovami alebo zlých poveternostných podmienok. Mobilné UAV ponúkajú jednu účinnú metódu, medzi inými, ako je diferenciálny GPS (D-GPS), na dosiahnutie robustnejšej UAV navigácie využitím mobilných signálov ako doplnku k GPS navigácii.

V neposlednom rade sú bunkovo pripojené UAV aj cenovo výhodné. Na jednej strane môžu znovu použiť milióny celulárnych základňových staníc (BS), ktoré sú už nasadené po celom svete, bez potreby budovania novej infraštruktúry vyhradenej len pre UAS, čím výrazne ušetria náklady na nasadenie siete. Na druhej strane to môže tiež pomôcť ušetriť prevádzkové náklady prostredníctvom spájania UAV C&C a iných početných typov komunikácií užitočného zaťaženia do bunkových systémov, čo vytvorí nové obchodné príležitosti pre mobilných operátorov aj operátorov UAV [8].

III. PREHLAD LACNÝCH A PROGRAMOVATELNÝCH DRONOV

Programovateľné drony sú čoraz populárnejšie, pretože ponúkajú veľké množstvo výhod pre spotrebiteľov aj podniky. Napríklad programovateľné drony možno použiť na fotografovanie, natáčanie videí a dokonca aj na mapovanie. Okrem toho je možné programovateľné drony naprogramovať tak, aby autonómne sledovali konkrétnu dráhu letu, vďaka čomu sú ideálne na použitie pri pátracích a záchranných misiách alebo iných nebezpečných prostrediach.

Medzi populárne modely patria DJI Tello, Ryze Tello a Parrot Mambo. Tieto drony sú lacné, prenosné a poskytujú možnosti programovania cez ich vlastné softvérové rozhranie alebo iné platformy ako napríklad Python. Okrem toho existujú aj programovateľné drony ako DJI Mavic Mini a DJI Phantom, ktoré sú vybavené kvalitnými kamerami a senzormi. Tieto drony sú schopné letieť na dlhšie vzdialenosti a vysoké výšky, čo umožňuje pokrytie väčšieho územia. Všetky tieto drony majú tiež rôzne bezpečnostné prvky, ktoré pomáhajú zabrániť nehodám, ako napríklad prekážkové senzory a GPS navigáciu. Pre ich použitie v dronovej sensorovej sieti je možné ich vybaviť rôznymi senzormi a kamerami, čím sa zvyšuje ich schopnosť zberu dát a použitia v rôznych aplikáciách [9].

A. DJI Matrice 300 RTK

DJI Matrice 300 RTK je kvalitný dron pre tých, ktorí potrebujú dlhý čas letu a dobrý dosah. Je tiež ideálny pre každého, kto chce používať dron na fotografovanie alebo natáčanie videa. Jedinou nevýhodou je, že je dosť veľký a ťažký, čo môže sťažiť prepravu. DJI Matrice 300 RTK má letový čas 55 minút. Batéria TB60 má kapacitu 9600 mAh, čo je výrazne vyššia ako pri väčšine batérií dronov na trhu. To umožňuje dronu zostať vo vzduchu oveľa dlhšie. DJI Matrice 300 RTK má letový čas 55 minút. Pri uvažovaní jeho dosahu musíme myslieť na dve veci, dosah ovládania a dosah streamovania videa v reálnom čase. Dron má dosah ovládania 15 km a dosah streamovania videa v reálnom čase 10 km. Takže budeme môcť letieť s dronom pomerne ďaleko a stále mať nad ním kontrolu. Rýchlejšie drony sa zvyčajne používajú na sút'áže alebo na získanie úžasných záberov. DJI Matrice 300 RTK má však maximálnu rýchlosť 23 m/s, čo je dostatočne rýchle pre väčšinu použití. Tento dron je viac zameraný na stabilitu a čas letu ako na rýchlosť. Kamera dronu má rozlíšenie 960p a dokáže nahrávať pri 30 fps (z angl. frames per second, fps). Má tiež 145° zorné pole, čo je veľmi dobre na získanie širokého záberu. Má hmotnosť 6,3 kg [10].

B. DJI Tello

Dron Tello vyrába spoločnosť Ryze Tech a poháňa ho DJI. Má hmotnosť 80 g, rozmery 98 mm × 92,5 mm × 41 mm a 76.2 mm vrtule. Dron je vybavený množstvom senzorov na palube vrátane optického senzora, akcelerometra, gyroskopu, barometra a senzora dosahu Time of Flight (ToF). Pripojenie s dronom je možné vykonať cez WiFi pomocou sprievodného SDK (poskytnutého spoločnosťou DJI). Senzor optického toku, ktorý sa nachádza v spodnej časti dronu smerom k zemi, dokáže vypočítať zmenu súradníc X a Z v reálnom čase. Tento snímač je v podstate obrazový snímač (fotoaparát) s nízkym rozlíšením, ktorý robí snímky v časovom

intervale, porovnáva po sebe idúce snímky a odvodzuje zmenu v každej súradnici. Senzor optického toku pomáha dronu vznášať sa nad daným miestom bez driftu a používa sa aj na výpočet relatívnej rýchlosti dronu. Optický snímač prietoku používa dron iba interne na výpočet rýchlosti a nedá sa k nemu priamo pristupovať. Snímač dosahu ToF, ktorý sa nachádza v spodnej časti dronu smerom k zemi, sa používa na určenie relatívnej výšky dronu od objektu pod ním, čo je nevyhnutné pre stúpanie po schodoch a vyhýbanie sa prekážkam zdola. Pomocou barometra dokáže dron zistiť relatívnu výšku vzhľadom na bod vzletu [11], [12].

C. Parrot Mambo

Minidron Parrot Mambo Fly je stabilný, bezpečný, malý a ľahko sa ovláda. Je vybavený ultrazvukovým senzorom, dole smerujúcou kamerou, tlakovým senzorom a IMU senzorom. Je taktiež vybavený 6-osovým stabilizačným systémom, vďaka čomu sa ľahko lieta a ovláda. Dokáže vykonávať rôzne triky, ako sú flipy, kotúľanie sudom (z angl. barrel roll) a 180 stupňové otočky, vďaka čomu je ideálny pre tých, ktorí sa chcú naučiť lietať s dronmi. Má tiež integrovanú kameru, ktorá zachytáva 720p HD video a fotografie. Dron Parrot Mambo je možné naprogramovať pomocou niekoľkých programovacích jazykov vrátane Scratch, Python a Swift. Táto programovacia schopnosť umožňuje používateľom vytvárať vlastné letové vzory, automatizovať úlohy [13], [14].

D. DJI Mavic Mini

DJI Mavic Mini je malý, ľahký a prenosný dron určený pre začiatočníkov, a každého, kto chce kompaktný a ľahko použiteľný dron. Vyrába ho DJI, čínska technologická spoločnosť, ktorá sa špecializuje na výrobu dronov, a kamier. Dron DJI Mavic Mini váži iba 249 gramov, čo z neho robí jeden z najľahších dronov na trhu. Napriek malým rozmerom je vybavený 3-osovým stabilizovaným fotoaparátom, ktorý dokáže zachytiť 2,7K video a 12-megapixelové fotografie. Fotoaparát je tiež vybavený 4 cm snímačom a dokáže snímať rýchlosťou 30 snímok za sekundu. Dron má maximálnu dobu letu 30 minút a môže letieť až 4 kilometre od ovládača, v závislosti od miestnych predpisov. Má tiež niekoľko inteligentných letových režimov, vrátane QuickShot, ktorý automaticky vytvára filmové zábery a Circle, ktorý obieha okolo konkrétneho objektu.

IV. ZÁVER

Oblasť bezpilotných lietajúcich prostriedkov má veľký potenciál a využitie v rôznych oblastiach. Ukazuje sa, že o tento vývoj v blízkej budúcnosti bude veľký záujem. Dôležité je vytvorenie komerčných nízko nákladových riešení pre mnoho odvetví. Domnievame sa tiež, že je potrebné venovať pozornosť vo výučbe komunikačných predmetov v študijnom programe Počítačové siete aj problematike návrhu ad-hoc sieti s využitím nepilotovaných lietajúcich prostriedkov.

POĎAKOVANIE

Výskum v tomto príspevku bol čiastočne podporený Vedeckou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky APVV-SK-TW-21-0002, APVV-22-0261 a APVV-22-0414 a Slovenskou akadémiou vied v rámci projektu VEGA 2/0165/21 a VEGA 1/0584/20.

V. LITERATÚRA

- [1]. M.Rouse, "Cellular Network", 2011. Dostupné na internete: <https://www.techopedia.com/definition/24962/cellular-network>.
- [2]. H. Marques, P.Marques, J.Ribeiro, T.Alves,L.Pereira, in: 2019 IEEE 24th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD). Experimental Evaluation of Cellular Networks for UAV Operation and Services, Október 2019. ISBN 978-1-7281-1016-5.
- [3]. Cellular Network, 2018. Dostupné na internete : <https://www.geeksforgeeks.org/cellular-networks/>.
- [4]. What is Cellular network or Mobile network?, 2020. Dostupné na internete : <https://www.samsung.com/in/support/mobile-devices/what-is-a-cellular-network-or-mobile-network/>.
- [5]. L.Gupta, R.Jain, G.Vaszkun, "Survey of Important Issues in UAV Communication Networks", IEEE Communications Surveys & Tutorials.vol. 18, no. 2, November 2015. 1123 – 1152 s.
- [6]. Y.Zeng, J.Lyu, R. Zhang, " IEEE Wireless Communications. Cellular-Connected UAV: Potential, Challenges, and Promising Technologies", vol. 26, no. 1, September 2018. 120 – 127 s.
- [7]. G.Geraci, L.G.Giordano, D.Lopez-Perez, A.E.Bjornson, Understanding UAV Cellular Communications: From Existing Networks to Massive MIMO, IEEE Access,vol. 6, November 2018. 67853 – 67865 s.

- [8]. L.Guezouli, K.Barka,A.Djehiche, "UAV's efficient controlled mobility management for mobile heterogeneous wireless sensor networks", Journal of King Saud University - Computer and Information Sciences vol. 34, no. 6, Jún 2022. 2461 – 2470 s.
- [9]. R.J.Gross, "10 Best Programmable Drones", 2023. Dostupné na internete : <https://www.propelrc.com/best-programmable-drones/>.
- [10]. DJI. MATRICE 300 RTK. Dostupné na internete : <https://www.dji.com/sk/matrice-300?site=developer>.
- [11]. S.Pikalov, E.Azaria, S.Sonnenberg, B.Ben-Moshe, A.Azaria, "Vision-Less Sensing for Autonomous Micro-Drones", 2021. Dostupné na internete : <https://www.mdpi.com/1424-8220/21/16/5293>
- [12]. B. Masharipov, "Sledování dronu Ryze Tello", České vysoké učení technické v Praze. Bakalářská práce. 2022.
- [13]. P.M.Casero, "A Vision-based Line Following Method for Micro Air Vehicles", 2021.
- [14]. P. Janda, "Návrh komponentov bunkových sietí na báze bezpilotných lietajúcich prostriedkov (UAV) pre monitorovanie mimoriadnych udalostí". Technická univerzita v Košiciach. Diplomová práca, , 2023

Tvorba korpusu textov pre úlohy detekcie nenávisťných prejavov, ofenzívneho jazyka a analýzy sentimentu

¹Zuzana SOKOLOVÁ, ²Maroš HARAHAUS, ³Ján STAS, ⁴Jozef JUHÁR

^{1,2,3,4}Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

¹zuzana.sokolova@tuke.sk, ²maros.harachus@tuke.sk, ³jan.stas@tuke.sk,
⁴jozef.juhar@tuke.sk

Abstrakt – Detekcia nenávisťných prejavov a analýza sentimentu sú v súčasnosti pomerne vyhľadávané a žiadané úlohy v oblasti spracovania prirodzeného jazyka. Pri práci na takýchto úlohách je nevyhnutné zhromaždiť korpus textov, ktorý slúži na tréning a testovanie jazykových modelov a algoritmov. Pri tvorbe korpusu textov pre úlohy detekcie nenávisťných prejavov alebo analýzy sentimentu je však potrebné dbať na niekoľko dôležitých kritérií. V príspevku sme opísali kritéria pre tvorbu korpusu textov v slovenskom jazyku.

Príležitostné slová – analýza sentimentu, kritéria, nenávisťné prejavy, ofenzívny jazyk, tvorba korpusu textov

I. ÚVOD

V súčasnosti sa na Internete stretávame s veľkým množstvom nenávisť. Ľudia na sociálnych sieťach komentujú, reagujú, diskutujú o roznych témach (napr. politika, klimatická kríza, rovnoprávnosť pohlaví, diskriminačné problémy a podobne). Spolu s tým vzniká veľké množstvo dát a textového obsahu, ktoré je potrebné moderovať a filtrovať na základe etických kódexov atď. Preto je dôležité vedieť nenávisťné prejavy moderovať rýchlo a efektívne.

ElSherief a kol. definujú nenávisťné prejavy ako „priamy a vážny útok na akúkoľvek chránenú kategóriu ľudí na základe ich rasy, etnicity, národného pôvodu, náboženstva, pohlavia, sexuálnej orientácie, zdravotného postihnutia alebo choroby“ [1].

Sociálne siete ako Facebook, Instagram, Youtube alebo Twitter už vedú nenávisťný obsah celkom promptne moderovať. Používajú sa na to rôzne jazykové modely a algoritmy. Detekcia nenávisťných prejavov a ofenzívneho jazyka je dôležitá, pretože negatívne myslenie môže viesť napríklad k samovraždám, sebapoškodzovaniu, psychickým poruchám, nerovnováhe, chorobám zo stresu, drogovej či alkoholovej závislosti a iným [2].

Úroveň detekcie nenávisťných prejavov a ofenzívneho, či toxického jazyka je pre anglický jazyk na veľmi dobrej úrovni. Avšak pre slovenský jazyk je výskum v tejto oblasti slabší. V rámci nášho výskumu sa zaoberáme technikami a jazykovými modelmi, pomocou ktorých vieme úspešne detegovať nenávisťné prejavy a ofenzívny jazyk v texte písanom v slovenskom jazyku.

V predchádzajúcom výskume [3] sme skúmali nami vytvorený súbor údajov pre úlohu analýzy sentimentu v slovenskom jazyku. Softvérové riešenie využívalo konvolučnú neurónovú sieť s prvkami rekurzívnej neurónovej siete. Okrem toho sme navrhli a implementovali dodatočný prístup na zisťovanie najnenávisťnejších komentárov pomocou lexikónu expresívnych slov. V článku [4] sme načrtli, ako by mal vyzeráť náš ďalší korpus textov, na ktorom práve pracujeme. Mal by byť triedne vyváženejší, presnejší, lepšie anotovaný s viacerými podkategóriami.

V tomto príspevku sa venujeme dátovému korpusu, ktorý je nevyhnutný pre tréning a testovanie jazykových modelov. Pre vytvorenie korpusu textov pre úlohu analýzy sentimentu a detekcie nenávisťných prejavov a ofenzívneho jazyka je potrebné dodržať niekoľko kritérií pri jeho tvorbe a následnej úprave a anotácii.

II. ZÁKLADNÉ KRITÉRIA PRE VYTVORENIE KORPUSU TEXTOV

V úlohách zameraných na strojové učenie (ML) je potrebný korpus textov, ktorý sa používa na tréovanie modelu pre vykonávanie rôznych úloh strojového alebo hlbokého učenia. Dôvodom, prečo je korpus textov potrebný je, že ML do veľkej miery od dát závisí. Čiže bez dát je umelá inteligencia (AI) neschopná učiť sa. Je to najdôležitejší aspekt, ktorý umožňuje tréovanie algoritmov ML.

Bez ohľadu na zručností, či vedomostí ľudí v tíme a veľkosti dátovej množiny, ak dátová množina nie je dostatočne kvalitná, celý projekt AI nedosiahne uspokojivé výsledky. Počas práce s AI je z veľkej časti potrebné spoliehať sa na korpus textov. Od tréovania, ladenia, výberu modelu až po testovanie používame korpus textov rozdelený do troch množín a to konkrétne na: tréovaciu, validačnú a testovaciu. Tréovacia množina sa používa na tréovanie modelu, validačná množina sa používa na nastavovanie rôznych váh a doladenie modelu, a testovacia množina sa používa na hodnotenie natréovaného modelu.

Mnohokrát nestačí len zhromaždiť údaje, práve naopak. Vo väčšine úloh v oblasti AI trvá klasifikácia a anotovanie korpusu textov väčšinu času, najmä pri korpusoch textov dostatočne presných na to, aby odzrkadľovali realistickú víziu sveta.

A. Výber textových dát

Súbor tréovacích dát sa používa na tréovanie algoritmu, aby sa oboznámil, ako používať koncepty, ako sú neurónové siete, učenie sa a vytváranie výsledkov. Zahŕňa vstupné údaje aj očakávaný výstup. Tréovacie dáta tvoria väčšinu celkových dát, okolo 60%. Pri validácii sú modely prispôbené parametrom v procese, ktorý sa nazýva nastavovanie váh. Množina testovacích údajov sa používa na vyhodnotenie toho, ako dobre bol algoritmus natréovaný na dátovej množine.

V úlohách AI nie je možné použiť testovacie dátové súbory vo fáze validácie, pretože algoritmus už vopred pozná očakávaný výstup, ktorý nie je požadovaným cieľom. Testovacie dáta predstavujú 20% v súbore údajov. Je potrebné zaistiť, že množina vstupných údajov je zoskupená s overenými správnymi výstupmi, zvyčajne to overuje človek. Často sa neodporúča, pokúšať sa o ďalšie prispôbenie po fáze validácie. Pravdepodobne by to viedlo k pretrénovaniu modelu. Je to vypracovanie analýzy, ktorá príliš úzko alebo presne zodpovedá konkrétnemu korpusu textov, a preto nemusí vyhovovať ďalším úlohám alebo spoľahlivo predikovať v rámci iných úloh.

Dôležitou otázkou je, koľko dát je potrebných na dosiahnutie postačujúcich výsledkov v úlohe ML. Všetky úlohy ML sú niečím výnimočné, jedinečné a často odlišné. Čím viac je úloha komplikovanejšia, tým viac dát je potrebných pre tréovanie. Z tohto dôvodu nie je presne špecifikované, aká veľká množina dát je vhodná [5].

Ďalšou podstatnou otázkou je, aký typ dát je potrebný pre zvolenú úlohu. Dôležité je ozrejmiť si, čo sa snažíme a chceme pomocou AI dosiahnuť. Je potrebné zvážiť, aké údaje sú skutočne potrebné na riešenie úlohy alebo problému, na ktorom pracujeme. Je vhodné vytvoriť si niekoľko predpokladov o dátach, ktoré sú požadované pre tréovanie algoritmu.

Keď sa nám podarí zhromaždiť požadovaný súbor údajov je potrebné myslieť na fakt, že všetky súbory údajov sú nepresné. V tomto momente je potrebné vykonať určitú prípravu dát, čo je veľmi dôležitý krok v procese ML. Účelom prípravy je dosiahnuť, aby bola množina dát vhodnejšia pre úlohu ML. Jedná sa o niekoľko krokov a úprav, ktoré zaberajú väčšinu času stráveného nad úlohami ML. Avšak aj keď už máme z nášho pohľadu pripravené dáta, stále môžeme naraziť na problémy s ich kvalitou a nedostatkami, ktoré sú v nich ukryté a na prvý pohľad pre nás bezvýznamné, či nedôležité. Zjednodušene povedané, kvalita tréovacích dát ovplyvňuje celkovú úspešnosť algoritmov ML.

Často sa stáva, že nemáme dostatok dát na aplikovanie riešenia AI. Zoskupenie dát napríklad v CSV súboroch pre pripravovanú AI vyžaduje čas. Preto je potrebné vyhradiť si dostatočné množstvo času na editovanie alebo anotáciu potrebných dát. Prípadne je možné vytvoriť modernú stratégiu na zhromažďovanie dát. Ak je už presne určený cieľ zvolenej úlohy ML, môžeme zapojiť viacero ľudí/kolegov do tvorby a anotovania korpusu textov. Jednoduchým príkladom je, keď chceme rozpoznať obraz, no nemáme žiadne obrázky. Je potrebné aby tím zhromaždil/nafotografoval dostatočné množstvo obrázkov, aby bolo možné vytvoriť korpus textov. Áno, bolo by možné tieto fotografie zhromaždiť z Internetu, no v dnešnom svete a GDPR politikou to už nie je jednoduché, priam nemožné.

Troma dôležitými aspektami ML je kvalita, rozsah a množstvo dát. Predovšetkým strojové učenie nie je len o veľkej množine dát. V skutočnosti nepotrebujeme algoritmus naplniť všetkými znalosťami dátových bodov v akomkoľvek súvisiacom poli. Algoritmus je potrebné naplniť so starostlivo upravenými dátami dúfajúc, že sa dokáže naučiť a možno rozšíriť na okraj vedomostí,

ktoré už ľudia nadobudli. Väčšina odborníkov a spoločností verí, že stačí zhromaždiť všetky možné údaje, skombinovať ich a nechať AI, aby našla poznatky.

Pri vytváraní korpusu textov by sme sa mali zamerať na rozmanitosť dát. Najlepšie je zhromaždiť interné aj externé údaje (tzn. k téme, ale aj mimo nej). Cieľom je vytvoriť jedinečný korpus textov. Algoritmy ML vyžadujú veľké množstvo dátových bodov, no to neznamená, že model musí brať do úvahy širokú škálu vlastností [6].

Výber dát pre výskum v oblasti detekcie nenávisťného prejavu a ofenzívneho jazyka väčšinou smeruje k sociálnym sieťam a jednotlivým komentárom, tweetom a správam. Vo väčšine štúdií v tejto oblasti si autori zhromažďujú niekoľko stoviek, tisícok a niekedy až milióny tweetov, či komentárov v rámci jedného korpusu textových dát.

B. Predspracovanie textových dát

V tomto kroku už máme zhromaždené údaje, ktoré považujeme za nevyhnutné, rozmanité a vhodné pre zvolenú úlohu AI. Predspracovanie textu zahŕňa výber správnych dát z celkovej množiny dát a zostavenie tréningového korpusu textov. Proces zostavovania dát je v tomto optimálnom formáte známy ako transformácia vlastností.

- Formát: Dáta sa môžu vyskytovať a šíriť v rôznych dátových typoch. Napríklad sa môžu vyskytovať v rôznych jazykoch, v odlišnej mene, atď. Tieto dáta je potrebné zhromaždiť a zjednotiť, aby sa vytvoril jednotný korpus textov.
- Čistenie údajov: V tomto kroku je cieľom vyriešiť chýbajúce hodnoty a odstrániť z dát nežiadúce znaky.
- Extrakcia vlastností: Zameriavame sa na analýzu a optimalizáciu počtu vlastností. Zistíme, ktoré vlastnosti sú dôležité pre predikciu, vyberáme ich pre rýchlejšie výpočty a nízku spotrebu pamäte. V oblasti detekcie nenávisťného prejavu a ofenzívneho jazyka sa autori vo svojich štúdiách pri tvorbe korpusu textov zameriavajú na odstránenie prázdnych riadkov, odstránenie URL adries, odstránenie veľkých písmen, a podobne.

C. Anotácia textových dát

Pri antovaní korpusu textov pre oblasť detekcie nenávisťného prejavu a ofenzívneho jazyka sa kladie dôraz na klasifikáciu a vytvorenie niekoľkých klasifikačných tried (napr. obrázok 1). Vo väčšine štúdií si autori roztriedili korpus zhromaždených tweetov/komentárov do 3 tried (nenávisťný, ofenzívny, ani jedno). Zároveň sa vo veľkej miere zamerali aj na prvky sociálnych sietí, ktorými používatelia vyjadrujú svoje pocity a zdôrazňujú svoje nálady.

Ano len do nich, nedajte im dychat , poukazujte , informujte prosim
verjnost- narod o svinstvach , podlostiach , klamstvach dnesnej
neoliberalne kaste na cele s psychopatom Matovicom , ktora vedie
krajinu do zahuby...

Positive 1

Neutral 2

Negative 3

ID: 16162 CREATION_DATE: Monday October 12 2020 2:51pm GMT USER_NAME: Jozef Bindas

Fig. 1 Ukážka anotácie korpusu textov pre analýzu sentimentu alebo detekciu nenávisťných prejavov a ofenzívneho jazyka

Tieto parametre sú dôležité najmä pri analýze sentimentu. Analýza sentimentu identifikuje sentiment vyjadrený v texte a potom ho analyzuje [7]. Cieľom analýzy sentimentu je nájsť názory, identifikovať pocity, ktoré ľudia vyjadrujú. Následne klasifikovať polaritu, podobne ako je znázornené na obrázku 2. Cieľom analýzy sentimentu na úrovni viet je klasifikovať sentiment vyjadrený v každej vete. Taktiež je pre presnejšiu klasifikáciu potrebné vytvorenie slovníka pre skratky, ktoré sa najčastejšie používajú na sociálnych sieťach. Práve kvôli tomu, že používatelia v nich ukrývajú svoje citové rozpoloženie. Taktiež pri anotácii korpusu textových dát môžeme pri každom tweete alebo komentári označiť, aký typ nenávisťných prejavov alebo ofenzívneho jazyka sa v ňom nachádza.

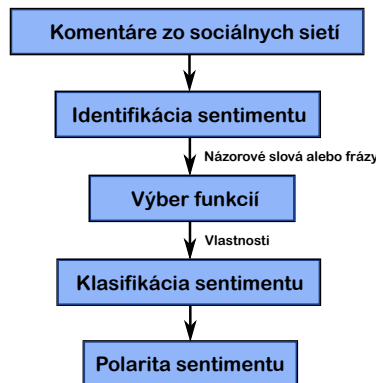


Fig. 2 Proces analýzy sentimentu

III. ZÁVER

Nenávistné prejavy sú veľkým problémom sociálnych sietí kvôli veľkému objemu dát, ktoré sú užívateľmi generované každý deň. V rámci nášho výskumu v oblasti detekcie nenávistných prejavov a ofenzívneho jazyka pre slovenský jazyk sa venujeme aj tvorbe korpusu textov. V tomto príspevku sme opísali postup a kritéria tvorby korpusu textov pre úlohy detekcie nenávistných prejavov a ofenzívneho jazyka. Opísali sme postup tvorby takéhoto korpusu textov a poznamenali sme dôležité vlastnosti a kroky, ktoré je potrebné dodržať. V ďalšom výskume by sme chceli prezentovať vytvorený korpus textov a experimentálne výsledky, ktoré pomocou neho dosiahneme.

POĎAKOVANIE

Výskum v tomto príspevku čiastočne podporilo Ministerstvo školstva, vedy, výskumu a športu SR v rámci výskumného projektu VEGA 2/0165/21, Agentúra na podporu výskumu a vývoja v rámci projektu bilaterálnej spolupráce APVV SK-TW-21-0002, výskumné projekty APVV-22-0414 a APVV-22-0261 a Fakulta elektrotechniky a informatiky TU Košice v rámci Grantu FEI-2023-95.

POUŽITÁ LITERATÚRA

- [1] M. ElSherief, V. Kulkarni, D. Nguyen, W. Y. Wang, and E. Belding, "Hate lingo: A target-based linguistic analysis of hate speech in social media," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 12, no. 1, 2018.
- [2] J. Whitlock, P. A. Wyman, and S. R. Moore, "Connectedness and suicide prevention in adolescents: Pathways and implications," *Suicide and Life-Threatening Behavior*, vol. 44, no. 3, pp. 246–272, 2014.
- [3] Z. Sokolová, J. Staš, and D. Hládek, "An introduction to detection of hate speech and offensive language in slovak," in *2022 12th International Conference on Advanced Computer Information Technologies (ACIT)*. IEEE, 2022, pp. 497–501.
- [4] Z. Sokolová, J. Staš, and J. Juhár, "Review of recent trends in the detection of hate speech and offensive language on social media," *Acta Electrotechnica et Informatica*, vol. 22, no. 4, pp. 18–24.
- [5] P. Domingos, "A few useful things to know about machine learning," *Communications of the ACM*, vol. 55, no. 10, pp. 78–87, 2012.
- [6] A. Gonfalonieri, "How to build a data set for your machine learning project," *Towards data science*, 2019.
- [7] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams engineering journal*, vol. 5, no. 4, pp. 1093–1113, 2014.

Explorácia v učení posilňovaním

¹Marián MACH

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

¹marian.mach@tuke.sk

Abstrakt — Pri riešení úloh učenia posilňovaním agent na základe svojej skúsenosti z minulosti (vo väčšej alebo menšej miere založenej na pôsobení typu pokus-omyl) sa snaží maximalizovať získavanú kumulatívnu odmenu. Článok sa zameriava na dilemu medzi exploračným skúmaním nových neznámych možností v podmienkach rizika, že tie možnosti sa ukážu ako neproduktívne, a čo najlepším využívaním už známej informácie pri riziku stagnácie v suboptimálnej oblasti. Dôraz je na exploračných stratégiách, ktoré umožňujú regulovať mieru exploraácie a tým riešiť dilemu hľadáním vhodnej rovnováhy medzi exploraáciou a využívaním známeho. V rámci experimentovania bolo realizovaných niekoľko experimentov s exploračnými stratégiami a ich riadiacimi parametrami v rámci zvoleného testovacieho problému.

Kľúčové slová — epsilon greedy politika, exploraácia, softmax, upper confidence bound, učenie posilňovaním

I. ÚVOD

Učenie posilňovaním reprezentuje rámec, v ktorom agent sekvenčne vyberá akcie a na základe toho získava odmenu. Snaží sa naučiť politiku výberu akcií, typicky reprezentovanú ako pravdepodobnostná distribúcia nad akciami, ktoré sú dostupné pre výber v danom okamihu. Cieľom jeho snaženia je naučiť sa politiku, ktorá maximalizuje kumulatívnu odmenu získanú výberom akcií.

V oblasti učenia posilňovaním jedným zo základných problémov je pomer medzi exploraáciou a využívaním získanej informácie. Príliš veľká miera exploraácie spôsobí zníženie akumulovanej odmeny (najmä skúmaním stavov, ktoré pre riešenie úlohu môžu byť menej dôležité a ich poznanie nie je príspevkom), zatiaľ čo príliš malá miera exploraácie spôsobí uviaznutie agenta v lokálnom optime.

Spôsoby exploraácie sa v zásade rozdeľujú do dvoch kategórií – na smerovú a nesmerovú exploraáciu. Nesmerová exploraácia je vedená náhodnosťou, keď agent úplne ignoruje aktuálny stav učenia a svoje akcie volí náhodným spôsobom. Smerová exploraácia reflektuje stav učenia pomocou nejakého pohľadu, ktorý môže byť založený napríklad na

- počte návštev jednotlivých stavov a počtoch výberov akcií v nich,
- veľkosti chýb odhadov kumulatívnych odmien,
- časovej distribúcii výberov akcií v navštevovaných stavoch.

Tento článok sa zameriava na metódy exploraácie v oblasti učenia posilňovaním. Štruktúra článku je nasledovná – v druhej kapitole sa venuje prehľadu metód exploraácie a v tretej kapitole je uvedená séria experimentov s danými metódami na úlohe hľadania sekvencie akcií pre dotlačenie autíčka na vrchol kopca.

II. EXPLORAČNÉ STRATÉGIE

A. Optimistická inicializácia

Táto metóda je založená na podpore exploraácie implicitným spôsobom, keďže pre výber akcie v nejakom stave s_t sa pravdepodobnosti výberu jednotlivých akcií určujú [1]

$$\pi(a_t | s_t) = \begin{cases} 1.0 & \text{ak } a_t = \arg \max_a Q(s_t, a) \\ 0.0 & \text{inak} \end{cases} \quad (1)$$

kde sa vždy vyberá akcia s najvyšším odhadom kumulatívnej odmeny.

Podstatou metódy je inicializovať odhady Q nadhodnotene („optimisticky“ ich odhadnúť vyššou hodnotou ako je skutočne dosahovaná kumulatívna odmena). Po iníciaálnom výbere

nejakej akcie a následná aktualizácia hodnoty $Q(s_t, a)$ pre danú akciu zníži pôvodný odhad, čím sa stane nižším ako odhady ostatných akcií v danom stave. Následkom toho pri nasledujúcej návšteve stavu s_t bude vybraná iná akcia. Takýmto spôsobom budú postupne pri opakovaných návštevách vyskúšané všetky akcie.

Takýmto spôsobom je exploračia podporovaná najmä na začiatku epizódy, s konvergenciou jednotlivých odhadov kumulatívnych odmien ku skutočným hodnotám jej účinok klesá, aj keď vplyv počiatočnej inicializácie bude viditeľný stále, pretože vzťah pre aktualizáciu hodnotovej funkcie na základe získanej odmeny

$$Q_{n+1}(s_t, a_t) = Q_n(s_t, a_t) + a(R_n - Q_n(s_t, a_t)) \quad (2)$$

sa dá prepísať na tvar

$$Q_{n+1}(s_t, a_t) = (1 - a)^n Q_1(s_t, a_t) + \sum_{i=1}^n a(1 - a)^{n-i} R_i \quad (3)$$

kde a je veľkosť učiaceho kroku. Čím je hodnota a menšia, tým silnejšie sa bude vplyv inicializácie prejavovať v neskorších výberoch.

B. ϵ -greedy politika

Jedna z najčastejšie používaných stratégií, explicitným spôsobom upravujúca mieru exploračie [1]. Používa parameter $0 \leq \epsilon \leq 1$ pre určenie, akým spôsobom má prebiehať výber vhodnej akcie. Pravdepodobnosti výberu jednotlivých akcií sa určujú

$$\pi(a_t | s_t) = \begin{cases} 1.0 - \epsilon + \frac{\epsilon}{|A(s_t)|} & \text{ak } a_t = \arg \max_a Q(s_t, a) \\ \frac{\epsilon}{|A(s_t)|} & \text{inak} \end{cases} \quad (4)$$

kde $A(s_t)$ je množina akcií, ktoré sú dostupné v stave s_t , a $| \cdot |$ označuje počet členov.

Daný vzťah sa dá interpretovať tak, že s pravdepodobnosťou $1 - \epsilon$ sa vyberá akcia s najvyšším odhadom kumulatívnej odmeny a s pravdepodobnosťou ϵ sa vyberá akcia náhodne (vrátane tej s najvyšším odhadom). Je zrejmé, že so zvyšovaním hodnoty ϵ stúpa miera exploračie pri výbere akcií.

Hodnota ϵ môže byť statická (nemenná) alebo dynamická. V prvom prípade je používaná rovnaká miera exploračie počas celej doby učenia. V druhom prípade je exploračia rôzna v rozličných fázach – typicky je najväčšia na začiatku epizódy a s postupom času miera exploračie klesá. Tento pokles môže byť lineárny alebo nelineárny, pričom zvyčajne sa štartuje z hodnoty $\epsilon = 1$ a pokles sa zastaví na nejakej hodnote ϵ_{min} , zabezpečujúcej nejakú minimálnu mieru exploračie aj v neskoršom štádiu.

Lineárny plán môže mať tvar

$$\epsilon_n = \max(d \epsilon_{n-1}, \epsilon_{min}) \quad (5)$$

kde d je útlmový koeficient z intervalu $(0, 1)$, zatiaľ čo nelineárny plán môže byť napríklad

$$\epsilon_n = \frac{1}{n} \quad (6)$$

kde n je počet krokov, už vykonaných v danej epizóde.

C. Softmax exploračia

Nevýhodou ϵ -greedy exploračie je to, že vyberá akcie úplne náhodne – pravdepodobnosť výberu najhoršej akcie je rovnaká ako pravdepodobnosť výberu druhej najlepšej akcie. Preto sa pre výpočet pravdepodobnosti akcií používa Boltzmannova distribučná funkcia [2]. V tomto prípade sa pravdepodobnosti akcií určia podľa

$$\pi(a_t | s_t) = \frac{e^{Q(s_t, a_t)/T}}{\sum_{i=1}^n e^{Q(s_t, a_i)/T}} \quad (7)$$

kde parameter T sa označuje ako teplota.

Pri tomto spôsobe aj naďalej je najpravdepodobnejší výber akcie s najvyšším odhadom kumulatívnej odmeny, avšak ostatné akcie sú zoradené podľa k nim prislúchajúcim odhadom. Čím je hodnota T vyššia, tým sú menšie rozdiely medzi akciami, a čím je táto hodnota naopak nižšia, tým sú preferovanejšie akcie s lepším odhadom kumulatívnej odmeny.

D. Sledovacia exploračia

Táto metóda [3] vychádza ako z aktuálnych hodnôt odhadov kumulatívnych odmien, tak aj z pravdepodobností výberu akcií v minulosti. Na rozdiel od predchádzajúcej metódy získava pravdepodobnosti výberu nie náhradou ale modifikáciou minulých pravdepodobností. Najprv na základe odhadov určí aktuálne najlepšiu akciu

$$a^* = \arg \max_a Q(s_t, a) \quad (8)$$

a následne modifikuje pravdepodobnosti výberu akcií

$$\pi_t(a_t | s_t) = \begin{cases} \pi_{t-1}(s_t, a_t) + \beta(1 - \pi_{t-1}(s_t, a_t)) & \text{ak } a_t = a^* \\ \pi_{t-1}(s_t, a_t) + \beta(0 - \pi_{t-1}(s_t, a_t)) & \text{inak} \end{cases} \quad (9)$$

kde β je koeficient z intervalu (0,1).

Pravdepodobnosť aktuálne najlepšej akcie je posunutá smerom k hodnote 1, zatiaľ čo pravdepodobnosti všetkých ostatných akcií sú posunuté smerom k hodnote 0. V prípade $\beta=1$ výber degraduje na selekciu akcie s aktuálne najvyšším odhadom kumulatívnej odmeny.

E. UCB1 explorácia

Princípom UCB (Upper Confidence Bound), na ktorom je táto stratégia založená, je zabezpečiť, aby v nejakom stave žiadna akcia nebola zanedbávaná a teda aby bola použitá dostatočne často (čo má význam hlavne v interakcii s nestacionárnym prostredím, kde vhodnosť použitia akcie sa môže s časom meniť) [4].

Počas interakcie s prostredím je sledovaný počet použitia každej akcie v každom stave, ak v danom stave je použitie danej akcie validné. Vždy, keď sa táto akcia použije, tak počet jej použití je inkrementovaný. Akcia získava bonus, ktorý je tým vyšší, čím je táto akcia menej používaná v danom stave oproti ostatným akciám. Tento bonus je určený podľa

$$b(s_t, a_t) = 100 C \sqrt{2 \frac{\log N(s_t)}{N(s_t, a_t)}} \quad (10)$$

kde $N(s_t)$ je počet navštívení stavu s_t , a $N(s_t, a_t)$ označuje počet, koľkokrát bola akcia a_t vybraná v stave s_t . Pravdepodobnosti výberu jednotlivých akcií sú

$$\pi_t(a_t | s_t) = \begin{cases} 1 & \text{ak } a_t = \arg \max_a (Q(s_t, a) + b(s_t, a)) \\ 0 & \text{inak} \end{cases} \quad (11)$$

kde výber prebieha deterministickým spôsobom.

III. VYBRANÉ EXPERIMENTY

A. Testovací problém

Pre experimentovanie s rôznymi metódami explorácie bol zvolený problém „Mountain Car“ [5] z kolekcie známych testovacích problémov určených pre testovanie metód učenia posilňovaním.

Jedná sa o deterministický MDP (Markovovský rozhodovací proces), keď autíčko, nachádzajúce sa na dne údolia, sa má dostať na vrchol kopca pomocou série akcelerácií. Priestor stavov je spojitý, definovaný dvomi reálnymi atribútmi (aktuálna pozícia a aktuálna rýchlosť), priestor akcií je diskretný s tromi rôznymi akciami (akcelerácia vľavo, žiadna akcelerácia, akcelerácia vpravo). Štartovací stav je volený náhodne.

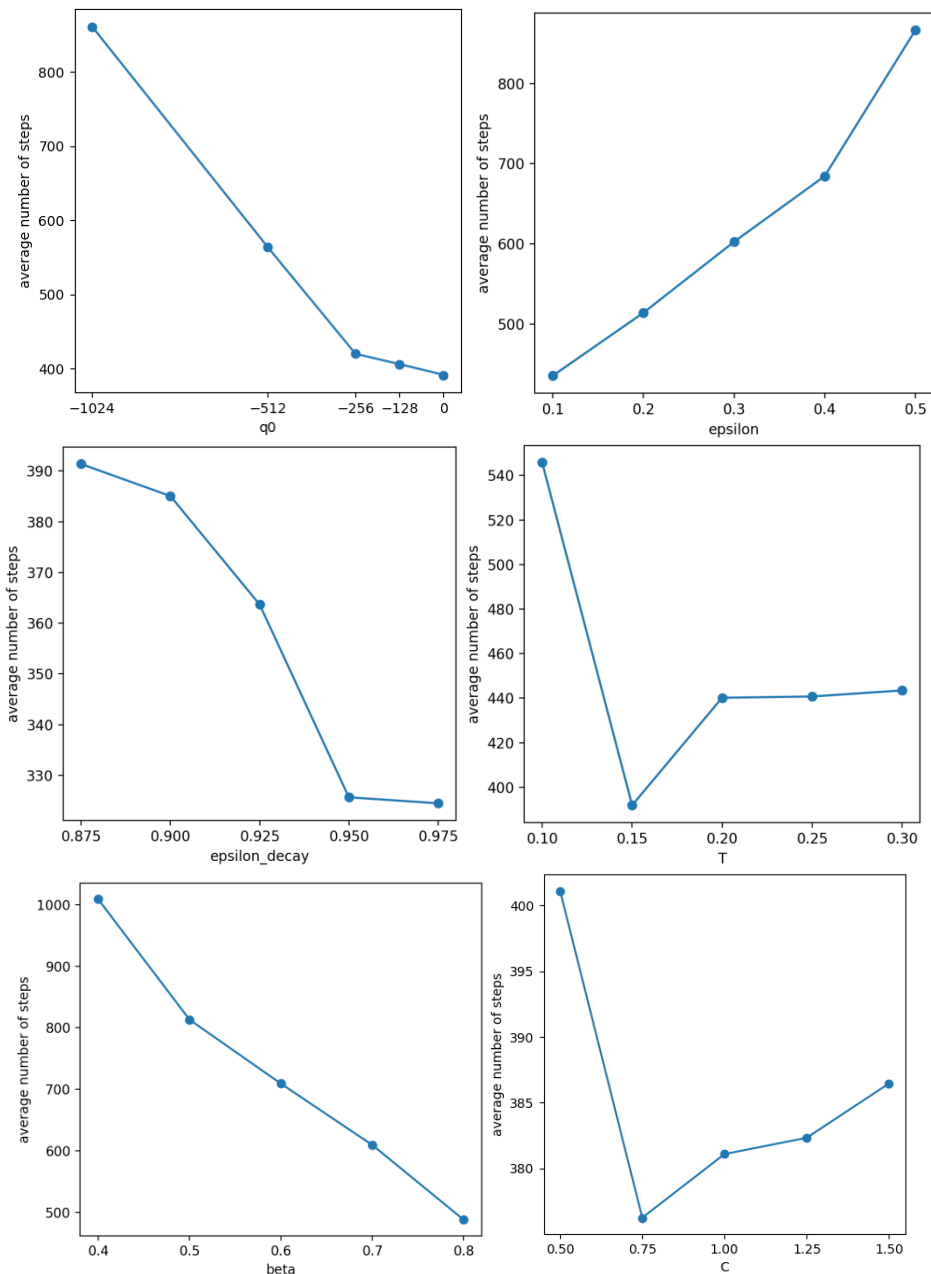
Keďže cieľom je dostať sa na vrchol kopca čo najskôr, tak odmena má výlučne charakter penalizácie – za každý časový krok je -1. Kumulovaná odmena je vždy hodnota záporná.

Úloha je epizodická – epizóda končí buď úspechom (autíčko dosiahne svoj cieľ) alebo neúspechom (vykoná sa maximálne povolený počet krokov).

B. Výsledky experimentov

Pre experimentovanie s exploračnými stratégiami bol použitý algoritmus Sarsa s tabuľkovou reprezentáciou hodnotovej funkcie, pričom stavový priestor bol diskretizovaný pomocou dlaždicového kódovania (osemnásobné delenie priestoru pomocou 10x10 dlaždíc). Rýchlosť učenia bola nastavená na hodnotu 0.1, diskontný faktor bol nastavený na 1.0.

Experiment bol vykonávaný pre rôzne hodnoty parametrov, ktoré jednotlivé metódy potrebujú k svojej činnosti. Bol sledovaný počet krokov, ktoré boli potrebné na dosiahnutie cieľa, pričom počet krokov v jednej epizóde bol obmedzený na 3000. Všetky experimenty boli vykonané 30 krát s následným spriemernením výsledkov. Dosiahnuté výsledky sú znázornené na obr. 1.



Obr. 1 Výsledky experimentov s exploračnými stratégiami

Pri optimistickej inicializácii (obr. 1, vľavo hore) boli skúmané jednak hodnoty, ktoré sú lepšie ako možné kumulatívne odmeny (každá odmena je menšia ako 0), ale aj iniciálne hodnoty, ktoré sú v pásme reálnych kumulatívnych odmien (hodnoty -128 a -256 sú často nad výkonom dosahovaným použitým algoritmom). Táto metóda pozitívne prekvapila, keď v kombinácii s deterministickým výberom akcie s aktuálne najvyšším odhadom kumulatívnej odmeny dokázala zabezpečiť použiteľné riešenie.

Stratégia ϵ -greedy ilustruje škodlivosť prílišnej miery exploračie (obr. 1, vpravo hore), kde so zvyšovaním statickej miery exploračie stúpa aj potrebný počet krokov pre dosiahnutie cieľa (a teda klesá kumulatívna odmena). Lepšie chovanie mala stratégia v dynamickej podobe s lineárnym plánom (obr. 1, vľavo v strede), kde koeficient ϵ z hodnoty 1.0 postupne klesal až na svoju konečnú hodnotu 0.01. Jasne však vidieť, že príliš rýchly pokles opäť zhoršuje výkon agenta – nie je možné v úvodných častiach epizód potláčať exploračiu priveľmi.

Softmax stratégia (obr. 1, vpravo v strede) nedokázala dostatočne znížiť mieru exploračie, preto sa nedokázala dostať na riešenia lepšie ako tie, ktoré boli dosiahnuté optimistickou inicializáciou. Navyše hľadanie vhodnej hodnoty parametra je obtiažne, pretože interval jeho teoreticky možných hodnôt je značne široký.

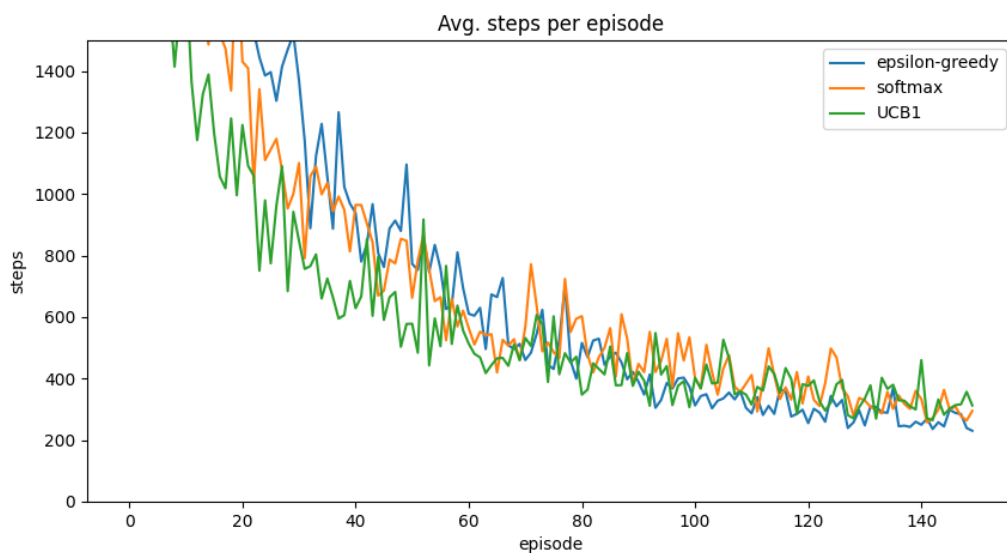
Vyslovene neúspešnou bola sledovacia stratégia (obr. 1, vľavo dole). Prezentované výsledky umožňujú predstavu extrapolácie, kde najlepšou hodnotou by bola hodnota parametra $\beta=1.0$, ktorá vlastne znamená deterministický výber akcie bez exploračie – čo vlastne vedie na prípad

výlučného použitia iba optimistickej inicializácie. Sledovacia stratégia takto vlastne viedla k zhoršovaniu výkonu. Je zrejmé, že daná stratégia sa nehodí pre typ úloh ako bola testovacia úloha (používanie iba penalizácií bez kladných odmien).

Posledná testovaná stratégia bola UCB1 (obr. 1, vpravo dole), ktorá dokázala vyvinúť mieru explorácie s pozitívnym prínosom pre nachádzané kumulatívne odmeny. Avšak podobne ako pri stratégii softmax, hľadanie vhodnej hodnoty riadiaceho parametra je obtiažny a dlhšie trvajúci proces.

Je potrebné zdôrazniť, že sledovacia stratégia a stratégia UCB1 využívali pamäť na sledovanie vybraných informácií (minulá pravdepodobnosť výberu, počet návštev v minulosti) pre každú dvojicu stav-akcia. Keďže stavový priestor testovaného prípadu bol pomocou dlaždicového kódovania diskretizovaný do značného počtu diskretných stavov, mohlo to mať vplyv na uvedené stratégie – návštevy nejakej dvojice stav-akcia neboli mnohokrát opakované, následkom čoho tieto stratégie získavali iba odhady sledovaných charakteristík.

Vzájomné porovnanie troch stratégií je na obr. 2. Je zrejmé, že UCB1 sa učila najrýchlejšie, zatiaľ čo učenie ϵ -greedy stratégie bolo najpomalšie, avšak v konečnej fáze dosahovala najlepšie hodnoty. Softmax bola v rýchlosti učenia strednou, avšak v druhej polovici dosahovala najväčšiu variáciu výsledkov.



Obr. 2 Porovnanie vybraných stratégií

IV. ZÁVER

Vykonalé experimenty potvrdili použiteľnosť prezentovaných metód explorácie, aj keď nie každá z testovaných stratégií sa ukázala rovnako prínosnou. Bol potvrdený intuitívny predpoklad, že pre hľadanie rovnováhy medzi exploráciou nových možností a stávaním na overených alternatívach je škodlivé vychýlenie ktorýmkoľvek smerom – či už potláčaním miery explorácie alebo naopak jej prílišným uvoľnením. Na druhej strane experimenty uvažovali iba jeden typ problému, charakterizovaného výskytom odmien iba vo forme záporných penalizácií. Na inom type problému môže byť výkonnosť jednotlivých stratégií iná než na použitom testovacom probléme.

LITERATÚRA

- [1] R.S. Sutton a A.G. Barto, *Reinforcement Learning- An Introduction* (2nd ed). MIT Press, 2018.
- [2] K. Saito, A. Notsu a K. Hond, “Discounted UCB1-tuned for Q-Learning”, Joint 7th Int. Conference on Soft Computing and Intelligent Systems and 15th Int. Symposium on Advanced Intelligent Systems, Kitakyushu, Japan, 2014
- [3] A.D. Tjisma, M.M. Drugal a M.A.Wiering, “Comparing Exploration Strategies for Q-learning in Random Stochastic Mazes”, IEEE Symposium Series on Computational Intelligence (SSCI), Athens, Greece, 2016.
- [4] K. Saito et al., “Performance Investigation of UCB Policy in Q-Learning”, IEEE 14th Int. Conference on Machine Learning and Applications, Miami, FL, 2015.
- [5] Mountain Car [online 26.5.2023] www.gymnasium.dev/environments/classic_control/mountain_car/

Detekcia Požiaru s Využitím Bunkových Sietí

¹Renát HALUŠKA, ¹Peter JANDA, ¹Stanislav MARCHEVSKÝ,
²Slavomír PILLÁR

¹ Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika
² TELESERVIS spol. s r.o., J. Hagaru 9, Bratislava - mestská časť Rača 831 51

¹renat.haluska@tuke.sk, ¹peter.janda@student.tuke.sk, ¹stanislav.marchevsky@tuke.sk,
²slavomir.pillar@gmail.com

Abstrakt — Článok obsahuje využitie UAV zariadení v kritických situáciách. Jednou z možných kritických situácií, ktoré môžu nastať je aj vznik požiarov. Požiar patrí k najsmrteľnejším a najnebezpečnejším prírodným katastrofám na svete. V rámci práce boli vytvorené neurónové siete na detekciu požiarov a následné vyhodnotenie snímok zhotovených pomocou UAV zariadenia Tello EDU.

Kľúčové slová — bezpilotné lietajúce prostriedky, mobilná sieť, neurónové siete, počítačové videnie

I. ÚVOD

Bunková sieť je rádiová sieť distribuovaná po zemi cez bunky, kde každá bunka obsahuje vysielač/prijímač s pevným umiestnením známym ako základňová stanica (GCS – angl. Ground Control Station), má základňovú stanicu analogickú s 802.11 AP, ktorá pomáha mobilným používateľom pripojiť sa k sieti. Medzi mobilnou a základňovou stanicou existuje vzdušné rozhranie fyzického protokolu a protokolu spojovej vrstvy. Tieto bunky spoločne poskytujú rádiové pokrytie väčších geografických oblastí. Používateľské zariadenie (UE – angl. User Equipment), ako sú mobilné telefóny, drony, sú preto schopné komunikovať, aj keď sa zariadenie počas prenosu pohybuje medzi rôznymi bunkami [1].

Bezpilotné lietadlá (UAV), tiež známe ako drony, získali významnú pozornosť v rôznych disciplínach vojenských a civilných služieb kvôli ich zvýšenej stabilite a odolnosti v niekoľkých operáciách. Aplikácie UAV sa výnimočne rozširujú vďaka ich pokročilému použitiu v internete vecí (IoT), 5G a B5G. UAV sa v poslednom desaťročí používali v rôznych aplikáciách vrátane detekcie a sledovania objektov, verejnej bezpečnosti, dopravného dozoru, vojenských operácií, prieskumu skrytých alebo nebezpečných oblastí, vnútornej alebo vonkajšej navigácie, snímania atmosféry, operácií po katastrofe, zdravotnej starostlivosti, zdieľanie údajov, riadenie infraštruktúry, núdzový a krízový manažment, nákladná doprava, monitorovanie lesných požiarov a logistika. UAV sú uzly celkového systému, ktorý je manažovaný cez vnútorný mechanizmus a Ad-Hoc komunikáciou medzi uzlami [2]. Taktiež sa môžu využiť v zdravotníctve na:

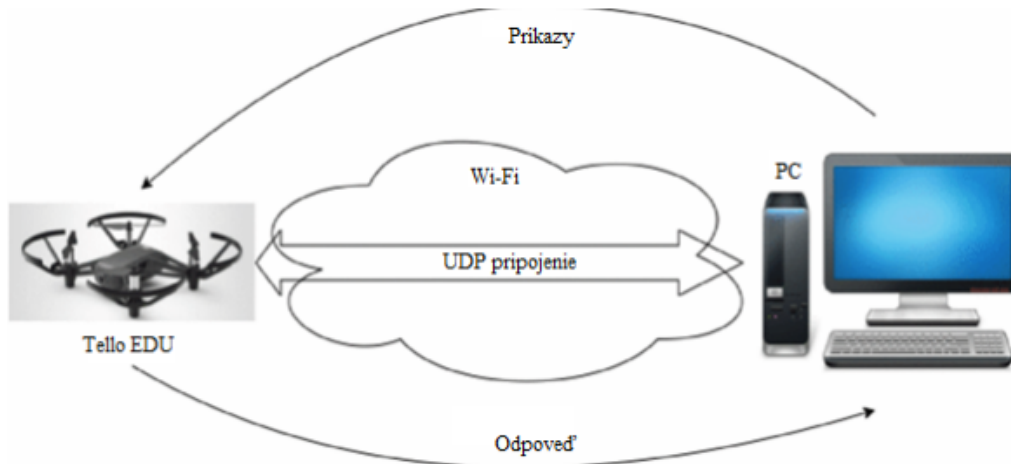
- prednemocničná pohotovostná starostlivosť,
- urýchlenie laboratórnej diagnostiky a testovania,
- monitorovanie.

Súčasne sa drony v zdravotníctve využívajú na doručovanie vakcín, hematologických produktov a automatizované externé defibrilátory [3].

II. KOMUNIKÁCIA MEDZI DRONOM A POZEMNOU STANICOU

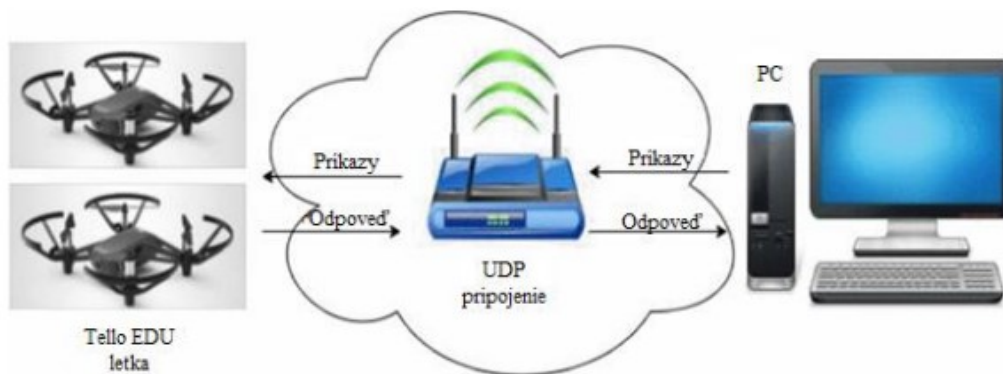
Bunkové siete ponúkajú rozsiahlu, vysokorychlostnú a bezpečnú bezdrôtovú konektivitu v rámci licencovaného spektra, čo môže výrazne zlepšiť kontrolu a bezpečnosť operácií malých

UAV a umožniť prevádzku za hranicou viditeľnosti (BVLOS – angl. Beyond the Visual Line Of Sight). Aby mohli UAV a aplikácia spolu komunikovať, všetky príkazy Tello SDK musia byť zabalené do funkcií Pythonu. Potom sa sprístupnia špeciálne príkazy pre riadenie UAV, ktoré možno použiť s akýmikoľvek algoritmickými konštrukciami a dátovými programovacieho jazyka. Všeobecná komunikačná schéma jedného UAV Tello EDU s aplikáciou je znázornená na Obr. 1 [4].



Obr. 1 Komunikácia Tello EDU s počítačom/aplikáciou [4]

Z dôvodu, že nie je možné priamo pripojiť viacero Tello EDU k počítaču, je potrebné použiť pomocné pripojenie, ktoré prepojí skupinu dronov a aplikáciu. Takýmto prepojovacím bodom je smerovač (angl. router). Program odošle príkazy smerovaču, ktorý následne pošle tieto príkazy všetkým Tello EDU v letke. Odpovede budú odoslané rovnakým spôsobom, iba v opačnom poradí - z UAV do počítača cez smerovač. Komunikácia je znázornená na Obr. 2.



Obr. 2 Komunikácia letky Tello EDU s počítačom/aplikáciou [4]

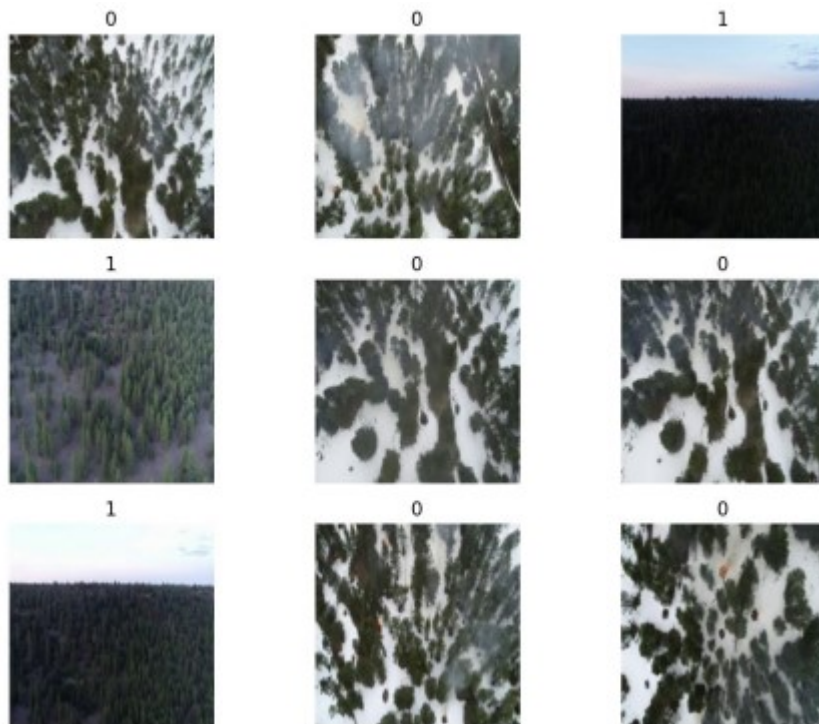
III. SIEŤ XCEPTION PRE ROZPOZNÁVANIE POŽIARU

XCception je architektúra hlbokoj konvolučnej neurónovej siete, ktorá zahŕňa hlbkovo oddeliteľné konvolúcie. Interpretuje moduly Inception v konvolučných neurónových sieťach ako medzikrok medzi pravidelnou konvolúciou a operáciou hlbkovo separovateľnej konvolúcie (hlbková konvolúcia nasledovaná bodovou konvolúciou) [5]. Dáta prechádzajú najskôr vstupným tokom (angl. entry flow), potom stredným tokom (angl. middle flow), ktorý sa opakuje osemkrát, a nakoniec výstupným tokom (angl. exit flow). Štandardná konvolučná vrstva neurónovej siete zahŕňa parametre $vstup * výstup * šírka * výška$, kde šírka a výška sú šírka a výška filtra. Pre vstupný kanál 10 a výstup 20 s filtrom $7 * 7$ to bude mať 2800 parametrov [6]. Toľko parametrov zvyšuje možnosť over-fitting-u. Overfitting je nežiaduce správanie strojového učenia, ku ktorému dochádza, keď model strojového učenia poskytuje presné predpovede pre tréningové údaje, ale nie pre nové údaje. Do týchto kategórií patrí hlbková konvolúcia a hlbkovo oddeliteľná konvolúcia. [7].

A. Trénovanie neurónovej siete

Na trénovanie neurónovej siete bolo potrebné stiahnuť set obrázkov (dataset), na ktorých sa nachádzal oheň/ požiar a obrázkov kde sa nenachádzal oheň/ požiar [8], [9], [10], [11]. V našej

neurónovej sieti bola priradená 0 pre obrázky kde sa nachádzal oheň a 1 pre obrázky kde sa nenachádzal oheň (Obr. 3).



Obr. 3 Vzorka použitých obrázkov pre učenie

Na tréningovanie neurónovej siete sa dataset rozdelí na obrázky určené na tréningovanie a obrázky určené na validáciu tréningu v pomere 80/20. Takže sme použili 31 500 obrázkov na tréning neurónovej siete a 7875 na validáciu. Validáčnne obrázky sa používajú v neurónových sieťach na vyhodnotenie výkonu modelu počas tréningového procesu a na ladenie hyperparametrov. Pri tréningovaní neurónovej siete je cieľom minimalizovať stratu tréningových údajov, ale konečným cieľom je vytvoriť model, ktorý sa dá vhodne zovšeobecniť.

Na Obr. 4 môžeme vidieť, že neurónová sieť prešla desiatimi epochami tréningu. Každá epocha trvala približne 430 až 440 sekúnd, za použitia ôsmich filtrov. Ak by sa použilo viac filtrov napr. 16, 32, 128, tak by to značne predĺžilo dobu tréningu. Použitím rôznych veľkostí filtrov v každom bloku sa neurónová sieť môže naučiť funkcie v rôznych mierkach. Počiatočné bloky používajú menšie filtre na učenie sa funkcií nízkej úrovne, ako sú okraje a rohy, zatiaľ čo neskoršie bloky používajú väčšie filtre na učenie sa funkcií vysokej úrovne, ako sú časti a vzory objektov.

```
985/985 [=====] - 434s 440ms/step - loss: 0.2337 - accuracy: 0.8834 - val_loss: 0.1788 - val_accuracy: 0.9575
Epoch 3/10
985/985 [=====] - 433s 440ms/step - loss: 0.2105 - accuracy: 0.8951 - val_loss: 0.1985 - val_accuracy: 0.9364
Epoch 4/10
985/985 [=====] - 434s 440ms/step - loss: 0.2005 - accuracy: 0.9033 - val_loss: 0.1553 - val_accuracy: 0.9549
Epoch 5/10
985/985 [=====] - 432s 438ms/step - loss: 0.1887 - accuracy: 0.9141 - val_loss: 0.1472 - val_accuracy: 0.9538
Epoch 6/10
985/985 [=====] - 433s 440ms/step - loss: 0.1772 - accuracy: 0.9220 - val_loss: 0.3910 - val_accuracy: 0.8358
Epoch 7/10
985/985 [=====] - 434s 441ms/step - loss: 0.1709 - accuracy: 0.9294 - val_loss: 0.4099 - val_accuracy: 0.7947
Epoch 8/10
985/985 [=====] - 433s 439ms/step - loss: 0.1671 - accuracy: 0.9341 - val_loss: 0.1238 - val_accuracy: 0.9453
Epoch 9/10
985/985 [=====] - 434s 440ms/step - loss: 0.1565 - accuracy: 0.9390 - val_loss: 0.0952 - val_accuracy: 0.9738
Epoch 10/10
985/985 [=====] - 440s 446ms/step - loss: 0.1514 - accuracy: 0.9404 - val_loss: 0.1460 - val_accuracy: 0.9449
```

Obr. 4 Epochy neurónovej siete

B. Vyhodnotenie neurónovej siete na obraze zachyteným z drona

Ako vstupné obrázky na vyhodnotenie neurónovej siete sme použili sadu obrázkov. Všetky obrázky boli vyhotovené pomocou Tello EDU. Proces zhotovenia množstva obrázkov bol nasledovný:

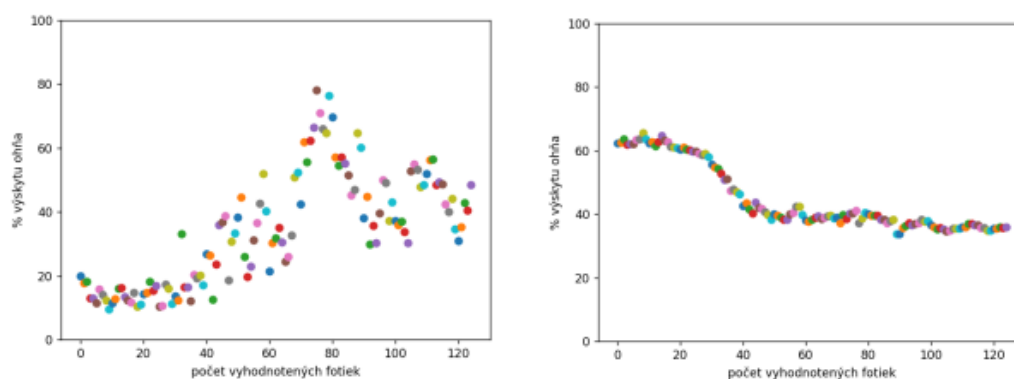
- nahralo sa video,
- následne bolo programom zostrihané na obrázky,
- zostrihané obrázky boli vyhodnotené neurónovou sieťou.

Na Obr. 5 môžeme vidieť obrázok ohňa počas obeda (naľavo) a obrázok ohňa večer (napravo). Ďalej tieto obrázky budeme testovať na rôznych natrénovaných neurónových sieťach. Natrénovali sme 3 rôzne neurónové siete s veľkosťou filtra 8 a 16. Tretiu neurónovú sieť sme natrénovali na väčšom množstve obrázkov.



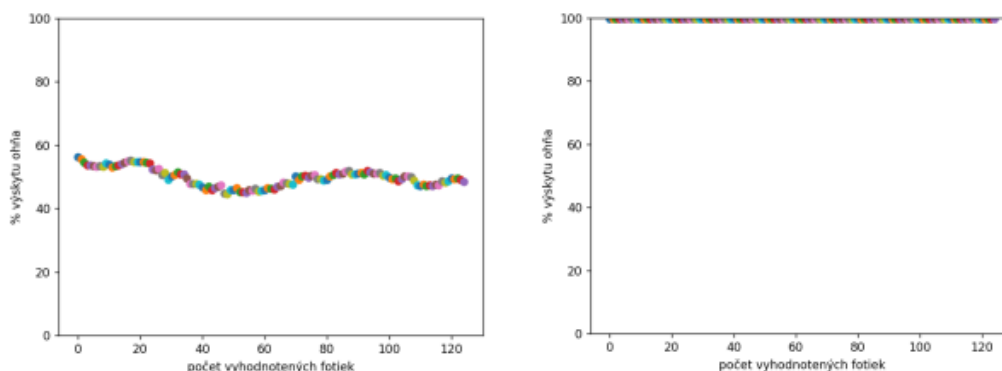
Obr. 5 Příklad detegovaného ohňa

Grafy (Obr. 6) zobrazujú závislosť výskytu ohňa od vyhodnotených fotiek. Naľavo je použitá veľkosť filtra 8 a napravo veľkosť filtra 16. Os x popisuje počet obrázkov, ktoré boli vyhodnocované a os y popisuje percento výskytu ohňa na jednotlivých obrázkoch. Dvojica grafov vyobrazuje závislosť obrázkov nafotených poobede. Ako môžeme vidieť, tak veľkosť filtra ovplyvnila výkon neurónovej siete. Neurónová sieť s veľkosťou filtra 16 vyhodnocovala obrázky s väčšou presnosťou ako a veľkosťou filtra 8.



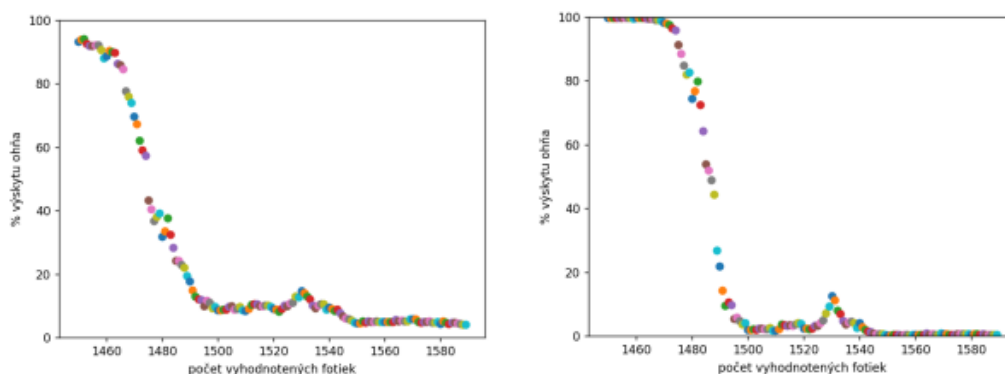
Obr. 6 Závislosť výskytu ohňa od vyhodnotených fotiek

Na grafoch (Obr. 7) môžeme vidieť naľavo použitú veľkosť filtra 8 a napravo veľkosť filtra 16. Taktiež sa vyhodnocovali rovnaké obrázky a to obrázky vyfotené večer. Ako môžeme vidieť, tak obe neurónové siete vyhodnocovali takmer so 100% úspešnosťou výskyt ohňa.



Obr. 7 Závislosť výskytu ohňa od vyhodnotených fotiek cez deň a večer

Na vytvorenie grafov (Obr. 8) sme použili neurónovú sieť, ktorá bola natrénovaná na väčšom množstve obrázkov. Graf naľavo predstavuje vyhodnotenie obrázkov, ktoré boli zhotovené poobede a graf napravo predstavuje vyhodnotenie obrázkov, ktoré boli zhotovené večer. Ako môžeme vidieť, obrázky, ktoré boli nafotené večer, neurónová sieť ich taktiež vyhodnotila takmer na 100%, že sa na nich nachádza oheň. Zlepšenie, v porovnaní s neurónovou sieťou s použitou veľkosťou filtra 8 a 16, nastalo pri vyhodnovení obrázkov, ktoré boli nafotené poobede. Nakoľko poobede nie je až tak viditeľný oheň tak má neurónová sieť ťažkosti vyhodnocovať s väčšou presnosťou.



Obr. 9 Závislosť výskytu ohňa od vyhodnotených fotiek s väčším datasetom

Na grafoch (Obr. 9) môžeme vidieť porovnanie vyhodnocovania neurónovej siete (s väčšou vzorkou pri tréningu) a s použitím veľkosti filtra 16. Ako môžeme vidieť neurónová sieť s veľkosťou filtra 16 presnejšie vyhodnocovala obrázky ak sme zmenili vstupné obrázky, ktoré obsahovali oheň na vstupné obrázky, ktoré neobsahovali oheň.

IV. ZÁVER

Prvá neurónová sieť, ktorá používala veľkosť filtra 8 mala problém rozpoznávať obrázky s ohňom počas svetla (poobede), ale vyhodnocovanie obrázkov za tmy (večer) robila s 98% úspešnosťou. Druhá neurónová sieť, ktorá používala veľkosť filtra 16, bola v porovnaní s neurónovou sieťou, ktorá využívala veľkosť filtra 8, lepšia o 37 % vo vyhodnovení obrázkov za svetla (poobede) a taktiež vyhodnocovanie obrázkov za tmy robila s takmer 100% úspešnosťou. Najlepšie vyhodnocovanie obrázkov ohňa počas svetla vyhodnocovala tretia neurónová sieť, ktorá bola natrénovaná na väčšej vzorke obrázkov.

POĎAKOVANIE

Výskum v tomto príspevku bol čiastočne podporený Vedeckou grantovou agentúrou Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky APVV-SK-TW-21-0002, APVV-22-0261 a APVV-22-0414 a Slovenskou akadémiou vied v rámci projektu VEGA 2/0165/21 a VEGA 1/0584/20.

LITERATÚRA

- [1]. Becker, R.A., Caceres, R., Hanson, K., Loh, J.M., Urbanek, S., Varshavsky, A. and Volinsky, C., 2011. A tale of one city: Using cellular network data for urban planning. *IEEE Pervasive Computing*, 10(4), pp.18-26.
- [2]. Mohsan, S. A. H., Khan, M. A., Noor, F., Ullah, I., & Alsharif, M. H. (2022). Towards the unmanned aerial vehicles (UAVs): A comprehensive review. *Drones*, 6(6), 147.
- [3]. Dhyani, K., Guhan, T., Gupta, P., Bhachawat, S., Ganapathy, G. P., & Srinivasan, K. (2022). Applications of IoT and Cloud Computing: A COVID-19 Disaster Perspective. In *New Frontiers in Cloud Computing and Internet of Things* (pp. 287-322). Cham: Springer International Publishing.
- [4]. Pohudina, O., Kovalevskiy, M., & Pyvovar, M. (2021, September). Group flight automation using Tello EDU unmanned aerial vehicle. In *2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT)* (Vol. 2, pp. 151-154). IEEE.
- [5]. Pan, D., Sun, L., Wang, R., Zhang, X., & Sinnott, R. O. (2020, December). Deepfake detection through deep learning. In *2020 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT)* (pp. 134-143). IEEE.
- [6]. Pandey, A. (2018). Depth-wise convolution and depth-wise separable convolution.
- [7]. Zhang, Z. Y., Sheng, X. R., Zhang, Y., Jiang, B., Han, S., Deng, H., & Zheng, B. (2022). Towards understanding the overfitting phenomenon of deep click-through rate prediction models. *arXiv preprint arXiv:2209.06053*.
- [8]. Shamsoshoara, Alireza. "The FLAME Dataset: Aerial Imagery Pile Burn Detection Using Drones (UAVs)." *IEEE DataPort*, 16 Apr. 2021, iee-dataport.org/open-access/flame-dataset-aerial-imagery-pile-burn-detection-using-drones-uavs.
- [9]. "Forest Fire." *Kaggle*, 15 Mar. 2021, www.kaggle.com/datasets/kutaykutlu/forest-fire.
- [10]. "FIRE Dataset." *Kaggle*, 25 Feb. 2020, www.kaggle.com/datasets/phylake1337/fire-dataset.
- [11]. "Fire and Smoke Dataset." *Kaggle*, 22 Apr. 2023, www.kaggle.com/datasets/dataclusterlabs/fire-and-smoke-dataset.

Anotačný nástroj korpusu textov pre úlohy opráv gramatických chýb v slovenskom jazyku

Maroš HARAHAUS¹, Zuzana SOKOLOVÁ², Matúš PLEVA³,
Daniel HLÁDEK⁴

¹Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovensko

¹maros.harachus@tuke.sk, ²zuzana.sokolova@tuke.sk, ³matus.pleva@tuke.sk,
⁴daniel.hladek@tuke.sk

Abstrakt – Témou tohto príspevku je problém nedostupnosti korpusov pre opravu gramatických chýb v textoch v slovenskom jazyku. V tomto príspevku sa zameriavame na opis korpusov dostupných v iných jazykoch. Na základe získaných informácií o korpusoch sme vytvorili webovú aplikáciu, pomocou ktorej plánujeme anotovať, čo najviac viet a tým pádom získať, čo najviac viet, ktoré sú gramaticky správne. Táto webová aplikácia je zameraná na určovanie, či je veta gramaticky správna alebo nie. Webová aplikácia má jednoduché prostredie, ktoré je ľahko pochopiteľné aj pre ľudí, ktorí sú menej zdatní v práci s počítačom.

Kľúčové slova – Anotačný nástroj, Gramatika, NLP, Textový korpus, Tvorba korpusu

I. ÚVOD

V dnešnej dobe sa technológie strojového učenia a umelá inteligencia stávajú stále viac zásadnými nástrojmi pre komunikáciu a vzdelávanie. Jednou z najdôležitejších aplikácií týchto technológií je automatická oprava gramatických chýb v texte. Táto oblasť kombinuje lingvistiku, dátové inžinierstvo a strojové učenie, aby sa zlepšila presnosť a efektívnosť písania a čítania v rôznych jazykoch.

Naším cieľom v tomto článku je ponúknuť čitateľom prehľad o korpusoch určených pre automatickú opravu gramatických chýb. Predstavíme rôzne korpusy, ktoré sa využívajú v tejto oblasti, a ukážeme, ako sa tieto dáta používajú pre tréning algoritmov strojového učenia.

Okrem toho sme vytvorili vlastnú webovú aplikáciu, kde sa zobrazujú vety a užívatelia môžu označiť, či je veta správna alebo nie, alebo, či si nie sú istí. Táto aplikácia nám umožňuje získavať cenné dáta pre tréning a testovanie našich modelov, a zároveň poskytuje užívateľom možnosť aktívne sa podieľať na vývoji technológií pre automatickú opravu gramatických chýb.

II. GRAMATIKA

Gramatika je súbor pravidiel, ktoré určujú, ako sa v danom jazyku tvoria a používajú slová, frázy a vety [1]. Tieto pravidlá sa môžu týkať rôznych aspektov jazyka, vrátane slovného druhu (napríklad podstatné meno, prídavné meno, sloveso), skloňovania (zmeny tvaru slova v závislosti od jeho funkcie vo vete), času (určuje, kedy sa daná akcia koná), rodu (mužský, ženský, stredný) a množstva (jednotné číslo, množné číslo).

Gramatika taktiež určuje syntax, čo je usporiadanie slov a fráz vo vete. Syntax stanovuje, ako sa slová a frázy kombinujú, aby vytvorili vety, ktoré majú zmysel [2].

Gramatika sa môže deliť na niekoľko rozličných oblastí podľa toho, na ktoré aspekty jazyka sa zameriava. Hlavné delenie gramatiky môže byť nasledovné [3]:

- **Fonológia:** Táto oblasť sa zaoberá zvukmi v jazyku a tým, ako sú tieto zvuky organizované a vzájomne interagujú.
- **Morfológia:** Morfológia skúma štruktúru slov a ich častí, ako sú korene, predpony, prípony a koncovky.

- **Syntax:** Syntax sa zaoberá pravidlami pre kombinovanie slov do fráz a viet.
- **Semantika:** Semantika sa zaoberá významom slov, fráz a viet.
- **Pragmatika:** Pragmatika skúma, ako kontext a situácia ovplyvňujú význam a použitie jazyka.

V niektorých prípadoch môže byť gramatika tiež rozdelená na deskriptívnu a prezskriptívnu. Deskriptívna gramatika opisuje, ako ľudia skutočne používajú jazyk, zatiaľ čo prezskriptívna gramatika stanovuje, ako by sa jazyk mal používať podľa určitých normatívnych pravidiel[4].

III. TEXTOVÝ KORPUS

Textový korpus je veľká a štrukturovaná zbierka textov. Texty sú zvyčajne zhromažďované a organizované na vedecké alebo analytické účely[5]. Najčastejšie sa využívajú v oblasti jazykovedy a spracovania prirodzeného jazyka (NLP). Korpusi môžu byť zložené z rôznych zdrojov, vrátane kníh, novín, webových stránok, akademických článkov, sociálnych médií a ďalších. Korpusy môžu byť zamerané na konkrétny jazyk, dialekt, sociálnu skupinu, obdobie času, žáner alebo tému. Niektoré korpusy sú "anotované", čo znamená, že boli obohatené o dodatočné informácie, ako je časť reči každého slova, syntaktické štruktúry, sémantické roly, atď. Anotované korpusy sú obzvlášť cenné v NLP, pretože môžu byť použité pre tréning modelov pre automatickú analýzu a porozumenie jazyku [6].

A. Korpus pre opravu gramatických chýb

Korpusy pre opravu gramatických chýb obsahujú príklady gramatických chýb a ich opráv, neoceniteľné. Tieto korpusy sú anotované tak, že označujú miesta, kde sa vyskytujú chyby. Môžu sa zaoberať rôznymi typmi chýb napr. v pravopise, syntaxi, morfológii, interpunkcii a konkrétnych problémoch pre ľudí, ktorí sa učia daný jazyk ako druhý jazyk.

- **Cambridge Learner Corpus (CLC):** Tento korpus obsahuje viac než 55 miliónov slov získaných z testov anglického jazyka napísaných študentmi, ktorí sa učia angličtinu ako druhý jazyk. CLC je anotovaný tak, že označuje a klasifikuje gramatické chyby, ktoré študenti robia. Táto anotácia poskytuje unikátne možnosti pre analýzu jazykových vzorcov a chýb, ktoré študenti robia počas procesu učenia sa angličtiny. CLC je tiež cenným zdrojom pre tvorcov testov a učebníc, ktorí chcú zlepšiť svoje materiály na základe skutočných chýb študentov [7].
- **Lang-8 Corpus of Learner English:** Lang-8 je online jazyková platforma, kde môžu používatelia písať texty v jazyku, ktorý sa učia, a natívni hovorcovia týchto jazykov môžu poskytnúť opravy. Korpus obsahuje tieto texty a opravy, a je tak cenným zdrojom pre štúdiá a opravu gramatických chýb. Tento korpus poskytuje jedinečnú príležitosť pre jazykových vedeckých pracovníkov sledovať, ako sa učia a vylepšujú študenti v reálnom svete, priamo na platforme [8].
- **French Learner Language Oral Corpora (FLLOC):** Tento korpus obsahuje transkriptované rozhovory francúzskych študentov a je anotovaný pre gramatické chyby. Tým, že sa zameriava na orálny jazyk, poskytuje jedinečný pohľad na špecifické výzvy, ktoré sa vyskytujú pri hovorenom jazyku, vrátane chýb v intonácii, rytmu a prízvuku, ktoré nie sú v písomnom texte tak zrejme [9].
- **NUS Corpus of Learner English (NUCLE):** Tento korpus obsahuje anglické eseje napísané študentmi na National University of Singapore. Je anotovaný pre gramatické chyby a ich opravy, a je zvlášť užitočný pre študentov, ktorí sa učia angličtinu ako druhý jazyk v Ázii. Korpus NUCLE poskytuje detailné informácie o tom, aké typy chýb študenti robia, a umožňuje učiteľom a výskumníkom vytvárať ciele cvičenia a metódy na ich nápravu [10].
- **GECCC - Grammar Error Corpus for Czech (GECCC):** je korpus písaného textu v češtine, ktorý obsahuje anotované gramatické, pravopisné a interpunkčné chyby. Bol vytvorený v roku 2022. Korpus pozostáva z 83 tisíc viet (949 tisíc slov). Korpus bol manuálne anotovaný 5 anotátormi. GECCC je významným prínosom do oblasti korpusovej lingvistiky, kde existuje potreba porozumieť a popísať význam a štruktúru gramatických chýb v češtine [11].

IV. WEBOVÁ APLIKÁCIA

Webová aplikácia pre anotovanie viet je určená na anotáciu, či je daná veta gramaticky správna alebo nie. Databáza viet obsahuje celkovo 90 tisíc viet. Na tvorbu webovej aplikácie sme použili štandardné jazyky (HTML, CSS, JavaScript a PHP). Pomocou HTML sme definovali štruktúru webovej stránky, CSS a Bootstrap zabezpečujú jednoduchý a prehľadný vzhľad. JavaScript bol použitý na zabezpečenie funkčnosti, napríklad na vytvorenie tlačidla, pomocou ktorých užívateľ

označuje, či je daná veta gramaticky správna alebo nie a podobne. Každá veta sa danému užívateľovi zobrazí iba raz. Jedna veta sa zobrazí maximálne trikrát. Podľa tohto princípu budeme ľahšie určovať, či užívatelia sa zhodli alebo nie.

Na obrázku číslo 1 môžeme vidieť, čo sa nám zobrazí po načítaní stránky. Pre rozpoznanie užívateľov každý užívateľ musí zadať mailovú adresu. Po zadaní mailovej adresy a stlačení tlačidla "Submit" bude užívateľ presmerovaný na hlavnú stránku.

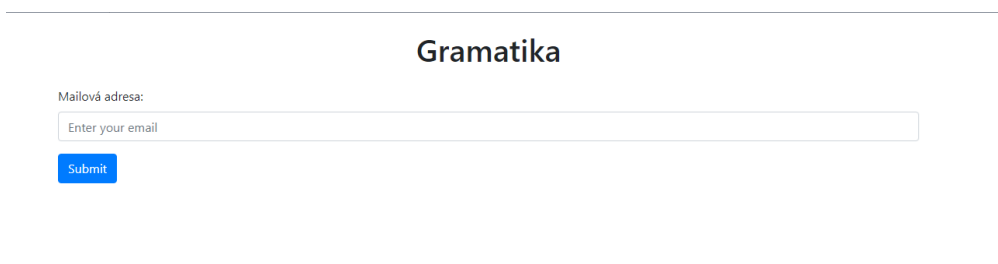


Fig. 1 Prihlasovacia obrazovka do webovej aplikácie

Ako je vidieť na obrázku číslo 2 po prejdení na hlavnú stránku sa užívateľovi zobrazí stránka, ktorej obsahom je náhodná veta načítaná z databázy viet. Pod vetou sa nachádza počítadlo viet. Pomocou tohto počítadla daný užívateľ vie koľko viet už spracoval. Počítadlo počíta spracované vety za celkový čas nie len za danú session. Nasledujú tlačidlá, pomocou ktorých užívateľ určuje, či je veta gramaticky správna. Na výber má z troch možností:

- Áno
- nie
- Neviem

Veta: Takže ja mám mozog, ale nemám funkciu, a preto nemôžem rozhodovať, a tí hore, čo mozog nemajú, ale majú funkciu, tí rozhodovať môžu?"

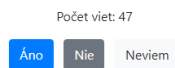


Fig. 2 Hlavná stránka webovej aplikácie

Užívateľ nemusí nič ukladať všetko sa automaticky uloží. Anotácie sa ukládajú do textového súboru, ako je to vidieť na obrázku číslo 3. Ako je možné vidieť na obrázku v súbore sa nachádza mailová adresa užívateľa, veta a voľba, ktorú užívateľ zvolil. Pomocou tohto súboru budeme môcť ľahko vybrať vety, ktoré sú gramaticky správne.

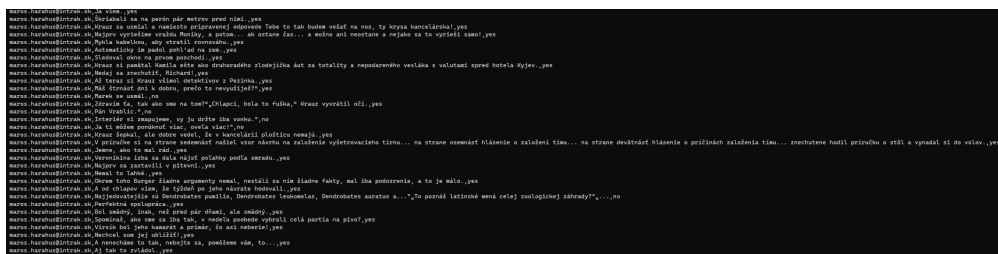


Fig. 3 Uložené anotácie v textovom súbore

V. ZÁVER

Automatická oprava gramatických chýb je významným nástrojom pre zlepšenie kvality textu a efektívnosti komunikácie. V tomto článku sme predstavili niekoľko korpusov určených pre automatickú opravu gramatických chýb, ako aj našu webovú aplikáciu pre anotáciu viet.

Korpusy pre opravu gramatických chýb poskytujú cenné dáta pre tréning a testovanie modelov strojového učenia a umožňujú nám lepšie porozumieť typom chýb, ktoré ľudia robia pri písaní a hovorení. Naša webová aplikácia umožňuje užívateľom aktívne sa podieľať na tomto procese tým, že anotujú vety podľa toho, či sú gramaticky správne alebo nie.

Dúfame, že naše práce pomôžu výskumníkom a vývojárom v oblasti umelého strojového učenia a spracovania prirodzeného jazyka v ich snahách vylepšiť automatickú opravu gramatických chýb a vytvoriť efektívnejšie nástroje pre komunikáciu a vzdelávanie.

POĎAKOVANIE

Táto práca bola podporená Agentúrou na podporu výskumu a vývoja SR v rámci grantu APVV-SK-TW-21-0002, APVV-22-0261 a APVV-22-0414, a čiastočne aj Fakultou elektrotechniky a informatiky TU Košice v rámci grantu FEI-2023-95 a v rámci grantu VEGA 2/0165/21 financovaného Ministerstvom školstva, vedy, výskumu a športu SR.

LITERATÚRA

- [1] N. Chomsky, "Three models for the description of language," *IRE Transactions on information theory*, vol. 2, no. 3, pp. 113–124, 1956.
- [2] A. Marantz, "No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon," *University of Pennsylvania working papers in linguistics*, vol. 4, no. 2, p. 14, 1997.
- [3] C. Anderson, *1.3 Mental Grammar*. Pressbooks, 2023, accessed on 2023-05-23. [Online]. Available: <https://pressbooks.pub/essentialsoflinguistics/chapter/1-3-mental-grammar/>
- [4] A. Reynolds, "Understanding prescriptive vs. descriptive grammar," 2023, the University of North Carolina at Chapel Hill. [Online]. Available: <https://amyrey.web.unc.edu/classes/ling-101-online/tutorials/understanding-prescriptive-vs-descriptive-grammar/>
- [5] D. Biber, *Discourse on the move: Using corpus analysis to describe discourse structure*. John Benjamins Publishing, 2007, vol. 28.
- [6] J.-D. Kim, T. Ohta, and J. Tsujii, "Corpus annotation for mining biomedical events from literature," *BMC bioinformatics*, vol. 9, pp. 1–25, 2008.
- [7] Z. Yuan and T. Briscoe, "Grammatical error correction using neural machine translation," in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2016, pp. 380–386.
- [8] T. Mizumoto, Y. Hayashibe, M. Komachi, M. Nagata, and Y. Matsumoto, "The effect of learner corpus size in grammatical error correction of esl writings," in *Proceedings of COLING 2012: Posters*, 2012, pp. 863–872.
- [9] F. Myles *et al.*, "French learner language oral corpora (floc)," *Oxford Text Archive Core Collection*, 2006.
- [10] D. Dahlmeier, H. T. Ng, and S. M. Wu, "Building a large annotated corpus of learner english: The nus corpus of learner english," in *Proceedings of the eighth workshop on innovative use of NLP for building educational applications*, 2013, pp. 22–31.
- [11] J. Náplava, M. Straka, J. Straková, and A. Rosen, "Czech grammar error correction with a large and diverse corpus," *Transactions of the Association for Computational Linguistics*, vol. 10, pp. 452–467, 2022.

Multi-Task Learning for Personalization

¹Jakub Dzurovčin, ²Ján Magyar

¹Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

²Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

¹jakub.dzurovcin@student.tuke.sk, ²jan.magyar@tuke.sk

Abstract – This paper explores the possibility of using multi-task reinforcement learning for personalization tasks. It describes practical approaches for combining multi-task reinforcement learning and personalization. We test our approach by implementing a multi-task agent IMPALA with modifications, using a form of normalization called PopArt, and various personalization-focused extensions. We use the IMPALA agent to solve model personalization tasks in the form of configurable mazes. Our results show that multi-task learning approaches face a number of key issues when dealing with personalization tasks, which the tested normalization methods were unable to completely solve.

Keywords – IMPALA, multi-task reinforcement learning, personalization, PopArt, reinforcement learning

I. INTRODUCTION

In recent years, personalization has become an increasingly important aspect of many applications and systems. From online shopping to social media, personalized recommendations have become a standard feature in software development. There are several approaches to personalization, and one of the most successful is the use of reinforcement learning (RL). Multi-task reinforcement learning (MTRL) is a natural evolution of reinforcement learning and deals with training agents that are capable of performing a larger number of similar or completely different tasks. This property represents an advantage in addressing the challenges of personalization, where the goal is to simultaneously learn multiple tasks that are related but not identical.

Section II provides a short overview of the state of the art in MTRL and personalization, section III describes our methodology. Section IV presents our experiments and results, and section V concludes the paper.

II. RELATED LITERATURE

Multi-task learning is a machine learning paradigm whose goal is to learn to perform multiple related tasks simultaneously. The goal of MTRL is to improve the overall performance of an agent across a larger number of similar tasks in the same environment. In the learning process, the agent builds a library of general knowledge that is applicable across tasks. The challenge is to find a balance between extracting common knowledge and specializing for specific tasks. This approach is also called **inductive knowledge transfer** [1]. Each task can have its own reward function and reward structure. Tasks can have shared or completely unique state spaces, action spaces, and transition functions. Tasks can differ in the number and type of states, actions, and transition rules between states.

Personalization is a process that changes the functionality, interface, access to information and content, or characteristics of a system to increase its personal relevance to an individual or category of individuals. While the software of these systems is identical for all users, the behavior of these systems can be adapted based on the experiences of individual users. Recently, reinforcement learning (RL) has gained considerable attention as an elegant paradigm for data-driven personalization. No explicit definition of the desired behavior of the model is required, as the RL agent learns it through interaction with the user. We have found three basic approaches to personalization in literature [2]:

- *one RL agent for every task* – This method of personalization involves defining a separate environment for each user and a separate RL agent for each user. The ultimate goal is to

find a set of optimal policies π_1, \dots, π_n for the set of user environments. The advantage and disadvantage of isolated environments is that differences in expected future rewards or current rewards for two different users are processed separately.

- *one RL agent for all tasks* – This approach involves defining a single RL agent and a single environment whose states contain user-specific information. The agent learns a single policy π for all users. In these applications, users may be described by a function that returns a vector representation of user features. This approach should be able to overcome the problem of negative knowledge transfer [1], however, the growth of the state space size may require an extremely large number of trajectories. Another disadvantage of this approach is the difference between the data used to represent the task and the users themselves. An alternative is to identify the user using an identifier, such as user's index [3]. There are several possible implementations for this approach, such as representing the user index using a trainable embedding. More complex techniques for creating personalized vectors include Task to Vector (Task2Vec) [4] and the use of vector fields [5].
- *hybrid approach* – The hybrid approach can be seen as a middle ground between learning a single policy π and learning all policies. Its idea is based on the assumption that users and their corresponding environments can be similar and is built upon clustering. For each group of features, an RL agent is defined that interacts with all environments.

The approach in which we train a single policy π for each user separately can generally achieve better results because learning is fully focused on a specific user. However, this approach requires a large amount of trajectories, computational power, and memory to store individual policies. This problem is compounded with each added user, while the possibility that some users may be similar is completely overlooked.

In addition to the combined approach with clustering, it is possible to combine the second approach (one policy for all users) with multi-task learning methodology. In theory, this approach can address many of the drawbacks associated with previous approaches. Learning multiple tasks requires a larger number of trajectories to learn a robust policy. MTRL algorithms take this aspect into account and use architectures to efficiently use resources in the learning process. The issue of positive and negative knowledge transfer between tasks, as well as the problem of dispersion due to different trajectory numbers, can be addressed using techniques such as PopArt [6], which ensures that all tasks are proportionally represented in the learning process.

IMPALA (Importance Weighted Actor-Learner Architecture) is a distributed agent architecture designed for multi-task learning [7] and addresses scalability issues such as:

- data load that an RL agent must process in a multi-task configuration, such as heterogeneous state spaces across tasks;
- long training times caused by a larger number of diverse tasks.

One of the key aspects of the IMPALA approach is its ability to efficiently utilize the resources of a single computer for training while scaling to multiple devices without sacrificing efficiency. It employs off-policy learning, where experience generation can be performed by an older policy than the current one to achieve relatively stable learning with separate modules for trajectory creation (actor) and learning (learner). Typically, IMPALA uses a single learner combined with multiple actors. In this system, each individual actor generates its own learning parameters called trajectories, which it sends to the learner through a queue. The learner collects trajectories from all actors and calculates a central policy π . Before each learning cycle, the actor updates its parameters according to the learner's policy. This approach is very similar to the RL algorithm A3C [8], and the IMPALA architecture was largely inspired by this algorithm.

PopArt is a method for improving MTLR by addressing the distraction dilemma [1]. PopArt as a method, can be added to existing algorithms, such as A3C or IMPALA. It operates on the principle of adapting the contributions of individual tasks during the agent's learning process. This ensures that each task will be properly represented throughout learning [6].

III. METHODOLOGY

In RL research, simple games are commonly used as environments. The original IMPALA publication utilized the DeepMind Lab library [7], featuring Atari games as individual tasks.

For our work, we established the following requirements:

- Simple, modifiable environments enabling quick learning and task definition;
- A standard interface for easy environment handling;
- Homogeneous state space for personalization, accommodating similar environments with varying details.

Considering these factors, we selected MazeLab [9] mazes as our environments, modifying them as needed. Mazes share some common features, such as identical goal, actor, free cell, and

obstacle colors for all tasks, and a consistent environment size of 20×20 cells. They differ in task layouts, obstacle shapes, obstacle densities, and reward schemes.

The used reward functions can be categorized into sparse, positive, and negative rewards:

- for positive rewards, agents are rewarded for moving towards the goal and penalized for moving away with twice the negative reward value;
- for negative rewards, agents are punished for both moving towards and away from the goal, with the latter having double the penalty value. This approach promotes reaching the goal in as few steps as possible, as agents always try to maximize rewards;
- for sparse rewards (zero), agents receive no rewards or penalties for moving towards or away from the goal.

In all cases, agents are significantly rewarded for reaching the goal. If they hit an obstacle, they receive a negative penalty and additional punishment for moving away from the goal. We define a score required to solve each environment, considering its complexity and reward scheme. An episode is considered solved if the agent reaches the goal in fewer than the maximum steps and with minimal collisions. Our experiments showed that the fixed positive or negative reward scheme based on approaching or moving away from the goal works best in our case.

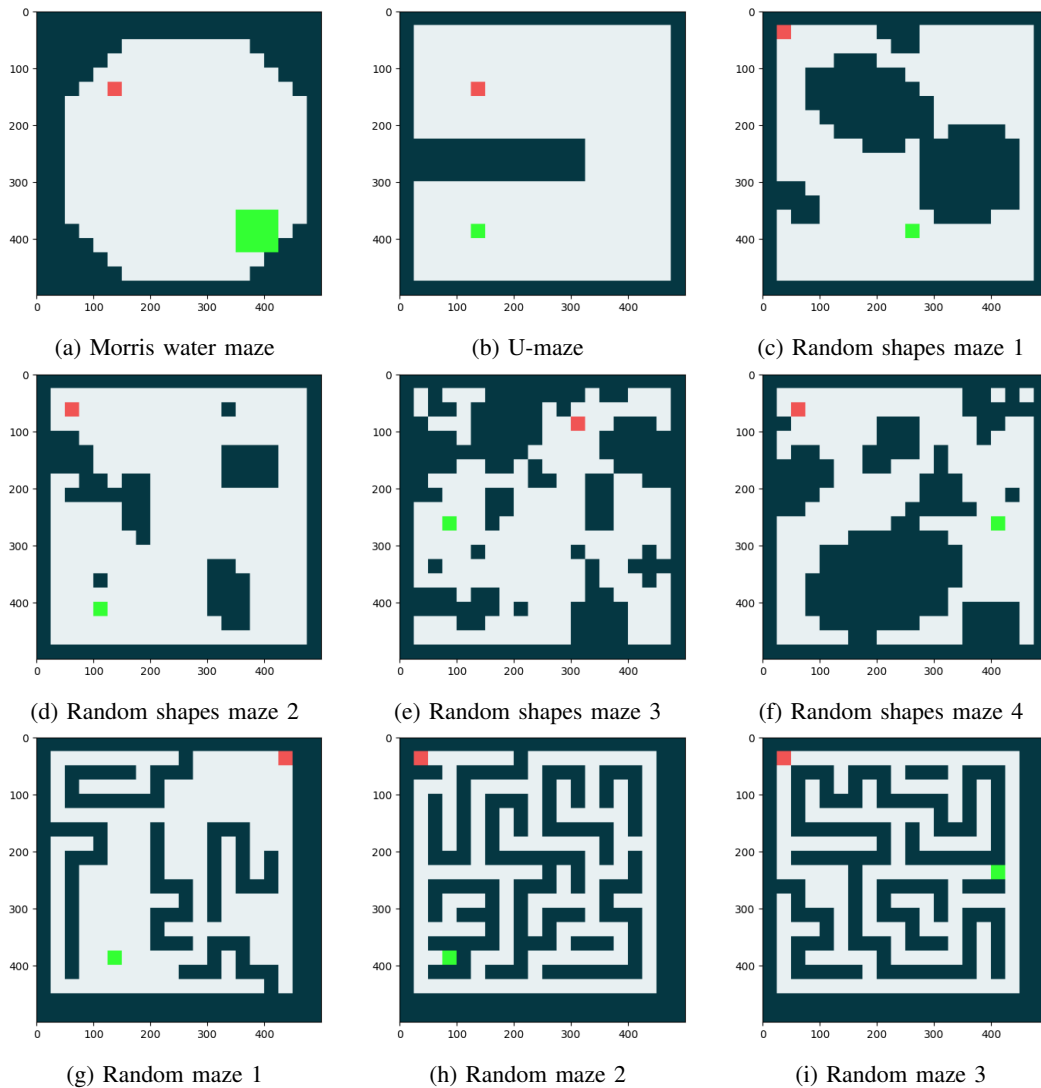


Fig. 1: Mazes

We use four basic types of mazes in combination with various reward schemes, see Fig. 1. The Morris Water Maze is the simplest, where the agent starts in the upper left corner and the goal is in the lower right corner. Obstacles are along the perimeter. This environment helps observe issues with negative knowledge transfer and distraction dilemma [1]. The U-maze is a harder version of the Morris Water Maze, where the agent must circumvent an obstacle to reach the goal below its starting point. Mazes with random shapes feature obstacles of random shapes and various starting positions and goals. This maze is harder than the U-maze due to more irregular obstacles. Random mazes are the most challenging mazes, with traditional obstacle shapes and

multiple versions varying in complexity and obstacle density. These mazes cater to different difficulty levels and allow us to study the agent’s behavior, knowledge transfer, and distraction dilemma issues.

IV. EXPERIMENTAL SETUP

Index	Environment	Movement penalty	Collision penalty	Goal reward	Solved when reward
0	U-maze	0	-	100	100
1	U-maze	0.01	-0.1	100	100
2	U-maze	-1	-5	100	20
3	U-maze	-0.1	-1	100	80
4	Morris water maze	0	-	50	50
5	Morris water maze	-0.01	-1	50	40
6	Morris water maze	-0.1	-5	50	40
7	Morris water maze	-1	-10	50	-10
8	Random maze 1	0	-	100	100
9	Random maze 1	0.01	-0.01	100	100
10	Random maze 1	-0.01	-1	100	80
11	Random maze 1	-0.01	-	100	80
12	Random maze 2	0.01	-0.1	100	100
13	Random maze 3	-	-	100	100
14	Random shapes maze 1	0	-	100	100
15	Random shapes maze 1	0.03	-0.01	100	100
16	Random shapes maze 1	-0.1	-3	80	80
17	Random shapes maze 2	-0.1	-1	100	80
18	Random shapes maze 1	-0.1	-1	100	80
19	Random shapes maze 3	-0.01	-0.1	100	80
20	Random shapes maze 4	-0.01	-1	100	80

Table 1: Task parameters

We conducted three types of experiments with different combinations of environments. Task specifics are described in Table 1:

- *4 tasks with sparse rewards* – We use all four environment types with a unified reward scheme. Two agents perform each task, totaling eight tasks during training.
- *4 random tasks, except sparse* – We use all four environment types with different randomly selected reward schemes. The environments differ in their state space and reward schemes.
- *8 random tasks, can also be sparse* – We use eight tasks, each environment type appearing twice with different reward schemes. This is the most difficult variant but includes the most diverse tasks.

We tested three agents: simple IMPALA, IMPALA in combination with PopArt, and IMPALA + PopArt with simple personalization in the form of scaling extracted convolution layer features by a task-specific value. Our expectations for each agent’s performance varied, with the last agent expected to achieve the best results.

We used the same neural network structure for all agent types with a batch size of 32, a buffer replacement rate of 0.3 and a learning rate of 0.00728. We trained five models for each experiment. Each model is trained for 5,000 training steps, later doubling the steps for the experiment with eight tasks.

Results confirmed our expectations, with the IMPALA agent learning to solve the easiest tasks while struggling with more complex ones. The IMPALA + PopArt agent showed unstable learning, and performance worsened for the task it learned to solve. The IMPALA + PopArt + simple personalization agent showed similar results.

After not being satisfied with these results, we conducted another experiment with different parameters and increased training steps. All agents still showed similar performance, struggling to complete tasks. We then attempted different personalization methods using embeddings methods, but observed no positive effect on the agents’ performance.

Overall, our experiments demonstrated that the tested agents struggled to solve complex tasks, and more advanced personalization techniques did not improve their performance. However, the agents still performed better than random in more complex tasks, showing some level of knowledge transfer between different tasks.

V. CONCLUSION

In this paper, we analyzed the problem of combining multi-task reinforcement learning and personalization, implementing an IMPALA agent with PopArt normalization and personalization enhancements. We faced challenges during implementation, such as framework-specific issues, and the need to train multiple models to find the best hyperparameters.

The experimental phase involved varying tasks, task numbers, and models with PopArt and personalization mechanisms. The pure IMPALA model performed the best, while additional modules destabilized the training process and worsened performance. Doubling training time and task numbers slightly improved results. The best outcomes were achieved using the pure IMPALA agent or IMPALA with the simplest personalization, but the distraction dilemma still persisted.

In the original publication, the IMPALA agent learned over 50 Atari games, benefiting from stabilization techniques and extensive training. Our agent might learn more tasks with longer training. Further research could involve reimplementing IMPALA or exploring alternative MTRL agents.

ACKNOWLEDGMENT

This paper was supported by the Slovak National Science Foundation project "Basic Research of Deep Learning for Image processing - DL4VISION" supported 2022-2025 under registration code 1/0394/22.

REFERENCES

- [1] V. Nelson and M. Qusay, "A survey of multi-task deep reinforcement learning," *Electronics*, vol. 9, p. 1363, 2020.
- [2] e. a. den Hengst, Floris, "Reinforcement learning for personalization: A systematic literature review," *Data Sci.*, vol. 3, pp. 107–147, 2020.
- [3] T. Sara, J. Natasha, N. Ehimwenma, S. Akane, and P. Rosalind, "Personalized multitask learning for predicting tomorrow's mood, stress, and health," *IEEE Transactions on Affective Computing*, pp. 1–1, 2017.
- [4] A. Achille, M. Lam, R. Tewari, A. Ravichandran, S. Maji, C. Fowlkes, S. Soatto, and P. Perona, "Task2vec: Task embedding for meta-learning," *arXiv:1902.03545v1*, 2019.
- [5] L. B, Y. S, Z. C, Y. J, and H. X, "Multi-task vector field learning," *Adv Neural Inf Process Syst*, vol. 2012, pp. 296–304, 2012.
- [6] M. Hessel, H. Soyer, L. Espeholt, W. Czarnecki, S. Schmitt, and H. van Hasselt, "Multi-task deep reinforcement learning with popart," *arXiv:1809.04474v1*, 2018.
- [7] E. Lasse, S. Hubert, M. Remi, S. Karen, M. Volodymir, W. Tom, D. Yotam, F. Vlad, H. Tim, D. Iain, L. Shane, and K. Koray, "Impala: Scalable distributed deep-rl with importance weighted actor-learner architectures," *arXiv:1802.01561v3*, 2018.
- [8] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. P. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, "Asynchronous methods for deep reinforcement learning," *arXiv:1602.01783v2*, 2016.
- [9] X. Zuo, "mazelab: A customizable framework to create maze and gridworld environments." <https://github.com/zuoxingdong/mazelab>, 2018.

Plánovanie trasy drona v inteligentnom priestore

¹Daniel HREHA, ²Ján MAGYAR

^{1,2}Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky,
Technická univerzita v Košiciach, Slovensko

¹daniel.hreha@student.tuke.sk, ²jan.magyar@tuke.sk

Abstrakt – Článok sa zaoberá problematikou plánovania trasy drona v interiéri pomocou dát získaných zo statických kamier. V našich experimentoch sme použili algoritmus A* a Rapidly-Exploring Random Tree pre naplánovanie trasy drona a porovnanie funkčnosti algoritmov medzi sebou. Navigačný systém bol otestovaný v rôznych situáciách, a dron v každej z nich dosiahol cieľ bez nárazu.

Kľúčové slová – dron, plánovacie algoritmy, plánovanie trasy

I. ÚVOD

Plánovanie trasy mobilného agenta v interiéri je jeden z najdôležitejších problémov, ktoré je potrebné vyriešiť pre úspešné dosiahnutie cieľového bodu. V tejto práci predstavíme jednoduchý používateľský systém, ktorý umožní používateľovi určiť štartovací a koncový bod letu drona v interiéri, ktorý je monitorovaný statickými kamerami. Následne systém vygeneruje optimalizovanú trasu drona pomocou algoritmu A* alebo Rapidly-Exploring Random Tree, a riadi drona, kým ten nepreletí vygenerovanou trasou.

V sekcii II sa zameriame na aktuálny stav riešenej problematiky, a na vybrané algoritmy, ktoré riešia problém plánovania trasy. Sekcia III pokračuje s popisom nášho vytvoreného systému, funkčnosť riešenia preukážeme v sekcii IV. Sekcia V uzaviera článok a načrtne možné smery budúceho výskumu.

II. PREHLAD PROBLEMATIKY PLÁNOVANIA TRASY MOBILNÝCH AGENTOV

Autonómne roboty sú vo vývoji už dlho, a rastúci dopyt po automatizovaných riešeniach pre reálne využitie takýchto robotov vo svete urýchlil výskum v tejto oblasti [1] [2] [3]. Problémy riešiteľné robotmi zahŕňajú úlohy v domácnosti ako vysávanie alebo kosenie trávy, riešenie problémov vo vojenskom priemysle na záchranu alebo na hliadkovanie, ale taktiež aj na úlohy spojené s donáškou balíkov k zákazníkovi alebo na roznášanie súčiastok po sklade.

Úlohou navigačného systému je naplánovať cestu k určenému cieľu a vykonať tento plán, respektíve upraviť ho podľa potreby tak, aby sa predišlo neočakávaným nárazom do novovzniknutých prekážok. Medzi informácie, potrebné hlavne pri rozširovaní navigačných systémov, patria informácie o polohe a pohybe mobilného robota v danom prostredí.

Globálne plánovanie cesty vyžaduje vopred naučený model, ktorý môže byť jednoduchým popisom skutočného sveta, a nemusí odrážať nedávne zmeny v prostredí. Tento globálny model musí poskytnúť plánovaciemu algoritmu sieť bodov, medzi ktorými sa môže robot pohybovať pomocou jednoduchých pohybov. Lokálny navigačný systém vykonáva kroky v globálnom pláne, udržiava odhad polohy robota vzhľadom na globálny priestor, a podľa potreby plánuje malé obchádzky, aby sa predišlo neočakávaným prekážkam. Zatiaľ čo globálna navigácia môže fungovať na vopred uloženej mape priestoru, lokálna navigácia pracuje s priestorom, ktorý odráža aktuálny stav prostredia vrátane zmien počas vykonávania plánu.

Plánovanie cesty robota je docielené rôznymi algoritmi, ktoré pomáhajú vytvárať optimálnu trasu pre robota, ktorý sa naviguje v prostredí. Navigačné algoritmy pracujú s mapou prostredia alebo pracovného priestoru. Správne alebo ideálne algoritmy dokážu splniť svoju úlohu dostatočne spoľahlivo a rýchlo, čo všeobecne znamená, že dokážu nájsť optimálnu trasu pre robota, po ktorej môže prechádzať, aj keď sa nachádza v prostredí, ktoré je preplnené prekážkami.

Ďalšou podstatnou úlohou pri navigácii robotov je zabezpečenie prísunu informácií, aby robot vedel, ako sa má dostať zo štartu do cieľa na základe trasy, ktorú mu vygeneroval algoritmus.

Na ovládanie robotov pomocou týchto algoritmov je potrebné dostať informácie o nájdenej trase pomocou algoritmov do ovládača robota.

Algoritmy plánovania trasy sú techniky, ktoré umožňujú autonómnemu mobilnému robotovi určiť postupnosť akcií na dosiahnutie cieľa, pričom sa robot zároveň vyhýba prekážkam. Môžeme ich deliť do niekoľkých skupín, ako:

- *vzorkovacie algoritmy* – využívajú pravdepodobnostné metódy na riešenie problému plánovania trasy. Hlavnou myšlienkou vzorkovacích algoritmov je vygenerovať veľké množstvo náhodných vzoriek a následne využiť tieto vzorky na vytvorenie reprezentácie prostredia. Následne algoritmus začne s prehľadávaním tohto prostredia za účelom nájdenia cesty, ktorá spĺňa zadané požiadavky [4] [5]. Vzorkovacie algoritmy sú výkonné a široko využívané na plánovanie ciest v zložitých podmienkach a prostrediach. Dokážu pracovať s veľkými priestormi, a taktiež dokážu rýchlo nájsť možné trasy v prostredí, v ktorom sa nachádza mnoho prekážok. Avšak tieto algoritmy nezaručujú vždy optimálnu cestu a zároveň môžu vyžadovať väčší čas pre zložitejšie problémy.
- *topologické algoritmy* – využívajú topologické znázornenie prostredia na plánovanie trasy. Algoritmy topologického plánovania pretvárajú mapu prostredia do grafu uzlov a hrán. Každý uzol v mape predstavuje oblasť prostredia, a každá hrana predstavuje cestu medzi regiónmi. Algoritmus následne prehľadáva tento graf za účelom nájdenia cesty [6] [1].
- *biologicky inšpirované algoritmy* – slúžia na nájdenie takmer optimálnej alebo optimálnej trasy zo štartu do cieľa, ktorá sa zároveň počas trasy vyhýba prekážkam. Biologicky inšpirované algoritmy sú alternatívou k tradičným algoritmom plánovania trasy a dokážu dobre zvládať zložité problémy, kde sa nachádza veľký počet prekážok alebo iných obmedzení.

III. IMPLEMENTÁCIA PLÁNOVACIEHO SYSTÉMU

Na navigáciu drona v interiéri sme sa rozhodli implementovať algoritmy A* a Rapidly-Exploring Random Tree (RRT). Dron, ktorý pomocou týchto algoritmov sme navigovali, bol dron Tello, ktorý sme ovládali pomocou knižnice tello-asyncio. Táto knižnica sa pripája na drona, získava údaje o stave drona, a odosiela príkazy na pohyb drona. Po odoslaní každého príkazu skript počká, kým nie je dokončený daný príkaz. K celkovej navigácii drona sú vykonané nasledovné kroky:

- 1) **výber kamery** – používateľ si vyberie, na ktorej kamere chce určiť koncové body letu drona;
- 2) **vytvorenie screenshotu z kamery** – pre jednoduchšiu definíciu koncových bodov vzhľadom na možné zmeny v priestore sa urobí screenshot, na ktorom používateľ bude označovať body;
- 3) **výber bodov** – používateľ si kliknutím vyberie koncové body letu;
- 4) **birdseye view** – z kamerového screenshotu sa vytvorí birdseye view ako vstup do plánovacieho algoritmu;
- 5) **úprava masky** – na obrázku sa automaticky detegujú prekážky, používateľ môže upraviť parametre generovania masky;
- 6) **výber veľkosti mriežky** – používateľ vyznačí priestor, v ktorom sa môže dron pohybovať;
- 7) **generovanie trasy drona** – používateľ má na výber z dvoch algoritmov: A* a RRT.

Systém sa pripája k statickým kamerám pomocou Real Time Streaming Protocol, a používatelia si môžu vybrať jednu kameru, cez ktorú môžu nastaviť scenár plánovania trasy - výber začiatočného a cieľového bodu letu a nastavenie obmedzení oblastí, v ktorej môže dron lietať ako jednoduchú mriežku. Aby sa táto voľba stala stabilnejšou, a obmedzil sa vplyv možných dočasných zmien v prostredí, v daný moment, keď používateľ vyberie kameru, sa z kamery zachytí statický obraz.

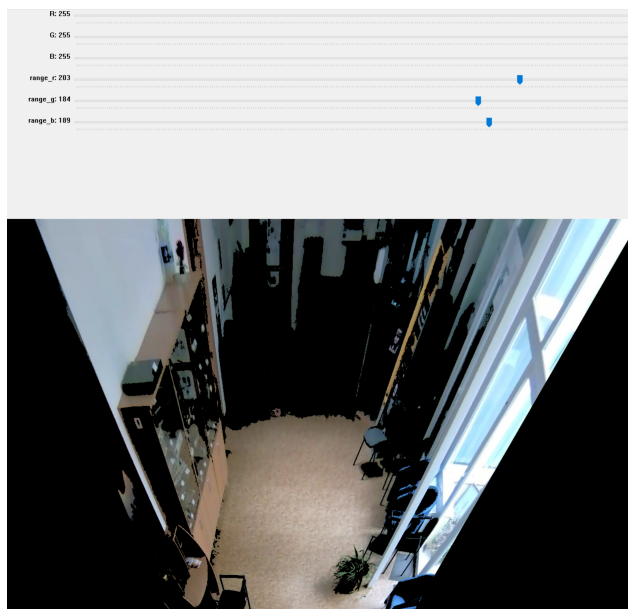
Výber začiatočného a cieľového bodu sa vykonáva priamo v používateľskom rozhraní, kde používateľ klikne na obrázok ľavým tlačidlom myši, aby nastavil začiatočný bod, a pravým tlačidlom myši, aby vybral cieľový bod.

Pre plánovanie trasy sú ideálne nastavenia kamery so záberom zhora na prostredie, čo umožňuje presné zisťovanie prekážok a ich lokalizáciu. Avšak, kamera so záberom zhora je zriedkavo dostupná v reálnych prostrediach, preto sme navrhli náš systém tak, aby dokázal spracovať obraz z akejkoľvek polohy kamery pomocou transformácie na pohľad z vtáčej perspektívy. Na túto transformáciu sme použili knižnicu *OpenCV*, definovaním dvoch zoznamov bodov (zdrojové a cieľové body), ktoré sa používajú na výpočet transformačnej matice a jej inverznej matice. Tiež prepočítavame polohu koncových bodov letu drona nastavených používateľom a vizualizujeme ich na pretransformovanom obraze (viď Obr. 1).

Z pohľadu z vtáčej perspektívy je tiež možné zisťovať prekážky, pre ktoré používame jednoduché analytické metódy počítačového videnia založené na prahovaní. Používateľ môže upravovať parametre rôznych farebných kanálov tak, aby proces zohľadňoval rôzne svetelné podmienky. Tieto hodnoty je možné nastaviť pomocou posúvačov priamo v používateľskom rozhraní, na základe čoho sa vizualizácia okamžite mení, a prekážky sa zobrazujú čiernou farbou, ako je možné vidieť na Obr. 2.



Obr. 1: Označené koncové body (zelený – štart, červený – cieľ) letu drona na kamerovom obraze a z vtáčej perspektívy



Obr. 2: Identifikované prekážky na obraze z vtáčej perspektívy

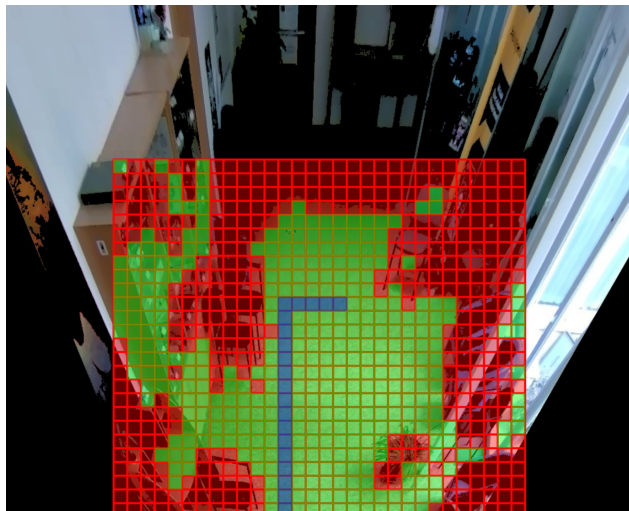
Ako posledný krok pred plánovaním trasy môže používateľ taktiež definovať obmedzenia oblasti, v ktorej môže dron lietať, pre dodatočnú bezpečnosť systému. Tento proces sa vykonáva podobne ako pri nastavovaní začiatkového a cieľového bodu na začiatku, jediný rozdiel je, že používateľ postupne definuje štyri body, ktoré slúžia ako rohy mriežky, v ktorej sa dron môže pohybovať. Táto mriežka je následne zobrazená na obraze z vtáčej perspektívy a každej bunke v nej je pridelený jedinečný identifikátor, ktorý neskôr definuje trasu letu drona. Na začiatku plánovania trasy sú začiatková a cieľová pozícia drona tiež zobrazené v tejto mriežke.

Pre samotné plánovanie trasy sme implementovali dva algoritmy, a to A* a Rapidly-exploring Random Trees (RRT). A* je topologický algoritmus, ktorý sa používa na nájdenie optimálnej trasy medzi dvoma bodmi vo váženom grafe so smerovaním k cieľovej pozícii [7]. Heuristická funkcia je použitá na približné určenie vzdialenosti každého bodu od cieľovej pozície, čo pomáha pri výbere nasledujúceho navštíveného uzla. Algoritmus začína inicializáciou nákladov štartovacieho uzla na 0 a približovaním vzdialenosti k cieľovej pozícii. Potom v každom kroku je vybraný uzol s najnižším nákladom od štartu po aktuálny uzol. Uzol je ďalej preskúmaný na základe svojich susedov, a algoritmus pokračuje, kým nepreskúma každý jednotlivý uzol alebo dosiahne cieľovú pozíciu.

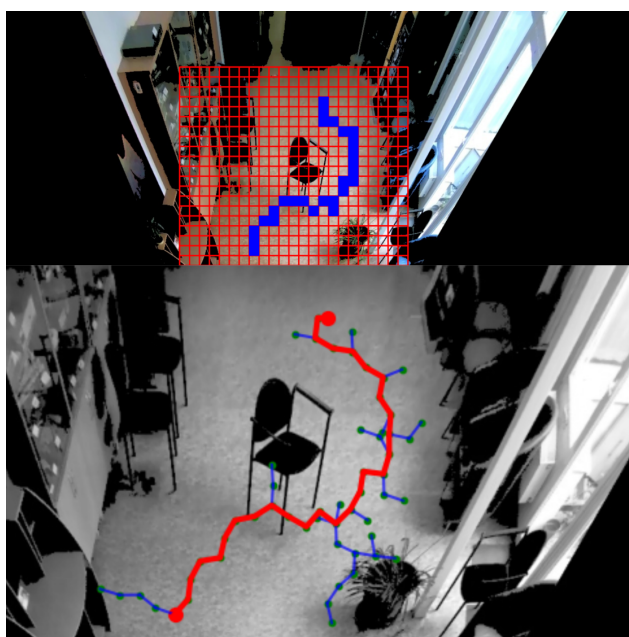
Algoritmus Rapidly-exploring Random Tree je založený na náhodnom preskúmaní priestoru a pridávaní uzlov do stromovej štruktúry, pričom nové uzly sú pripojené k predchádzajúcemu uzlu [8]. Algoritmus uprednostňuje predtým nepreskúmané časti priestoru, čo ho robí účinnejším pri hľadaní dostupných trás v prostrediach s viacerými prekážkami. Algoritmus začína so začiatkovou pozíciou a považuje ju za koreň stromu. Potom sa vyberie náhodná vzorka bodov a ak sú dostupné, ku stromu je pridaný uzol smerujúci k vzorke. Po vygenerovaní stromu algoritmus nájde uzol najbližší cieľovej pozícii, a nájde najkratšiu cestu od koreňa k tomuto uzlu.

IV. VALIDÁCIA PLÁNOVANIA TRASY

Systém sme testovali v priestoroch Ústavu výpočtovej techniky pri Technickej univerzite v Košiciach. Pre plánovanie trasy oba algoritmy dokázali nájsť bezpečné trasy vyhýbajúce sa prekážkam. Ako sa očakávalo, A* v priemere našiel kratšie trasy, avšak RRT bol viac vhodný na manipuláciu



Obr. 3: Trasa generovaná algoritmom A*: červená – prekážka, zelená – prázdna pozícia, modrá – generovaná trasa



Obr. 4: Trasa generovaná RRT, mriežka hore, generovaný strom dole

s vyhýbaním sa prekážkam. Na obr. 3 je zobrazená ukážková cesta vygenerovaná algoritmom A*, kde červené bunky mriežky predstavujú zistené prekážky (viď stoličku v strede a kvetináč vpravo dole), zelené bunky predstavujú voľné pozície, a cesta je zobrazená modrou farbou. Obrázok však ukazuje aj niektoré obmedzenia detekcie prekážok, kde prahovanie nebolo schopné identifikovať prekážky, ktoré mali rovnakú farbu ako pozadie.

Na obr. 4 je znázornené, ako algoritmus RRT pracoval s podobným nastavením. V hornej polovici obrázka môžete vidieť trasu zobrazenú na mriežke: vygenerovaná cesta je zobrazená modrou farbou (detekcia prekážok bola rovnaká ako pre A*). Diagonálne pohyby sa vykonávali v dvoch krokoch. Vygenerovaný strom je zobrazený v dolnej časti obrázka a konečná cesta je zobrazená červenou farbou.

Pre vyhodnotenie dvoch implementovaných algoritmov sme vykonali 50 experimentálnych behov, pričom v každom z nich sme vygenerovali náhodný štartovací a koncový bod v priestore s rozmermi $3m \times 3m$. Túto oblasť sme pre jednoduchšie spracovanie rozdelili na mriežku, v ktorej každý štvorček mal rozmery $20cm \times 20cm$ (celkovo 15×15 štvorčekov). Toto rozdelenie bolo opodstatnené skutočnosťou, že dron pri každom presune urobil krok s dĺžkou $20cm$ v jednom zo štyroch kardinálnych smerov, pričom výška letu bola konštantná.

Pri našich experimentoch priemerná vzdialenosť medzi štartovacím a koncovým bodom bola $132,2cm$, čo zodpovedá $6,62$ krokom s konštantnou dĺžkou. Algoritmus A* vygeneroval trasy s priemernou dĺžkou $8,54$ krokov (zodpovedá $170,8cm$), pričom RRT v priemere vygeneroval o

niečo dlhšie trasy s 9,1 krokmi (zodpovedá 182cm). Dron sa však vyhýbal prekážkam úspešne pri každom experimente.

V. ZÁVER

V tomto článku sme priblížili implementáciu jednoduchého systému, ktorý umožňuje naplánovať trasu interiéneho drona. Systém podporuje dva vybrané plánovacie algoritmy, konkrétne A* a RRT. Systém bol validovaný v priestore so statickými kamerami, a oba algoritmy boli schopné vygenerovať trasy vyhýbajúce sa prekážkam. V našich testoch sme použili drona Tello.

Výhodou tohto navigačného systému je to, že dokáže nájsť cestu efektívne a jednoducho aj napriek tomu, že umiestnenie kamery nie je ideálne. Výhodou tohto systému je aj to, že je možné vybrať si z dvoch navigačných algoritmov, podľa požiadaviek používateľa. Avšak tento systém by sa dal vylepšiť lepším umiestnením kamery, ktorá sníma priestor, a taktiež aj vylepšením drona, ktorý by mal možnosť hýbať sa aj diagonálne. Taktiež by sa dal systém vylepšiť použitím iných navigačných algoritmov.

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.

REFERENCES

- [1] 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.
- [2] T. Sasaki and H. Hashimoto, "Human observation based mobile robot navigation in intelligent space," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2006, pp. 1044–1049.
- [3] R. S. Pol and M. Murugan, "A review on indoor human aware autonomous mobile robot navigation through a dynamic environment survey of different path planning algorithm and methods," in *2015 International conference on industrial instrumentation and control (ICIC)*. IEEE, 2015, pp. 1339–1344.
- [4] N. Sariff and N. Buniyamin, "An overview of autonomous mobile robot path planning algorithms," in *2006 4th student conference on research and development*. IEEE, 2006, pp. 183–188.
- [5] A. Tahirovic and G. Magnani, "A roughness-based rrt for mobile robot navigation planning," *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 5944–5949, 2011.
- [6] F. Gul, W. Rahiman, and S. S. Nazli Alhady, "A comprehensive study for robot navigation techniques," *Cogent Engineering*, vol. 6, no. 1, p. 1632046, 2019.
- [7] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.
- [8] S. M. LaValle *et al.*, "Rapidly-exploring random trees: A new tool for path planning," 1998.

Lokalizácia drona v inteligentnom priestore

¹Marek TÓTH, ²Ján MAGYAR

^{1,2}Katedra kybernetiky a umelej inteligencie, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovensko

¹marek.toth.2@student.tuke.sk, ²jan.magyar@tuke.sk

Abstrakt – V tomto článku predstavíme systém, ktorý rieši problém lokalizácie drona v inteligentnom priestore, ktorý je snímaný v reálnom čase vhodne umiestnenými kamerami. Popíšeme model hlbokého učenia pre detekciu drona na kamerovom obraze, ako aj kalibráciu kamery pomocou lasera pre presnú lokalizáciu. Vytvorený systém bol otestovaný a validovaný v priestoroch Ústavu výpočtovej techniky, pričom v nakalibrovannej oblasti lokalizácia drona bola presná a konštantná.

Kľúčové slová – dron, kalibrácia kamery, lokalizácia

I. ÚVOD

Úspešná a bezpečná manipulácia s dronom (UAV) v interiéri vyžaduje vyriešenie niekoľkých úloh, ktoré boli definované pre riadenie mobilných agentov, ako je navigácia, lokalizácia, mapovanie a plánovanie trasy. V tomto článku popíšeme systém, ktorý využíva statické kamery na snímanie priestoru, v ktorom sa nachádza dron. UAV je následne detegovaný z kamerového obrazu pomocou neurónovej siete, a jeho poloha je určená na základe kalibrácie kamier.

Článok je štruktúrovaný nasledovne: sekcia II popisuje laserové zariadenie, ktoré bolo navrhnuté a vytvorené pre podporu kalibrácie kamier pre presnú lokalizáciu drona. Lokalizácia drona a vývoj neurónovej siete sú popísané v sekcii III. Sekcia IV obsahuje vyhodnotenie systému a diskusiu dosiahnutých výsledkov, a načrtne možné smerovanie ďalšieho vývoja.

II. KALIBRÁCIA KAMERY

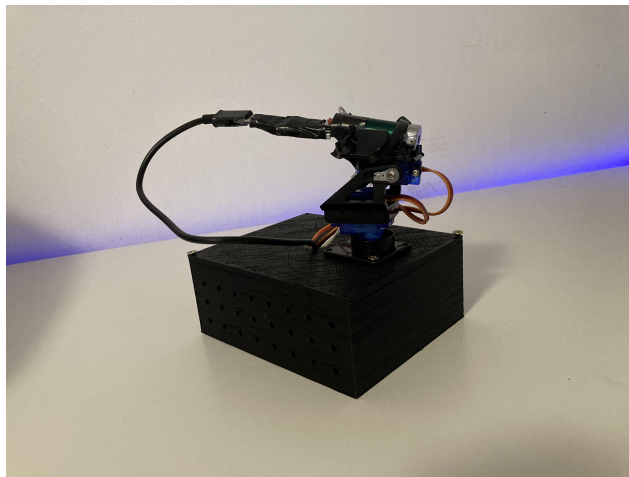
Kalibrácia kamery je rozhodujúcim procesom v počítačovom videní, ktorý zahŕňa určenie vnútorných a vonkajších parametrov kamery s cieľom umožniť zachytenie presných obrázkov a videí napríklad pre detekciu a sledovanie objektov, ako aj 3D rekonštrukciu.

Kalibrácia kamery sa používa na korekciu skreslenia, ku ktorému dochádza pri snímaní obrazu kamerou. Toto skreslenie môže byť spôsobené rôznymi faktormi, ako sú nedokonalosti objektívu, nelinearita snímača, a skreslenie perspektívy. Tieto skreslenia môžu spôsobiť, že zachytené obrázky budú vyzeráť odlišne od skutočných scén, čo vedie k nepresným meraniam a odhadom.

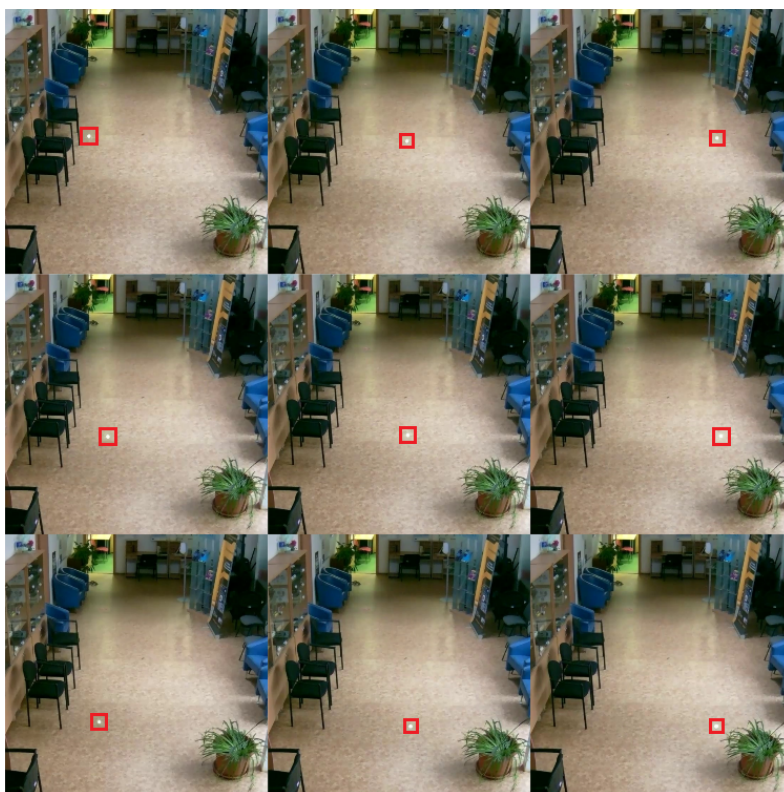
Náš systém používa techniku kalibrácie kamery pomocou lasera, táto metóda bola popísaná v [1]. Pomocou lasera sa dá jednoducho pohybovať kalibračným bodom po celej rovinatej ploche (priestore), bez potreby konštantného premiestňovania kalibračnej pomôcky, ako napríklad šachovnice. To umožňuje širšie využitie kalibračnej metódy aj pre viaceré kamery, kým je zabezpečený pohľad z kamery na kalibračný bod.

Pre naše účely sme navrhli a vytvorili vlastné laserové ukazovátka, ktoré sme umiestnili na statické miesto. Zariadenie sa ovláda pomocou vzdialeného pripojenia, hlavnou riadiacou jednotkou je mikropočítač Wemos D1 Mini založený na mikrokontroléri ESP8266. Pohyb lasera v dvoch osiach je riešený cez dva servomotory typu SG90. Okrem nasmerovania lasera vzdialený prístup umožňuje laser zapínať a vypínať, ako aj nastaviť stav napájania a vypínať servomotory, aby sme predišli ich prehrievaniu. Pre ovládania lasera bola vytvorená jednoduchá web stránka, príkazy sa posielajú ako HTTP požiadavky. Model lasera je možné vidieť na obrázku 1.

V našom riešení sme použili kalibráciu na základe 9 bodov, ktoré boli postupne nasnímané v mriežke 3×3 , ako je možné vidieť na obrázku 2. Po automatickom ukladaní týchto snímkov sa spustí kalibračný skript, ktorého cieľom je vygenerovať maticu kamery, ktorá udáva hodnoty potrebné na neskoršiu transformáciu obrazu z danej kamery.



Obr. 1: Laserové zariadenie určené na kalibráciu kamier



Obr. 2: Ukážkové pozície laserových bodiek pre kalibráciu kamery (bodky zvýraznené v červenom rámečku)

III. LOKALIZÁCIA DRONA

Lokalizácia mobilného agenta je úloha, ktorej cieľom je určiť pozíciu agenta v snímanom prostredí vzhľadom na vybraný referenčný bod. V exteriéri sa najčastejšie používa lokalizácia pomocou GPS (Global Positioning System), pričom pre interiér takisto existuje niekoľko štandardných riešení, ako napríklad ultraširokopásmová lokalizácia [2], a lokalizácia pomocou signálu wi-fi [3], bluetooth [4] alebo radio frequency identification (RFID) [5]. Tieto metódy však buď nemajú dostatočnú presnosť, alebo vyžadujú umiestnenie špeciálnych snímačov v priestore, alebo na samotnom agentovi, čo v prípade drona môže narušovať rovnováhu zariadenia počas letu.

Omnoho širšie aplikovateľný prístup je lokalizácia drona z kamerovej snímky, keďže interiérové priestory v priemysle už často disponujú s kamerovým vybavením. Ako sme už spomínali, pred lokalizáciou je však potrebné tieto kamery nakalibrovať pre dosiahnutie zjednoteného súradnicového systému. Po úspešnej kalibrácii kamery a získaní kalibračnej matice je možná lokalizácia drona, ktorá sa skladá z dvoch základných krokov:

- 1) detekcia drona na obraze;
- 2) určenie polohy drona k referenčnému bodu.



Obr. 3: Detekcia drona na kamerovom obraze

A. Detekcia drona

Detekciu drona z kamerového záberu sme sa rozhodli riešiť pomocou neurónovej siete, konkrétne prístupom transfer learningu, pričom sme použili model YoloV5. YOLO (You Only Look Once) je populárny algoritmus detekcie objektov v reálnom čase, ktorý vyvinul Joseph Redmon a jeho tím [6]. YOLOv5 je piata iterácia tohto algoritmu vyvinutá spoločnosťou Ultralytics v roku 2020, ktorý je založený na podpornej sieti CSP (Cross Stage Partial) s hlavou SPP (Spatial Pyramid Pooling) a PAN (Path Aggregation Network).

Pre tréning drona sme použili agregovaný dataset, v ktorom sa nachádzalo 16 278 snímok dronov z voľne dostupných zdrojov s označením pozície drona na obrázku. Z datasetu sme použili 80% dát na tréning, 10% na validáciu a 10% na testovanie. Model sme trénovali po 100 epoch s učiacim parametrom 0,01. Konečná presnosť dosahovala 85%. Ukážka detekcie drona na kamerovom obraze je možné vidieť na obrázku 3.

B. Určenie polohy drona

Po detekcii drona na obraze je možné vypočítať jeho pozíciu v skutočnom svete použitím kalibračnej a transformačnej matice. Lokalizácia drona v našom systéme funguje priamo na RTSP (Real Time Streaming Protocol) vysielaní kamery. Lokalizačný skript pracuje s už spomínanou transformačnou maticou z kalibrácie kamery, a ďalej s x -ovou a y -ovou súradnicou rohov ohraničeného detegovaného objektu. Následne sa tieto body konvertovali do homogénnych koordinátov pomocou prídania čísla 1 do každého riadku matice bodov.

Pre určenie skutočných súradníc drona sme použili inverznú maticu kamerovej matice, a body sveta sme dostali pomocou delenia súradníc súradnicou z :

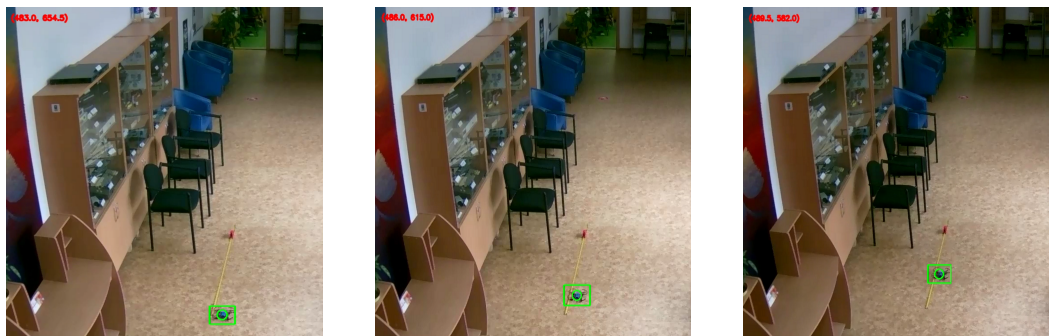
$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p1 & p2 & p3 & p4 \\ p5 & p6 & p7 & p8 \\ p9 & p10 & p11 & p12 \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

Výsledné súradnice drona (x a y) sú neskôr vyobrazené v ľavom hornom rohu grafického prostredia, ako je možné vidieť na obrázku 4. Pozíciu drona po osi z sme považovali za konštantnú. Pri našich experimentoch sme dron položili na ľubovoľnú pozíciu v rámci nakalibrovannej oblasti, a posunuli sme ho niekoľkokrát po x -ovej resp. y -ovej osi a spozorovali sme zmenu v súradnicovej hodnote po danej osi. Súradnicové hodnoty pri každom pokuse reprezentovali konštantnú vzdialenosť, na ukázkovom prípade posun drona o 25 cm zodpovedal 7 jednotkám po x -ovej osi (obrázok 4) a 3 jednotkám po y -ovej osi (obrázok 5).

IV. ZÁVER

Implementovaný systém sme testovali v priestoroch Ústavu výpočtovej techniky pri Technickej univerzite v Košiciach. Priestor disponuje statickými IP kamerami namontovanými na strope chodby. K testovaniu sme použili dron Ryze Tello. Pohyb drona bol implementovaný posunmi v kardinálnych smeroch o 25 centimetrov.

Testovanie potvrdilo detekciu drona z kamerového obrazu s vysokou úspešnosťou, a lokalizácia bola presná v rozmedzí kalibračných bodov. Táto skutočnosť bola potvrdená konštantnými zmenami po jednotlivých osiach pri presune drona o konštantnú vzdialenosť. Lokalizácia drona bola



(a) Dron na štartovacej pozícii.
 $y = 483$

(b) Dron po presune o 25 centimetrov. $y = 486$

(c) Dron po presune o 50 centimetrov. $y = 489.5$

Obr. 4: Prevedenie polohy drona na reálne súradnicové hodnoty (y -ová os)



(a) Dron na štartovacej pozícii.
 $x = 553$

(b) Dron po presune o 25 centimetrov. $x = 546$

(c) Dron po presune o 50 centimetrov. $x = 539$

Obr. 5: Prevedenie polohy drona na reálne súradnicové hodnoty (x -ová os)

neúspešná iba v prípade, ak dron bol viac vzdialený od kamery alebo sa nachádzal v priestore so zlými svetelnými podmienkami.

Systém sa dá rozšíriť určovaním z súradnicovej hodnoty pozície drona, čo by umožnilo efektívnejšie vyhýbanie sa prekážkam (napr. dron by dokázal preletieť nad nízkou prekážkou). K takémuto rozšíreniu by bolo potrebné vytvoriť systém, ktorý by dokázal dron detegovať na viacerých kamerách naraz, čo by ešte viac spresnilo určovanie jeho pozície.

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.

POUŽITÁ LITERATÚRA

- [1] T. Svoboda, D. Martinec, and T. Pajdla, "A convenient multicamera self-calibration for virtual environments," *Presence: Teleoperators & virtual environments*, vol. 14, no. 4, pp. 407–422, 2005.
- [2] L. Zwirello, T. Schipper, M. Harter, and T. Zwick, "Uwb localization system for indoor applications: Concept, realization and analysis," *Journal of Electrical and Computer Engineering*, vol. 2012, pp. 1–11, 2012.
- [3] J. Biswas and M. Veloso, "Wifi localization and navigation for autonomous indoor mobile robots," in *2010 IEEE international conference on robotics and automation*. IEEE, 2010, pp. 4379–4384.
- [4] R. Bruno and F. Delmastro, "Design and analysis of a bluetooth-based indoor localization system," in *Personal Wireless Communications: IFIP-TC6 8th International Conference, PWC 2003, Venice, Italy, September 23-25, 2003. Proceedings 8*. Springer, 2003, pp. 711–725.
- [5] T. Sanpechuda and L.-o. Kovavisaruch, "A review of rfid localization: Applications and techniques," in *2008 5th international conference on electrical engineering/electronics, computer, telecommunications and information technology*, vol. 2. IEEE, 2008, pp. 769–772.
- [6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

Modelovanie šírenia nedôveryhodných správ o pandémii CoViD-19 na sociálnych sieťach

¹Kristína MACHOVÁ, ²Branislav KOVALČIN

^{1,2} Katedra Kybernetiky and Umelej Inteligencie, Fakulta Elektrotechniky a Informatiky, Technická Univerzita v Košiciach, Slovenská Republika

¹kristina.machova@tuke.sk, ²branislav.kovalcin@student.tuke.sk

Abstrakt — Zanedlho potom čo sa koronavírus Covid-19 rozšíril do celého sveta, výrazne vzrástlo používanie sociálnych médií na komunikáciu medzi používateľmi. To zjednodušilo a zrýchlilo komunikáciu no taktiež vytvorilo výborné podmienky pre všetkých, ktorí šíria nepravdivé informácie. Práve sociálne siete ako napríklad Facebook, Twitter alebo YouTube sa v období pandémie stali hlavným zdrojom dezinformácií. Tieto nepravdivé informácie je potrebné odhaľovať a odstraňovať z online priestoru. Nemenej dôležitým spôsobom boja je aj pochopenie takýchto správ, kto ich vytvára, kto ich šíri, ako dlho trvá kým ich väčšina používateľov odhalí a podobne. Práve pochopením ich šírenia sa zaoberá táto publikácia. Snaží sa analyzovať šírenie nedôveryhodných správ. To koľko ľudí daným správam uverilo a ako dlho sa šíрили môže mať pozitívny efekt pri boji s podobným typom dezinformácií v budúcnosti. Článok sa zameriava na skúmanie možnosti použitia priehradkových modelov na modelovanie šírenia nedôveryhodných správ v online priestore.

KLúčové slová — falošné správy, dezinformácie, sociálne médiá, modelovanie šírenia dezinformácií, priehradkové modely

I. ÚVOD

Dezinformácie môžu umocniť najväčšie problémy ľudstva. Výrazným príkladom z poslednej doby je pandémia Covid-19, ktorá plodí množstvo nepravdivých informácií, aj keď pravda sa čoraz viac stáva otázkou života a smrti. Pandémia Covid-19 predstavuje zásadnú výzvu pre zachovanie globálneho ľudského blahobytu. Nanešťastie sa v sociálnych médiách rozšírili dezinformácie o Covid-19. V prípade Covid-19 majú tieto dezinformácie rôzne podoby - od konšpiračných teórií o tom, že vírus bol vytvorený ako biologická zbraň v Číne, až po tvrdenia, že kokosový olej vírus zabíja. V najhoršom prípade môžu dezinformácie tohto druhu spôsobiť, že sa ľudia obrátia k neúčinným (a potenciálne priamo škodlivým) prostriedkom, ako aj to, že budú zareagujú prehnane (napr. hromadením tovaru), alebo, čo je nebezpečnejšie, zareagujú nedostatočne (napr. úmyselným rizikovým správaním a neúmyselným šírením vírusu).

Dôležitou výskumnou výzvou je teda určiť, ako ľudia vyhľadávajú informácie, ako sa im vyhýbajú alebo ako sa samotné informácie šíria naprieč online priestorom. Najmä v tejto dobe, kedy je rozmach sociálnych médií silnejší každým dňom. Platformy sociálnych médií, ako sú YouTube a Twitter, poskytujú priamy prístup k bezprecedentnému množstvu obsahu a môžu posilňovať fámy a pochybné informácie. Niektoré štúdie poukázali na to, že falošné správy a nepresné informácie sa môžu šíriť rýchlejšie a mať väčší dosah ako správy založené na faktoch. Obzvlášť dôležitým to je pri celosvetovo závažných témach, ako je napríklad aj aktuálna pandémia. Zapríčinené je to tým že množstvo šírených správ je také veľké a dopyt po nich zo strany používateľov neustále rastie, z toho dôvodu je prakticky nemožné podchytiť tieto dezinformácie v čase ich uverejnenia.

II. FALOŠNÉ SPRÁVY

Falošné správy sú nepresné, niekedy senzačné správy, ktoré sú vytvorené s cieľom získať pozornosť, zavádzať, klamať alebo poškodiť dobré meno. Falošné správy sú vytvorené s úmyslom niekoho alebo niečo zmanipulovať. Môžu sa rýchlo šíriť, keď poskytujú dezinformácie, ktoré sú v súlade s názorom publika, pretože takýto obsah pravdepodobne nebude sponchybný alebo znevážnený.

V posledných rokoch internet poskytuje lacný distribučný kanál pre falošné správy. Uverejňovanie falošných správ na diskusných fórach, v komentároch k webovým stránkam, blogoch a postoch sociálnych médií si vyžaduje len malé, ak vôbec nejaké technické znalosti.

Najmä sociálne médiá sa ukázali ako jednoduché miesto na šírenie falošných správ. Falošné príbehy sa dajú napríklad tweetovať alebo zverejniť zo smartfónu a rýchlo sa šíria medzi veľké publikum prostredníctvom retweetov a zdieľania. Niektorí tvorcovia a distribútori falošných správ majú politické alebo sociálne ciele, iní sú viac podnikaví a používajú falošné správy, ktoré oslovujú príjemcov na emocionálnej úrovni, aby zarobili peniaze z digitálnej reklamy umiestnenej okolo obsahu. Keď sa falošné správy používajú na šírenie propagandy, môžu byť nebezpečné. Okrem toho, že formujú verejnú mienku a správanie, môžu tiež vyvolať nedôveru, podporiť nesúhlas a odvieť pozornosť od skutočných správ.

Správy na internete, ktoré sú vytvorené s úmyslom klamať, majú často zlú gramatiku a nesprávne napísané slová, môžu používať rasistický slovník alebo nadmerný počet veľkých písmen a výkričníkov. Na overenie správnosti správy môže byť užitočné si vyhľadať, že o danej správe informujú aj legitímne, tradičné spravodajské zdroje a to z toho dôvodu, že falošné správy majú často len jeden zdroj. Ďalšou stratégiou na identifikáciu falošných správ je kontrola názvu domény a adresy URL hostiteľskej stránky. Často sa zdá, že falošné správy majú legitímne znejúci názov domény, ale ich adresa URL sa končí na .com.co alebo inú nezvyčajnú koncovku.

A. Virálnosť Falošných Správ

Ako už bolo spomenuté, za vytváraním falošných správ môže byť viacero dôvodov, a jedným z nich je to, aby sa tieto správy dostali k čo najväčšiemu počtu používateľov. Tento dosah sa nedosiahne len tým, že článok uverejníte a potom čakáte, kým ho ľudia uvidia. Stane sa, že súbor používateľov tento článok navštívi a potom ho niektorí z nich zdieľajú s ďalšími používateľmi a proces sa opakuje. Prečo však ľudia tieto správy vôbec zdieľajú, keďže sú falošné? Čo ich k tomu vedie? Prvým dôvodom je skutočnosť, že je čoraz ťažšie identifikovať zdroj správ ako falošný. Ako už bolo spomenuté, väčšina webových stránok s falošnými správami je navrhnutá tak, aby vyzerali dôveryhodne (napríklad po vizuálnej stránke, ale aj názvami, tónom písania a adresami URL) s cieľom získať si dôveru čitateľa. Hranica medzi dôveryhodnými a overenými zdrojmi informácií a zdrojmi falošných informácií sa teda čoraz viac stiera. Po druhé sa jedná o skutočnosť, že väčšina falošných správ hrá na naše emócie. Väčšina titulkov, ktoré vidíme na webových stránkach s falošnými správami, má tendenciu vyvolať v nás nejakú emóciu: prekvapenie, nenávisť, strach, hnev, nádej, ..., a to platí aj pre obrázky. Takže tým, že v čitateľoch je prebudená nejaká emócia, je pravdepodobnejšie, že budú s článkom komunikovať a zdieľať ho. Toto je to, čo sa nazýva "clickbaiting", termín na označenie procesu a zámeru umiestnenia senzáčných titulkov a/alebo obrázkov, ktoré majú čitateľa zaujať a nechať ho v napätí s cieľom prilákať kliknutia s konečným cieľom získať príjmy z online reklamy.

Napokon je tu ešte jeden psychologický aspekt, ktorý nás vedie k interakcii s titulkami falošných správ: kognitívne predsudky. Zjednodušene povedané, tie spočívajú v odchýlke od racionálneho myslenia, keď si jednotlivci vytvárajú svoju subjektívnu realitu, čo môže viesť k správaniu, ako sú nepresné úsudky alebo nelogické interpretácie. A práve to sa stane, keď falošný článok zdieľa toľko ľudí, na ktorých berieme ohľad, ako sú blogeri, celebrity, rodina a priatelia, že máme tendenciu veriť, že tento článok skutočne hovorí pravdu. To sa stáva aj vtedy, keď vidíme tie isté falošné články mnohokrát: na prvý raz sa nám to môže zdať smiešne, ale pri druhom, treťom alebo desiatom opakovaní sa ten článok stáva čoraz hodnovernejším. A nielen to, ale keď vidíme falošné spravodajské články, ktoré potvrdzujú niektoré naše presvedčenia, máme tiež tendenciu ľahšie ich prijať a v niektorých prípadoch zdiskreditovať zdroje, ktoré ich nepotvrdzujú, čím sa dostávame do tzv. komory ozveny (ang. "echo chamber").

III. MODELOVANIE ŠÍRENIA FALOŠNÝCH SPRÁV

A. Model Založený na Dôveryhodnosti

Z dôvodu presnejšej reprezentácie šírenia správ po sieti je nutné vytvoriť orientovaný graf siete ($A=(N,E)$), v ktorom jednotlivé uzly grafu predstavujú používateľov a hrany predstavujú prepojenia medzi nimi [1]. Zároveň vytvoríme orientovaný graf dôveryhodnosti ($C=(N,E',c)$), ktorý bude obsahovať hrany opačného smeru ako pôvodná sieť. Dôvodom je že ak uzol pošle správu, tak prijímajúci uzol si vytvorí názor na odosielateľa pomocou váhy $c \in [-1,1]$, kde -1 označuje najvyššiu dôveryhodnosť, zatiaľ čo 1 označuje najnižšiu.

V orientovanom grafe je základným predpokladom to, že medzi párami jednotlivých uzlov nie je nevyhnutne spoločný záujem o zverejnené príspevky, takže vstupné a výstupné hrany siete majú rozdielne roly: odchádzajúce hrany predstavujú spojenie so záujemcami o to, čo zverejňujeme, zatiaľ čo prichádzajúce sú zdrojom inšpirácie pre naše záujmy. Z tohto dôvodu zavádzame do siete založenej na dôveryhodnosti siete dve premenné: dôveru a spoľahlivosť.

Dôvera (trust - T) v jednotlivca naznačuje, do akej miery je jeho susedmi považovaný za dôveryhodného. Vysoké hodnoty dôvery naznačujú, že kto je s ním v kontakte, oceňuje obsah,

ktorý zverejnil, a považuje ho za osobu, ktorá overuje správy pred ich opätovným zverejnením (súvisí to s prichádzajúcimi hranami siete založenej na dôveryhodnosti).

Spôľahlivosť (reliability - R) jednotlivca ukazuje jeho schopnosť vybrať si, od ktorých susedov bude správy prijímať a preposielať ich ďalej, preto tento parameter nepriamo ovplyvňuje jeho schopnosť zdieľať skutočné správy, inými slovami, súvisí s odchádzajúcimi hranami siete založenej na dôveryhodnosti.

V rámci sociálnych sietí má veľa používateľov tendenciu nadväzovať kontakty s ostatnými, ktorí zdieľajú rovnaké správy, zatiaľ čo používatelia so zameraním na šírenie dezinformácií si vytvárajú viac účtov na preposielanie správ. Členovia týchto dvoch skupín účtov si navzájom zvyšujú dôveru, takže zvyšku siete, je možné priradiť nízku úroveň spoľahlivosti, aby sa vyrovnala vysoká úroveň dôvery.

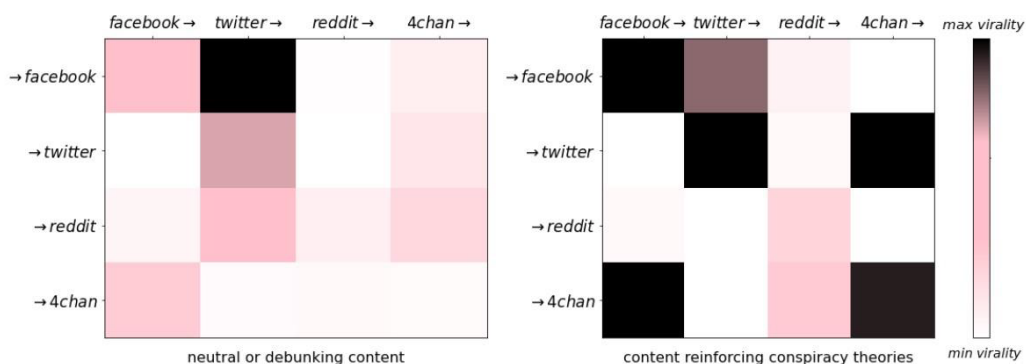
Poznáme viaceré modely na simuláciu šírenia správ[1]:

- Nekonkurenčný (No-competitive) model šírenia
- Konkurenčný (Competitive) model šírenia
- Konkurenčný (Competitive) model šírenia s oneskorením (Competitive with Delay)

B. Vplyv Moderovania Obsahu na Šírenie Konšpirácií

V štúdiu [2] sa zameriavali na pochopenie toho, ako sa konšpiračné teórie šíria na začiatku krízy COVID-19, na základe toho odhaľovali možnosti a problémy kontroly faktov a moderovania príspevkov v systéme sociálnych médií. Aby to dosiahli, zisťovali počet výskytov príbehov posilňujúcich konšpiračné teórie na štyroch platformách: Facebook, Twitter, Reddit a pod-sekcii 4chan s názvom “politically incorrect” (politicky nekorektné), čo je významné fórum konšpiračných teoretikov. Počas skúmaní zistili, že príbehy posilňujúce konšpiračné teórie sa stali virálnejšie ako príbehy, ktoré ich buď odhaľujú, alebo majú neutrálny postoj. Väčšina príbehov posilňujúcich konšpiračné teórie pochádzala z alternatívnych zdrojov, osobných blogov a príspevkov na sociálnych sieťach. Aj keď alternatívne a iné zdroje boli hlavnými autormi konšpiračných teórií, bežné zdroje mali väčší vplyv na šírenie konšpiračných teórií. V práci boli preskúmané spôsoby moderovania jednotlivých platforiem, ktoré sa do istej miery líšili. Twitter a YouTube odstránili príbehy, ktoré podporovali konšpiračné teórie, zatiaľ čo Reddit a Facebook ich odstránili alebo označili. Odstránenie alebo nahlásenie v službe Reddit záviselo od pravidiel každej subkomunity, zatiaľ čo na Facebooku to, či spoločnosť skontrolovala samotné príbehy (odstránené), alebo sa spoľahla na overovače faktov tretích strán (označené).

Záverom tejto štúdie bolo, že postupy moderovania jednotlivých platforiem výrazne znížili pravdepodobnosť opätovného objavenia sa príbehov v celom systéme. Teoreticky by teda okamžité odstránenie alebo filtrovanie konšpiračných teórií zastavilo ich šírenie. Moderovanie obsahu je však zložitý a časovo náročný proces. Spočíva to jednak vo veľkom množstve obsahu, ktorý treba skontrolovať, ale taktiež v povahe obsahu, ktorý je často ťažké klasifikovať ako konšpiračnú teóriu alebo nie. Platformám sa podarilo skontrolovať iba niečo medzi 15% až 50% príspevkov obsahujúcich príbehy posilňujúce konšpiračné teórie, pričom v mnohých prípadoch k tomu došlo týždeň po tom, čo sa stali virálnymi. Obrázok 1 ilustruje heatmapu zobrazujúca viralitu konšpiračných teórií o pôvode COVID-19. Čím je farba tmavšia, tým vyššia je viralita webových adries. Heatmapa popisuje, ako každá platforma v stĺpci ovplyvnila ostatné platformy. Napríklad prvý stĺpec ukazuje, ako príbehy na Facebooku ovplyvnili viralnosť podobných príbehov na Facebooku, Twitteri, Reddite a 4chane.



Obr. 1. Heatmapa virálnosti fake news na sociálnych sieťach.

C. Šírenie Medicínskych Falošných Správ na Sociálnych Médiách

V štúdiu [3] skúmali najčastejšie zdieľané články v poľskom jazyku pomocou kľúčových slov súvisiacich s chorobami považovanými za najdôležitejšie príčiny úmrtia v Poľsku - mozgová

mŕtvica, ischemická choroba srdca a rakovina (spolu tvoria > 60% celkových úmrtí), prevládajúce chronické choroby - arteriálna hypertenzia a cukrovka mellitus (32%, respektíve 9,5% z celkovej populácie) a navyše - očkovanie a HIV / AIDS. V z celej množiny dát bolo 40% odkazov vyhodnotených ako falošné správy. Za zmienku stojí, že viac ako 20% z nich bolo generovaných z jedného zdroja, zvyčajne s dostatkom nepravdivých informácií. Podobné opisy chorôb a modely riadenia sa našli vo viacerých kategóriách. Veľká časť zdieľaných správ sa týkala alternatívnej medicíny, ale bolo zaznamenané aj trvalé popieranie AIDS alebo očkovanie.

Takáto analýza sociálnych médií by mohla významne prispieť k správe webových zdrojov zverejňujúcich lekárske dezinformácie. Niektoré stránky obsahujú hlavnú časť celkového obsahu zameraného na niekoľko málo populárnych tém, ako sú vakcíny alebo HIV / AIDS. V týchto konkrétnych prípadoch môže byť zavedené ciele blokované alebo jednoducho varovania pre konkrétne pseudovedecké domény. Obmedzenia uvedené na týchto webových stránkach môžu významne znížiť celkový podiel nebezpečných informácií. Ďalej by poskytovatelia vyhľadávačov mohli pre tieto stránky nastaviť nižšiu prioritu, aby sa znížil počet celkových návštev webu.

IV. MODELOVANIE ŠÍRENIA SPRÁV O COVID-19

A. Popis Dátových Množín

Množinu dát využívanú v prezentovaných experimentoch vytvorila skupina výskumníkov, Knuutila a kol, [4]. Dôvodom pre jej vytvorenie bolo následne využitie v rámci výskumnej činnosti zameranej na súčasnú pandemickú situáciu. Konkrétne na šírenie dezinformácií naprieč viacerými sociálnymi sieťami. Jedná sa o množinu údajov obsahujúcu metadáta o videách z platformy YouTube súvisiacich s COVID, ktoré kolovali na verejných sociálnych sieťach, no YouTube ich nakoniec odstránil, pretože obsahovali nepravdivé informácie. Obsahuje popisné údaje o 8 122 videách, ktoré boli zdieľané medzi novembrom 2019 a júnom 2020. Súbor údajov obsahuje jedinečné identifikátory účty na sociálnych sieťach, ktoré videá zdieľali a metadáta, ako sú názvy videí a počty zhliadnutí.

B. Použité Metódy

Na modelovanie šírenia správ o CoViD-19 boli zvolené priehradkové modely SIR, SIRS, SEIR a SEIRS.

Keďže šírenie falošných správ možno prirovnať k spôsobu vzniku epidémie, môžeme použiť modely epidémie na lepšie pochopenie toho, ako falošné správy kolujú po internete a aký dosah majú. Jedným z týchto modelov je model SIR [5], epidemiologický model, ktorý zjednodušene opisuje prenos infekčnej choroby prostredníctvom jednotlivcov v populácii. Použitie modelu SIR má však problém: predpokladá, že všetci používatelia, ktorí prestali zdieľať falošnú správu majú celoživotnú imunitu. Znamená to, že všetci používatelia, ktorí raz uvidia falošnú správu, ju buď neuvidia znova, alebo ju budú ignorovať. V skutočnosti však, ak sa zamyslíme hlbšie, používatelia, ktorí už predtým videli alebo zdieľali falošnú správu, ju stále môžu znova zdieľať alebo hovoriť o danej téme, čo znamená, že sú náchylní na to, aby sa znova nakazili a znova preniesli falošnú informáciu na ostatných. To znamená, že nedôjde k úplnej obnove populácie, pretože uzdravení používatelia sa môžu vrátiť do medzi náchylných a zopakovať proces šírenia. Toto nie je znázornené v grafe vyššie, kde sme skončili so všetkými „vylicenými“ používateľmi a žiadnym z nich náchylným. To vôbec neznamená, že použitie modelu SIR je úplne nesprávne, pretože to závisí od našej interpretácie a modelovania problému, ale pre presnejšiu reprezentáciu reality vieme zdefinovať aj odlišné prístupy.

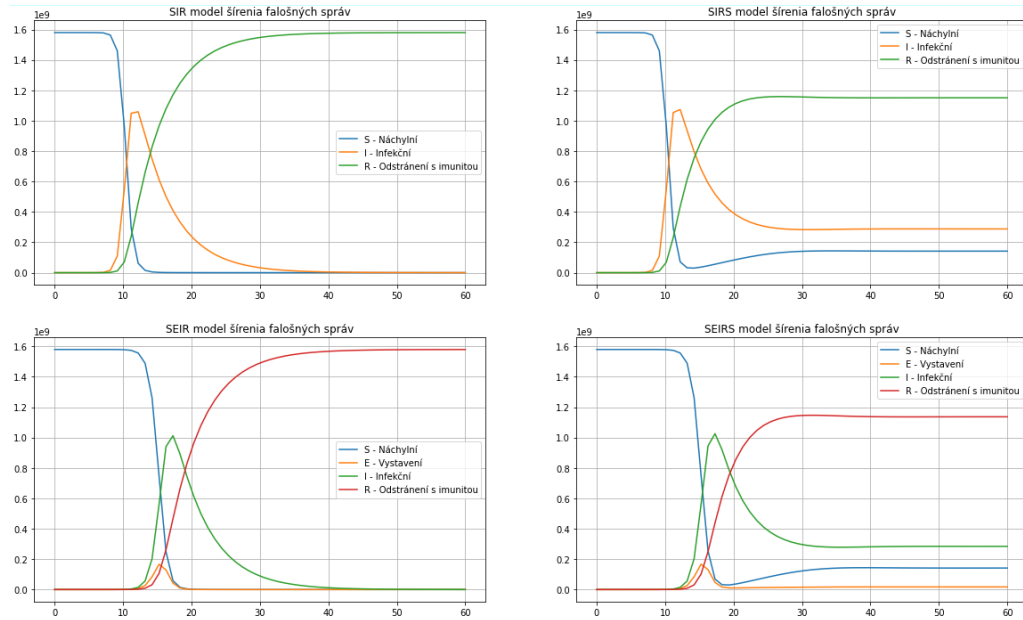
Model SIRS [6] je veľmi podobný modelu SIR, ale pridáva nový krok, označený dodatočným "S". To znamená, že podobne ako v prípade SIR, vylicený používateľ získava imunitu voči chorobe, avšak SIRS predpokladá, že je imúnny len na určité obdobie, po ktorom môže byť náchylný na opätovné ochorenie.

Zloženie modelu SEIR [7] je vo veľkej časti podobné ako pri predchádzajúcich modeloch, či už to je SIR alebo SIRS. V tomto konkrétnom prípade je ale pridaná nová kategória do ktorej môžu používatelia patriť. Danou kategóriou sú vystavení (exposed) používatelia a označujeme ju ako E. V prípade modelov o šírení choroby by osoby nachádzajúce sa v tejto kategórii boli vytavené ochorením a práve by sa nachádzali v inkubačnej dobe ochorenia. Analogicky pre model šírenia falošných správ vieme túto skupinu používateľov popísať ako používateľov, ktorí boli vystavení dezinformácii ale ešte sa ju nerozhodli šíriť ďalej.

Ďalším možným krokom po zedefinovaní predchádzajúcich modelov je ich spojenie pre rozšírenie funkcionality. Spojením jednotlivých častí z predchádzajúcich modelov dostaneme model SEIRS [8], ktorý posúva celú simuláciu bližšie k realite. Ako už aj názov napovedá, výsledný model dostaneme spojením modelov SEIR a SIRS. Jedná sa o ďalšiu možnú modifikáciu základného modelu SIR, je však rozšírený o skupinu „Exposed“, čiže používateľov,

ktorí boli vystavení dezinformácii ale ešte sa ju nerozhodli šíriť ďalej. Taktiež bola pridaná miera straty imunity, označená ako ξ .

Nasleduje konkrétna simulácia záznamu číslo 4209. Dané video bolo online desať dní a v priebehu toho dosiahlo takmer 7 miliónov zhladnutí. Ak ale nastavíme základné parametre simulácie podľa daného článku, čiže parameter šírenia $\beta = 0.3$ a mieru zotavenia $\gamma = 0,11$. Celkový počet používateľov bude už vyššie spomenutých 1,58 miliardy a infikovaný bude na začiatku simulácie len jeden používateľ. Výsledky simulácie môžeme vidieť na Obrázku 2.



Obr. 2. Simulácia modelov SIR, SIRS, SEIR SEIRS pri manuálnom nastavení rýchlosti učenia (S – Náchylní (modrá farba), I – Infekční (žltá pre SIR a SIRS a zelená pre SEIR a SEIRS), R – Odstránení s imunitou (zelená pre SIR a SIRS a červená pre SEIR a SEIRS), a E – Vystavení (žltá pre SEIR a SEIRS)).

Grafy na Obrázku 2 nám zobrazujú všetky štyri využívané typy modelov. Os y reprezentuje počet používateľov a os x čas, ktorý uplynul od začiatku simulácie.

Všetky modely nám ukazujú rýchly nárast šírenia falošnej správy a postupný pomalý úpadok jej popularity. Prvý väčší rozdiel si vieme všimnúť medzi modelmi ktoré nevyužívajú kategóriu používateľov, ktorí sú vystavení falošnej správe a tými ktoré áno. V prípade SIR a SIRS modelu je nárast popularity správy omnoho strmší, začína skôr a dosahuje vyšší počet infikovaných používateľov. Je to zapríčinené tým že každý používateľ najprv určitú dobu zostáva v kategórii vystavených a neprejde automaticky do kategórie infikovaných. Toto nám približuje model reality a to z toho dôvodu že nie každý kto uvidí nejakú správu jej začne bezprostredne potom veriť a šíriť ju ďalej.

V prípade modelov SIRS a SEIRS vidíme že počty infikovaných sa nedostanú do nuly ako v prípade SIR a SEIR. Týmto vieme simulovať fakt že vždy sa nájde určitá vzorka populácie ktorá neuverí skutočnosti a zostanú presvedčení o pravdivosti falošnej správy. Z toho nám vyplýva, že model SEIRS je spomedzi týchto štyroch modelov najbližšie reality, čo je aj očakávaný fakt keďže sa jedná o najzložitejší model zo spomínaných modelov.

V. ZÁVER

V rámci publikácie sú priblížené možnosti modelovania šírenia nedôveryhodných správ o pandémie Covid-19 v online priestore. Boli použité priehradkové modely (ang. Compartmental models), konkrétne modely SIR, SIRS, SEIR a SEIRS. Nedôveryhodné správy o pandémie boli reprezentované pomocou množiny videí z YouTube, ktoré boli odstránené lebo obsahovali nepotvrdené informácie a tým porušovali pravidlá platformy.

Dané modely boli otestované na vybraných záznamoch z datasetu a pri rôznych metódach nastavenia parametrov šírenia. Najdôležitejším z parametrov sa stala rýchlosť šírenia falošnej správy. Ako najpresnejší spôsob sa v našom prípade javí jeho manuálne nastavenie aby v momente zmazania správy zodpovedali výsledky modelu záznamom z datasetu. Následne vieme odhadnúť aký veľký mohli mať dané správy dosah ak by neboli mazané a nebola nijaká kontrola pravosti obsahu na internete.

Avšak z testov sa ukázalo že zvolená metóda nie je vhodná na presnú reprezentáciu šírenia dezinformácii naprieč online priestorom. Poskytuje nám skôr približný popis doby šírenia a ich dosahu, na tento účel sú priehradkové modely vhodné pre nízku zložitosť a nevelké nároky na vstupné údaje.

POĎAKOVANIE

Táto práca vznikla s podporou Slovenskej grantovej agentúry Ministerstva školstva a SAV Slovenskej republiky v rámci projektu VEGA 1/0685/21.

LITERATÚRA

- [1] V. Carchiolo, A. Longheu, M. Malgeri, G. Mangioni, M. Previti. *Credibility-based Model for News Spreading on Online Social Networks*. [Online] 2020. [19.05.2023] <https://www.scitepress.org/Papers/2020/93405/93405.pdf>. ISBN: 978-989-758-427-5
- [2] O. Papakyriakopoulos, J. C. Medina Serrano, S. Hegelich. *The spread of COVID-19 conspiracy theories on social media and the effect of content moderation*. [Online] 17.8.2020. [Dátum: 19.05. 2023] DOI: <https://doi.org/10.37016/mr-2020-034>
- [3] P. M. Waszak, W. Kasprzycka-Waszak, A. Kubanek. *The spread of medical fake news in social media – The pilot quantitative study*. [Online] 27.4.2018. [Dátum: 19.05.2023] DOI: <https://doi.org/10.1016/j.hlpt.2018.03.002>
- [4] A. Knuutila, A. Herasimenko, H. Au, J. Bright, P. N. Howard. *A Dataset of COVID-Related Misinformation Videos and their Spread on Social Media*. [Online] [Dátum: 19.05.2023] DOI:<https://doi.org/10.5281/zenodo.4557828>
- [5] Mathur , C. P. Gupta. *Dynamic SEIZ in Online Social Networks: Epidemiological Modeling of Untrue Information*. [Online] 8.2020. [Dátum: 15.05.2023] DOI: 10.14569/IIACSA.2020.0110771
- [6] E. Oktaviansyah, A. Rahman. *Predicting hoax spread in Indonesia using SIRS model*. [Online] 3.2020. [Dátum: 15.05.2023] DOI: 10.1088/1742-6596/1490/1/012059
- [7] D. Suyalatu, D. Yan-Bin, H. Yong-Chang. *SEIR Model of Rumor Spreading in Online Social Network with Varying Total Population Size*. [Online] 9.2017. [Dátum: 19.05.2023] DOI: 10.1088/0253-6102/68/4/545
- [8] O.N. Bjørnstad, K. Shea, M. Krzywinski, N. Altman. *The SEIRS model for infectious disease dynamics*. [Online] 4.6.2020. [Dátum: 19.05.2023] DOI: <https://doi.org/10.1038/s41592-020-0856-2>

Viacportový menič založený na MAB

¹Adrián MARCINEK, ²Marek PÁSTOR

¹ Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,
Technická univerzita v Košiciach, Slovenská republika

² Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,
Technická univerzita v Košiciach, Slovenská republika

¹adrian.marcinek@tuke.sk, ²marek.pastor@tuke.sk

Abstrakt — Článok predstavuje návrh modulárneho viacportového meniča, ktorý využíva multi-aktívne mostíky (MAB). Príspevok popisuje modulárnu štruktúru a začína vykonaním analýzy ustáleného stavu na vyhodnotenie charakteristík meniča. Následne sa odvodí rovnice pre výkon dodávaný/odoberaný jedným portom a tiež maximálny výkon meniča. V závere príspevku sú uvedené možnosti konfigurácie topológie MAB meniča pre dosiahnutie vyššieho výkonu prenášaného jednotlivými portami.

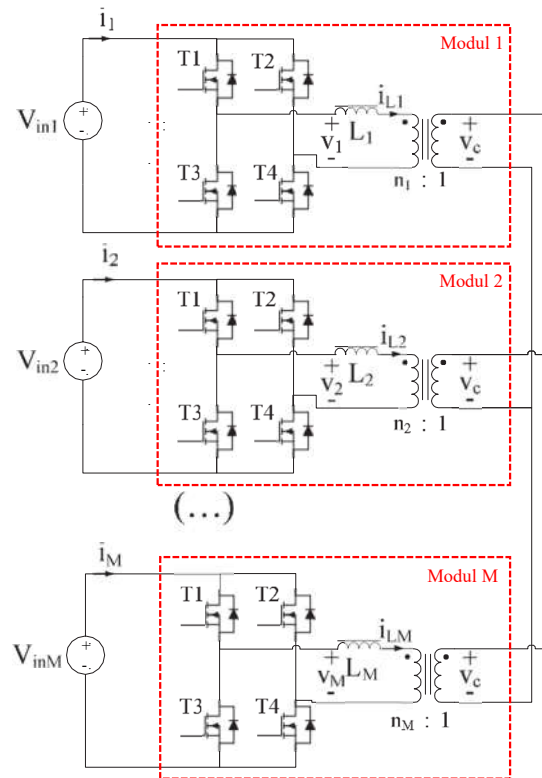
Kľúčové slová — viacportový menič, multi-aktívny mostík, MAB

I. ÚVOD

Súčasný stav systémov manažmentu elektrickej energie súvisí s využívaním rôznych zdrojov energie, ich vysokou flexibilitou a lokálnejším využívaním a výrobou elektrickej energie. V tejto súvislosti sú veľmi dôležité výkonové architektúry založené na viacportových výkonových meničoch, skladovaní energie a viacerých energetických tokoch.

Viacportové meniče energie boli v posledných rokoch široko používané ako kľúčový prvok pri budovaní hybridných zdrojov energie [1], elektrických vozidlách [2][3], leteckých aplikáciách [4][5], obnoviteľných zdrojoch energie [6][7], atď. Typický viacportový systém pozostáva z hlavného zdroja energie (palivový článok, FV panely, usmerňovač pripojený k elektrickej sieti atď.), úložných prvkov (batérie, superkondenzátory atď.) a záťaží, ktoré môžu byť v niektorých prípadoch aktívne (napr. motory). Hybridné zdroje energie umožňujú integráciu pomalých dynamických zdrojov (napr. palivové články), nekonštantných zdrojov energie (napr. FV panely) alebo zdrojov nízkeho špičkového výkonu s akumuláčnými prvkami, ktoré môžu dodávať energiu do záťaže, keď to nie je schopný zabezpečiť hlavný zdroj. V iných aplikáciách tieto systémy umožňujú rekuperáciu energie z aktívnych záťaží (napr. motorov).

Jedným z hlavných problémov v systémoch s viacerými zdrojmi energie je topológia napájania, t. j. rozhranie na pripojenie záťaží, zdrojov a úložných prvkov. Mnoho aplikácií je založených na topológiách so spoločnou jednosmernou zbernicou (DC BUS), kde sú všetky prvky, vrátane prvkov na ukladanie energie, prepojené cez jednosmerné meniče [4]. Hlavnými požiadavkami na tieto meniče je obojsmerný tok energie a galvanické oddelenie. Jedným z najznámejších typov takýchto meničov je tzv. Dual Active Bridge (DAB), ktorý umožňuje ako obojsmerný tok energie tak aj galvanické oddelenie [8]. V posledných rokoch sa však skúmali aj modulárne prístupy k návrhu výkonových topológií [9][10][11]. Modulárny prístup pre viacportový systém spočíva v modularizácii prvkov systému ako sú: zdroj hlavného napájania, úložné prvky a záťaž. Modularizácia takýchto systémov umožňuje veľkú flexibilitu: v elektrických vozidlách, kde pre rôzne jazdné cykly je možné vyžadovať rôznu veľkosť prvkov pre ukladanie energie. Modulárny energetický systém je zaujímavý aj pre komerčné použitie, pretože s jedným zariadením je možné pokryť mnoho rôznych špecifikácií toku výkonu.



Obr. 1 Topológia multi-aktívneho meniča

Viacportové systémy založené na meničoch odvodených od DAB boli opísané v [12] a [13]. DAB je založený na jedinom transformátore, ktorý spája všetky porty. Akonáhle je však systém navrhnutý a postavený, jeho architektúru už nie je možné meniť.

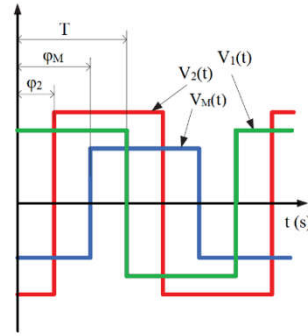
V tomto článku je analyzovaný tzv. Multi Active Bridge (MAB) (Obr. 1) menič. Je odvodený od DAB a je založený na prepojení viacerých rôznych modulov. Každý modul obsahuje tzv. plný mostík a transformátor, ktorý je rozdelený na sériovú indukčnosť predstavujúcu rozptylovú indukčnosť transformátora a prevod transformátora. Moduly sú navzájom prepojené a je medzi nimi možný obojsmerný tok výkonu. Tento systém je navyše plne škálovateľný a modulárny. Topológia MAB meniča je na Obr.1 a je elektricky ekvivalentná s [13], avšak magnetická väzba medzi mostíkmi je dosiahnutá pomocou fyzicky nezávislých transformátorov. Predstavená topológia umožňuje vysokú flexibilitu, pretože počet modulov v systéme možno ľahko zmeniť pridaním alebo odstránením polovodičového mostíka a zodpovedajúceho transformátora. Analýza multi-aktívneho mostíka v [13] je založená na interakcii portov medzi sebou. V [14] a [15] je analyzovaný a modelovaný podobný menič ako v tomto príspevku.

Príspevok je zameraný na analýzu pomerov v obvode potrebnú pre návrh takéhoto meniča. Predložená analýza meniča MAB je zameraná na ustálený stav, kde sa analyzuje vzťah medzi strednou hodnotou prúdu každého modulu a fázovými posunmi medzi nimi. Opísaný je aj jednoduchý postup návrhu na dosiahnutie požadovaného výkonu v každom porte a v závere je uvedená krátka analýza zmeny vo výkonových pomeroch pri rôznych možnostiach zapojenia multi-aktívneho mostíka so štyrmi portami.

II. MULTI-AKTÍVNY MOSTÍK (MAB)

Náhradná schéma pre MAB menič s počtom modulov, ktorý je označený ako „M“, je znázornená na Obr. 1. Na Obr. 2 sú znázornené priebehy napätí pre tri moduly. Všetky moduly pracujú s rovnakou spínacou frekvenciou a vetvy daného mostíka sú komplementárne spínané. V tejto analýze sa neuvažovalo o riadení so zmenou striedy. Napätie k – teho mostíka je reprezentované pravouhlým priebehom s amplitúdou rovnou V_{ink} a s určitým fázovým posunom vzhľadom na ostatné mostíky (Obr. 2). Energia prenášaná medzi modulmi je založená na fázovom posune napätí generovaných mostíkmi. Fázové posuny $\varphi_2, \varphi_k, \text{atd.}$, sú definované ako časové oneskorenie pri spínaní daného mostíka v rámci polperiódy spínania T , čo je tiež možné vidieť na Obr.2. Mostík 1 predstavuje časovú referenciu, preto $\varphi_1 = 0$. V tomto prístupe uvažujeme, že všetky mostíky majú na vstupe zdroj napätia zodpovedajúci zdroju energie s akumulárnym

prvkom (superkondenzátory, batérie atď.). Počas analýzy sa tieto zdroje napätia považujú za ideálne a magnetizačná indukčnosť transformátorov sa považuje za nekonečnú.



Obr. 2 Časové priebehy napätí MAB meniča

A. Popis MAB meniča v ustálenom stave

Cieľom tejto analýzy je nájsť závislosť strednej hodnoty vstupných prúdov vzhľadom na fázový posun φ_k , vstupné napätia V_{ink} a ďalšie parametre obvodu (n_k, L_k, T).

Prvým analyzovaným priebehom je napätie v_c na sekundárnej strane transformátorov, ktoré sú zapojené paralelne. Napätie v_c je neregulované striedavé napätie. Vyjadrenie v_c ako funkcie mostíkových napätí v_k a konštrukčných parametrov transformátorov (n_k, L_k, T) je (1).

$$v_c = \frac{\frac{v_1}{n_1} \cdot \frac{1}{L_1} + \frac{v_2}{n_2} \cdot \frac{1}{L_2} + \dots + \frac{v_M}{n_M} \cdot \frac{1}{L_M}}{\frac{1}{L_1} + \frac{1}{L_2} + \dots + \frac{1}{L_M}} = \frac{\sum_{k=1}^M v_k \cdot \frac{n_k}{L_k}}{\frac{1}{L_{eq}}} \quad (1)$$

$$L_{eq} = \left(\sum_{k=1}^M \frac{n_k^2}{L_k} \right)^{-1} \quad (2)$$

Kde, v_c je vážený priemer napätí mostíkov. Jednotlivé koeficienty závisia od pomeru závitov každého transformátora n_k , rozptylovej indukčnosti každého transformátora L_k a parametra L_{eq} (2) nazývaného aj ako ekvivalentná indukčnosť. L_{eq} môžeme určiť ako paralelnú kombináciu rozptylových indukčností prepočítaných na sekundárnu stranu transformátorov. Znalosť v_c je dôležitá pre návrh transformátorov, pretože je z nej možné určiť tzv. volt-sekundu pôsobiacu na vinutie transformátora a vhodne navrhnuť jadro a počet závitov. Strednú hodnotu vstupného prúdu mostíka k možno vypočítať lineárnou superpozíciou napätí ostatných mostíkov v_j . Potom, časť vstupného prúdu i_k vzhľadom na mostík j , môžeme nazvať i_{kj} a je daný vzťahom (3). Tento výraz je veľmi podobný prúdu v DAB meniči [8], ale v tomto prípade sa napätie sekundárnej strany prejaví na všetkých indukčnostiach, cez parameter L_{eq} (4). Nakoniec je možné vypočítať celkový prúd mostíka k pripočítaním jednotlivých častí prúdu i_k vytvorených dôsledkom iných mostíkov (5).

$$i_{kj} = K_{kj} \cdot V_{inj} \cdot (\varphi_j - \varphi_k) \cdot (1 - 2 \cdot |\varphi_j - \varphi_k|) \quad (3)$$

$$K_{kj} = T \cdot \frac{n_k \cdot n_j}{L_k \cdot L_j} = \frac{n_k \cdot n_j \cdot L_{eq}}{L_k \cdot L_j \cdot f} \quad (4)$$

$$i_k = \sum_{j=1}^M K_{kj} \cdot V_{inj} \cdot (\varphi_j - \varphi_k) \cdot (1 - 2 \cdot |\varphi_j - \varphi_k|) \quad (5)$$

Rozšírením rovníc (3)-(5) pre ďalšie porty je výraz týkajúci sa stredných hodnôt vstupných prúdov i_k , fázových posunov φ_k a ostatných parametrov (n_k, L_k, T) uvedený v (6). Keďže mostík 1 predstavuje časovú referenciu, potom $\varphi_1 = 0$ a platí tiež rovnica pre rovnováhu výkonu v MAB meniči (7).

$$\begin{bmatrix} i_1 \\ i_2 \\ \vdots \\ i_M \end{bmatrix} = \begin{bmatrix} 0 & K_{12} \cdot (\varphi_2 - \varphi_1)(1 - |\varphi_2 - \varphi_1|) & \cdots & K_{1M} \cdot (\varphi_M - \varphi_1)(1 - |\varphi_M - \varphi_1|) \\ K_{21} \cdot (\varphi_1 - \varphi_2)(1 - |\varphi_1 - \varphi_2|) & 0 & \cdots & K_{2M} \cdot (\varphi_M - \varphi_2)(1 - |\varphi_M - \varphi_2|) \\ \vdots & \vdots & \ddots & \vdots \\ K_{M1} \cdot (\varphi_1 - \varphi_M)(1 - |\varphi_1 - \varphi_M|) & K_{M2} \cdot (\varphi_2 - \varphi_M)(1 - |\varphi_2 - \varphi_M|) & \cdots & 0 \end{bmatrix} \cdot \begin{bmatrix} V_{in1} \\ V_{in2} \\ \vdots \\ V_{inM} \end{bmatrix} \quad (6)$$

$$\sum_{k=1}^M i_k \cdot V_{ink} = 0 \quad (7)$$

B. Analýza toku výkonu v MAB meniči

Z predchádzajúcej analýzy možno usúdiť, že maximálny prúd (resp. výkon), ktorý je modul schopný poskytnúť, je obmedzený napäťovými úrovňami a konštrukčnými parametrami (počet závitov, rozptylová indukčnosť a spínacia frekvencia). Táto problematika bola hlbšie analyzovaná v [13] a je veľmi dôležitá pre modulárny prístup. Kľúčovým bodom je poznať menovitý výkon modulu a vplyv na výkon, ktorý môže modul spracovať v závislosti od ich počtu.

Pre ďalšiu analýzu uvedieme niektoré podmienky a vzťahy kde $\frac{V_{ini}}{n_i} = \frac{V_{inj}}{n_j} = V_0$ a $\frac{L_i}{n_i^2} = \frac{L_j}{n_j^2} = L_0$, ďalej $\varphi_k = 0$ a $\varphi_j = 0,25$ pre všetky $j \neq k$. Čo znamená, že zdroj predstavuje iba port k a všetky ostatné porty predstavujú záťaž. Výkon prenášaný z portu k na ďalšie porty je potom maximálny. Pri takýchto podmienkach je potom výkon prenášaný portom k rovný (8), kým maximálny možný výkon prenášaný jedným portom je (9).

$$P_{kmax} = V_{ink} \cdot i_{kmax} = \frac{(M-1)}{M} \cdot \frac{V_0^2 \cdot T}{4 \cdot L_0} \quad (8)$$

$$P_{max} = \frac{V_0^2 \cdot T}{4 \cdot L_0} \quad (9)$$

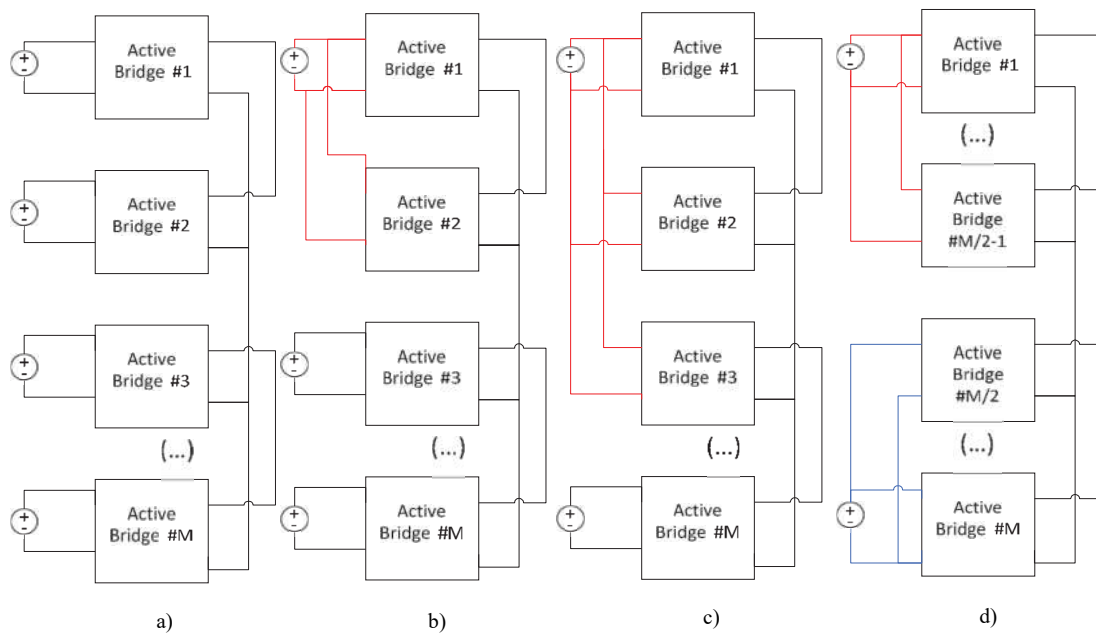
Čím vyšší je počet modulov, tým vyšší je maximálny možný výkon v jednom porte. Táto závislosť je však asymptotická a prírastok pre P_{kmax} je nízky pre stredný alebo vysoký počet modulov. V MAB meniči nie je možné ľubovoľne priamo presmerovať výkon z jedného portu na druhý. Úprava fázových posunov umožňuje upraviť distribúciu energie, ale medzi modulmi vždy existuje vzájomná závislosť. V MAB meniči je energia poskytovaná ďalším novým modulom vždy rozdelená medzi ostatné moduly.

V topológií MAB meniča môže nastať situácia, kedy dva porty meniča sú pripojené k rovnakému zdroju napájania, aby sa zvýšil výkon poskytovaný týmto zdrojom napájania. V tomto prípade $\varphi_k = \varphi_{k+1} = 0$ pričom $\varphi_j = 0,25$. Pomocou rovnakého prístupu ako pre rovnicu (8) môžeme pre výkon prenášaný portom k a portom $k+1$ uviesť rovnice (10) a (11).

$$P_k = V_{ink} \cdot i_k = \frac{(M-2)}{M} \cdot \frac{V_0 \cdot T}{4 \cdot L_0} \quad (10)$$

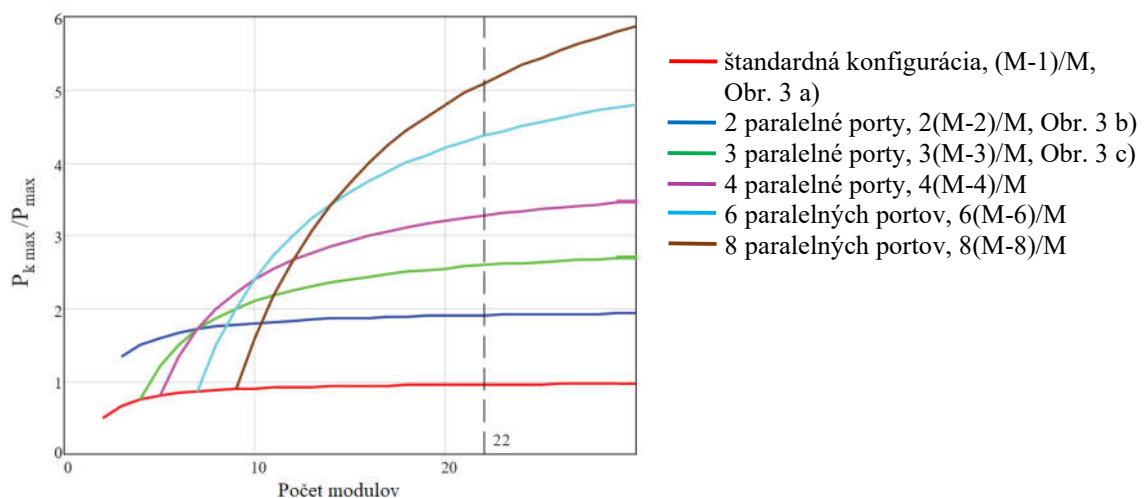
$$P_{k+1} = V_{ink+1} \cdot i_{k+1} = \frac{(M-2)}{M} \cdot \frac{V_0 \cdot T}{4 \cdot L_0} \quad (11)$$

Systém s veľkým počtom portov je potom možné prekonfigurovať tak, aby využíval menej zdrojov, resp. záťaží, ako je reálny počet portov. Maximálny výkon spracovaný zdrojom alebo záťažou tak možno výrazne zvýšiť. Pri popise MAB meničov uvažujeme, že porty meniča sú obojsmerné a teda zariadenie pripojené na port môže byť ako zdroj energie, tak aj spotrebič energie, prúd teda môže tiecť oboma smermi. Smer toku prúdu potom závisí na požadovaných tokoch energie v systéme. Na Obr. 3 sú znázornené rôzne konfigurácie na zmenu výkonu MAB meniča pomocou zmeny zapojenia na portoch meniča. Obr. 3 a) je štandardná konfigurácia, kde každý port môže mať pripojenú záťaž alebo zdroj energie. Na Obr. 3 b) sú dva porty zapojené paralelne, aby sa zvýšil výkon dodávaný/odberaný do/zo záťaže/zdroja, ktorý je k nim pripojený. Na Obr. 3 c) sú tri porty spojené dohromady a nakoniec topológia na Obr. 3 d) zodpovedá už vyššie spomínanému DAB meniču. Vo všetkých topológiách uvedených na Obr. 3 je možné energiu z každého portu odoberať a taktiež portu energiu dodávať. Dôležitá je potom už len samotná požiadavka na druh (dodávaná/odoberaná) a veľkosť energie na tom-ktorom porte, kde na základe takejto požiadavky port odoberá alebo dodáva energiu.



Obr. 3 Rôzne konfigurácie topológie MAB meniča pre zvýšenie výkonu pri nezmenenom počte portov

Na obr. 4 je znázornený maximálny možný výkon spravovaný skupinou paralelných portov ako funkcia počtu modulov. Vertikálna os zodpovedá výkonu spravovanému skupinou paralelných portov (modrá čiara tri porty, zelená čiara štyri porty atď.), ktoré sa správajú ako zdroj ($\varphi = 0$), zatiaľ čo ostatné porty sa správajú ako záťaž ($\varphi = 0,5$), v závislosti od celkového počtu modulov. Nárast výkonu je uvedený ako podiel výkonu prenášaného portom $k - P_{kmax}$ a maximálneho možného výkonu v jednom porte P_{max} (9). Obr. 4 porovnáva, o koľko sa zvýši výkon pri paralelnom zapojení portov. Ak sa napríklad použije veľmi vysoký počet modulov (napr. 30), pripojenie dvoch portov k jednej záťaži/zdroju zvýši výkon o 86 %. Pri strednom počte modulov (napr. 10) sa výkon zvýši o 60 %. Ak sú pripojené tri porty, výkon sa zvýši o 110 % pre 10 modulov. Paralelné porty teda poskytujú viac energie, avšak nejde o lineárnu závislosť. Pre výpočet percentuálneho zvýšenia výkonu v porovnaní so štandardnou konfiguráciou na jednotlivých portoch platia závislosti uvedené v legende k Obr. 4, kde zvýšenie výkonu závisí na počte paralelne prepojených portov a na celkovom počte modulov. Celkový počet modulov vo viacportovom meniči je označený ako „M“. Napríklad, pre spojenie dvoch paralelných portov je to výraz: *zvýšenie výkonu v %* = $2(M - 2)/M$.



Obr. 4 Maximálny možný výkon portu alebo skupiny paralelných portov v závislosti od ich počtu

III. ZÁVER

Tento príspevok predstavuje modulárny MAB menič s portami pre obojsmerný tok energie. Menič je zostavený z modulov, ktoré pozostávajú z plného mostíka a transformátora. Model transformátora je reprezentovaný ako rozptyľová indukčnosť zaradená v sérii a prevodový pomer transformátora. Článok predstavuje teoretickú analýzu a poskytuje rovnice pre návrh meniča s požadovaným výkonom každého portu. Prenos výkonov medzi jednotlivými portami je riadený pomocou fázových posunov ostatných mostíkov od referenčného mostíka s fázovým posunom $\varphi = 0$. Príspevok tiež popisuje možnosti úpravy topológie MAB pre zvýšenie výkonu prenášaného jedným portom.

Ďalší výskum bude zameraný na simulačné overenie matematického popisu MAB meniča a na overenie možnosti výpočtu fázového posunu v prípade, že užívateľ pozná žiadaný prúd konkrétneho mostíka.

POĎAKOVANIE

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23.

REFERENCIE

- [1] Haimin Tao; Duarte, J.L.; Hendrix, M.A.M., "Multiport converters for hybrid power sources," Power Electronics Specialists Conference, 2008. PESC 2008. IEEE , vol., no., pp.3412,3418, 15-19 June 2008 doi: 10.1109/PESC.2008.4592483
- [2] Xing Zhang; Shengyong Liu; Hongbin Chen; Xie Jun; Haibin Guo, "Multiple-port DC/DC converter for power flows management in electric vehicles," Power Electronics for Distributed Generation Systems (PEDG), 2010 2nd IEEE International Symposium on , vol., no., pp.727,730, 16-18 June 2010 doi: 10.1109/PEDG.2010.5545770.
- [3] Burnett, M.B.; Borle, L.J., "A power system combining batteries and supercapacitors in a solar/hydrogen hybrid electric vehicle," Vehicle Power and Propulsion, 2005 IEEE Conference , vol., no., pp.7 pp., 7-9 Sept. 2005 B. Smith, "An approach to graphs of linear forms (Unpublished work style)," unpublished.
- [4] Karanayil, B.; Arregui, M.G.; Agelidis, V.G.; Ciobotaru, M., "Bi- directional isolated multi-port power converter for aircraft HVDC network power transfer," IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society , vol., no., pp.3394,3399, 25-28 Oct. 2012 doi: 10.1109/IECON.2012.6389354.
- [5] Oliver, J.A.; Zumel, P.; Sanz, M.; Raga, C.; Izquierdo, D.; Garcia, O.; Barrado, A.; Prieto, R.; Azcona, R.; Delicado, B.; Cobos, J.A., "High level decision methodology for the selection of a fuel cell based power distribution architecture for an aircraft application," Energy Conversion Congress and Exposition, 2009. ECCE 2009. IEEE , vol., no., pp.459,464, 20-24 Sept. 2009.
- [6] Jianwu Zeng; Wei Qiao; Liyan Qu; Yanping Jiao, "An Isolated Multiport DC-DC Converter for Simultaneous Power Management of Multiple Different Renewable Energy Sources," Emerging and Selected Topics in Power Electronics, IEEE Journal of , vol.2, no.1, pp.70,78, March 2014 doi: 10.1109/JESTPE.2013.2293331
- [7] Jianwu Zeng; Wei Qiao; Liyan Qu, "A LCL-resonant isolated multiport DC-DC converter for power management of multiple renewable energy sources," Energy Conversion Congress and Exposition (ECCE), 2013 IEEE , vol., no., pp.2347,2354, 15-19 Sept. 2013 doi: 10.1109/ECCE.2013.6647001
- [8] Alonso, A.R.; Sebastian, J.; Lamar, D.G.; Hernando, M.M.; Vazquez, A., "An overall study of a Dual Active Bridge for bidirectional DC/DC conversion," Energy Conversion Congress and Exposition (ECCE), 2010 IEEE , vol., no., pp.1129,1135, 12-16 Sept. 2010
- [9] Reese, B.; Schupbach, M.; Lostetter, A.; Rowden, B.; Saunders, R.; Balda, J., "High voltage, high power density bi-directional multi-level converters utilizing silicon and silicon carbide (SiC) switches", Applied Power Electronics Conference and Exposition, 2008. APEC 2008
- [10] Siri, K.; Willhoff, M.; Haibing Hu; Batarseh, I., "High-voltage-input, low-voltage-output, series-connected converters with uniform voltage distribution", Energy Conversion Congress and Exposition, 2009. ECCE. IEEE 20-24 Sept. 2009 Pp.:541 – 547
- [11] Cilio, E.; Mitchell, G.; Schupbach, M.; Lostetter, A., "SiC intelligent multi module DC/DC converter system for space applications", IEEE Aerospace conference, 2009
- [12] Falcones, S.; Ayyanar, R.; Xiaolin Mao, "A DC-DC Multiport- Converter-Based Solid-State Transformer Integrating Distributed Generation and Storage," Power Electronics, IEEE Transactions on , vol.28, no.5, pp.2192,2203, May 2013
- [13] Haimin Tao; Kotsopoulos, A.; Duarte, J.L.; Hendrix, M.A.M., "Transformer-Coupled Multiport ZVS Bidirectional DC-DC Converter With Wide Input Range," Power Electronics, IEEE Transactions on , vol.23, no.2, pp.771,781, March 2008 doi: 10.1109/TPEL.2007.915129.
- [14] Ari, A.M.; Lingxi Li; Wasynczuk, O., "Modeling and analysis of N-port DC-DC converters," Power and Energy Conference at Illinois (PECI), 2012 IEEE , vol., no., pp.1,7, 24-25 Feb. 2012
- [15] Ari, A.M.; Lingxi Li; Wasynczuk, O., "Modeling and analysis of N-port DC-DC converters using the cyclic average current," Applied Power Electronics Conference and Exposition (APEC), 2012 Twenty-Seventh Annual IEEE , vol., no., pp.863,869, 5-9 Feb. 2012

Batériové systémy elektromobilov a predstavenie modulárneho systému

¹ Daniel MARCIN, ² Milan LACKO

^{1,2}Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹daniel.marcin@tuke.sk, ²milan.lacko@tuke.sk

Abstrakt — Rastúci dopyt po efektívnych a cenovo dostupných batériových systémoch podnietil výskum v oblasti elektrických vozidiel. V oblasti elektromobility sú úložiská energie najväčšou oblasťou výskumu. Riešenia, ktoré by vedeli doceliť zvýšenie dojazdu, zníženie celkovej ceny alebo zlepšenie v oblasti údržby či výmeny batérie sú nevyhnutné pre budúci vývoj týchto vozidiel. Modularita a jednosť v oblasti batérií má veľký potenciál splniť mnohé z týchto očakávaní.

Kľúčové slová — batéria, systém manažmentu batérie, balansovanie článkov, elektromobil, Li-ion, modulárny batériový systém

I. ÚVOD

V posledných rokoch vzrástla popularita elektromobilov pretože predstavujú riešenie na zníženie závislosti od fosílnych palív a obmedzenie environmentálneho vplyvu dopravy. Účinnosť elektromobilov je výrazne ovplyvnená účinnosťou a spoľahlivosťou ich batériového systému. Ten je v elektromobiloch kľúčovým faktorom pri určovaní dojazdu, výkonu a celkovej spoľahlivosti vozidla. S narastajúcim dopytom po elektromobiloch rastie aj potreba pokročilých batériových riešení, ktoré sú efektívne a hlavne cenovo dostupné [1], [2].

Modulárny batériový systém, pozostávajúci z viacerých vymeniteľných batériových modulov, prináša veľký potenciál pre splnenie energetických požiadaviek elektrických vozidiel. Tento systém poskytuje pružnosť prispôbovať sa meniacim požiadavkám, či už ide o prispôbenie používateľským preferenciám alebo jazdným podmienkam, ktoré vyžadujú rôzne dojazdy na jedno nabitie. Ďalším typom modulu môžu byť superkapacity, ktoré môžu pomôcť batérii pri rekuperatívnom brzdení. Okrem toho možnosť pridávať alebo odstraňovať moduly zlepšuje údržbu batériových systémov. Tieto modulárne systémy majú potenciál splniť výkonové a cenové požiadavky elektromobilov a zároveň poskytujú pružnosť prispôbiť sa meniacim sa potrebám do budúcnosti.

Systém manažmentu batérie (BMS) je neoddeliteľnou súčasťou každého batériového systému elektromobilov. Zabezpečuje bezpečnú prevádzku a optimálny výkon. BMS monitoruje stav jednotlivých článkov a vykonáva funkcie, ako napríklad monitorovanie napätia a teploty, balansovanie článkov, detekciu porúch alebo ochranu. Balansovanie článkov je kľúčovým aspektom BMS. Zabezpečuje rovnomerné rozdelenie energie pri nabíjaní alebo vybíjaní medzi článkami batériového packu. Tým sa zabráni preťaženiu jednotlivých článkov, čo môže viesť k zníženému výkonu, zvýšenému opotrebeniu a potenciálnym bezpečnostným rizikám. V súčasnosti je pasívne balansovanie článkov široko využívanou technikou v elektromobiloch [1]. Alternatívnym riešením je aktívne balansovanie, ktoré je účinnejšie a zahŕňa prenos energie medzi článkami, čím sa zabráni plytvaniu energiou [2].

Tento článok sa v úvode zaoberá trendami a lítium-iónovými batériami s potrebnými obvody pre ich manažment. Nasleduje prehľad v metódach balansovania článkov. Spomenuté sú metódy pasívneho balansovania. Väčší dôraz je kladený na metódy aktívneho balansovania a rôznym obvodom a topológiám. Na záver je predstavený modulárny batériový systém pre elektromobily.

II. BATÉRIOVÝ SYSTÉM V ELEKTRICKÝCH VOZIDLÁCH

Jednou z kľúčových oblastí výskumu pri elektromobiloch je zvýšenie dojazdu batériových systémov. To zahŕňa vývoj nových technológií, ktoré dokážu uchovať viac energie v menšom a ľahšom prevedení. Ďalším trendom v oblasti batérií pre elektromobily je narastajúci dopyt po efektívnych systémoch manažmentu batérií. Pasívne balansovanie článkov, ktoré premieňa

energiu na teplo, sa dlhodobo uplatňuje ako štandardná metóda. Avšak, keďže batériové systémy sa stávajú väčšími a komplexnejšími, aktívne balansovanie článkov sa stáva populárnejšou voľbou.

A. *Lítium-iónové (Li-ion) batérie*

Li-ion batérie sa stali kľúčovou súčasťou rastu elektromobilov. Vyznačujú sa vysokou energetickou hustotou, dlhou životnosťou, zanedbateľným pamäťovým efektom a nízkym samo vybíjaním [1].

Li-ion batérie pre elektrické vozidlá sa vyrábajú vo forme cylindrických článkov, hranolových článkov alebo vreckových článkov. Cylindrické články sú najbežnejšie používaným typom (Obr. 1). Hranolové články majú plochý a obdĺžnikový tvar, zatiaľ čo vreckové články (Obr. 1) sú flexibilné a môžu byť tvarované do rôznych tvarov [2], [2], [4].



Obr. 1 Batériové články, batériový modul, batériový pack [5]

Rozloženie batériových článkov v batériovom balíku je prispôbené podľa požadovaného napätia a kapacity batérie elektrického vozidla. Články sú zapojené do série, aby bolo dosiahnuté požadované napätie, a keď je potrebná vyššia kapacita, pridávajú sa ďalšie články zapojené paralelne. Táto konfigurácia, ktorá je znázornená na Obr. 1, umožňuje vytvorenie batériového modulu z viacerých zapojených článkov a spojením viacerých batériových modulov vznikne konečný batériový pack [1].

B. *Systém manažmentu batérie*

Napriek tomu, že li-ion batérie majú široké spektrum aplikácií a mnoho výhod, môžu byť značne nebezpečné, ak nie sú prevádzkované v rámci svojho bezpečného prevádzkového rozsahu (SOA) [6]. Obzvlášť li-ion batérie používané v elektromobiloch vyžadujú vhodné bezpečnostné opatrenia, najmä kvôli zvýšenému riziku vplyvu meniacich sa jazdných podmienok alebo teploty prostredia. SOA je stanovená prevádzkovým rozsahom napätia, prúdu a teploty [7]. Pre li-ion články je bezpečné prevádzkové napätie v rozsahu od 2,4 V do 4,2 V [8]. Teplota je bezpečná v rozmedzí od -20 °C do 60 °C. Nabíjací a vybíjací prúd je vyjadrený v pomere ku kapacite článku, tzv. C-rate [9].

BMS je kľúčovou súčasťou elektromobilov a zodpovedá za zabezpečenie bezpečnej a efektívnej prevádzky batériového packu. Kľúčové funkcie BMS sú nasledovné [1]:

- monitorovanie stavu batériových článkov,
- monitorovanie vstupno-výstupného prúdu a napätia,
- riadenie nabíjania a vybíjania,
- odhad stavu (SOC, SOH),
- ochrana batérie,
- balansovanie článkov,
- riadenie prevádzkových teplôt a manažment tepla,
- komunikácia s okolitými systémami,
- ukladanie a získavanie dát,
- diagnostika a hodnotenie porúch.

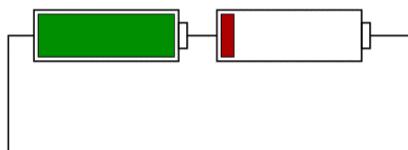
Tieto funkcie sú nevyhnutné pre zabezpečenie spoľahlivosti batérií, zvýšenie vybíjacieho prúdu, predĺženie dojazdu, maximalizáciu výkonu a predĺženie celkového životného cyklu batériového packu [10], [11].

Presné odhadnutie stavu batérie je kľúčovým aspektom prevádzky BMS. Ide o odhad stavu nabitia (SOC) a stavu zdravia (SOH). SOC batérie indikuje stav nabitia, ktoré zostáva v batérii. SOH batérie sa vzťahuje na kapacitu batérie v porovnaní s jej pôvodnou kapacitou, keď bola nová [12].

C. *Balansovanie článkov*

Balansovanie li-ion článkov je dôležitým procesom v rámci riadenia batériových systémov. Li-ion články majú tendenciu mať odlišné stavy nabitia, čo môže viesť k nerovnováhe v batériovom packu. Nerovnováha môže spôsobiť skracovanie životnosti batérií, zníženie výkonu a dokonca znemožniť používanie batérie, čo je znázornené na Obr. 2. Bez balansovania sa tento batériový

pack stáva nevyužitelným, pretože jeden článok by bol počas nabíjania nadmerne nabitý, zatiaľ čo druhý článok by bol počas vybijania nadmerne vybitý [14].



Obr. 2 Batériový pack s článkami, ktoré sú nevybalansované [15]

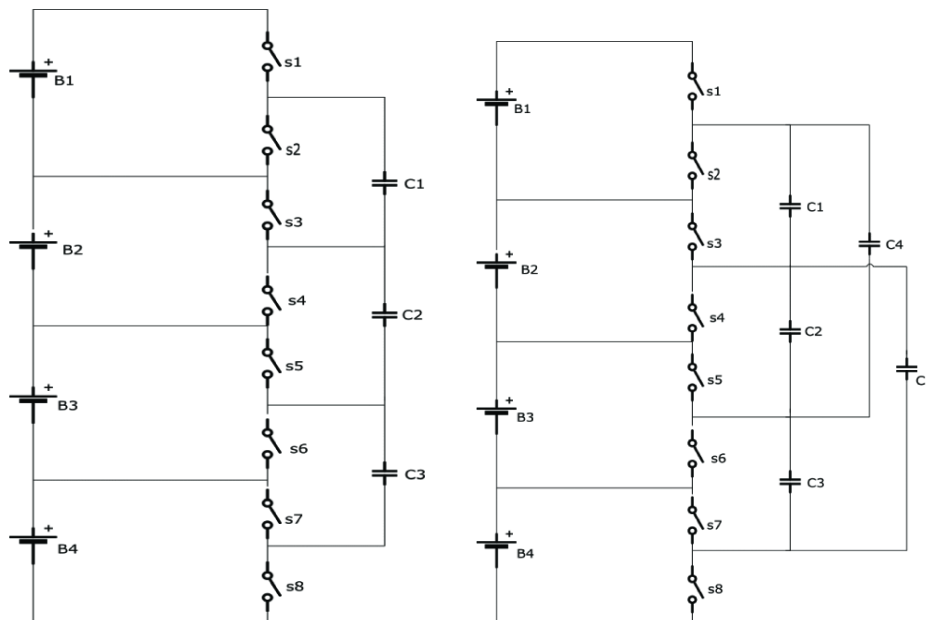
Cieľom balansovania li-ion článkov je dosiahnuť rovnomerné rozdelenie nabitia medzi všetkými článkami batériového packu. Existujú rôzne metódy balansovania, vrátane pasívneho balansovania a aktívneho balansovania.

Pri pasívnom balansovaní sa využíva premena energie na teplo pomocou rezistorov. Pasívne balansovanie je najjednoduchšou a najlacnejšou metódou balansovania. Články s vyšším SOC sú vybijané pomocou paralelného odporu, ktorý je riadený spínačom [17], [16]. Ďalšou formou pasívneho balansovania je použitie pevného odporu bez dodatočného spínača. Pre všetky články majú odpory rovnakú hodnotu. Balansovanie článkov môže byť teda dosiahnuté bez potreby aktívneho riadenia.

Na druhej strane, aktívne balansovanie vyžaduje použitie dodatočných obvodov a elektroniky na presné riadenie rozdeľovania nabitia medzi článkami. Táto metóda umožňuje efektívnejšie a presnejšie vyrovnanie nabitia článkov. Hlavnou myšlienkou aktívneho balansovania je využitie komponentov, ako sú kondenzátory, cievky alebo transformátory, na ukladanie energie z článkov s vysokým SOC a prenos tejto energie do článkov s nízkym SOC. Toto umožňuje presnejšiu kontrolu procesu balansovania a znižuje straty energie, čo vedie k vyššej energetickej účinnosti. Avšak aktívne balansovanie článkov je zvyčajne zložitejšie a drahšie v porovnaní s pasívnym balansovaním článkov a nemusí byť najlepšou voľbou pre všetky aplikácie [17].

1) Aktívne balansovanie na báze kondenzátorov

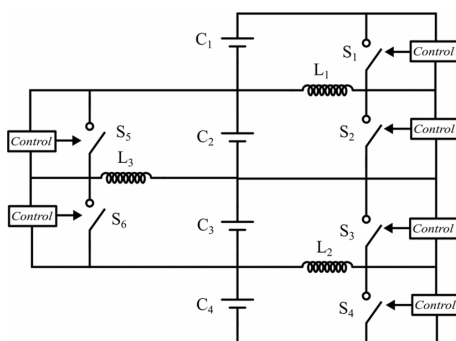
Zo všetkých možností úložiska energie pri aktívnom balansovaní článkov je kondenzátor najviac preferovaný. Nemá magnetické komponenty, je malých rozmerov a cenovo výhodný. Metóda spínaných kondenzátorov (Obr. 3) je jednou z najbežnejších techník používaných pri tejto metóde aktívneho balansovania článkov [17].



Obr. 3 Kondenzátorové balansovanie: (vľavo) Jednoúrovňový (Single tiered) spínaný kondenzátor (vpravo) Dvojurovňový (Double tiered) spínaný kondenzátor [14]

2) Aktívne balansovanie na báze cievok

Balansovanie článkov založené na cievkach je ďalšou bežnou metódou používanou pri balansovaní článkov. Podobne ako balansovanie s kondenzátormi, aj tu sa využívajú spínače a podobné topológie obvodov s jednou alebo viacerými cievkami (Obr. 4). Táto metóda je známa svojím rýchlym časom balansovania v porovnaní s balansovaním pomocou kondenzátorov [19], [20]. Avšak cievky sú drahšie ako kondenzátory.



Obr. 4 Aktívne balansovanie so spínanými cievkami [20]

3) Aktívne balansovanie na báze transformátora

Ďalšou metódou prenosu energie pri aktívnom balansovaní článkov je použitie transformátorov. Magnetické pole vytvorené transformátorom sa využíva na prenos energie z článkov s vyšším SOC na články s nižším SOC. Jednoduchý transformátor s primárnym a sekundárnym vinutím sa môže použiť vo forme jedného blokujúceho meniča [14], [21]. Inou možnosťou prenosu energie je použitie transformátorov s viacerými vinutiami [21]. Primárne vinutie transformátora je zvyčajne pripojené na stranu batériového balansovacieho obvodu. Sekundárne vinutie je pripojené ku každému článku v balansovanom obvode [23].

III. MODULÁRNY BATÉRIOVÝ SYSTÉM

V poslednej dobe sa zvyšuje dopyt po energeticky úsporných a cenovo dostupných batériových systémoch. To vedie výskumníkov k vývoju nových technológií v oblasti batériových systémov pre elektromobily. V rámci toho sa vyvíjajú aj modulárne batériové systémy, ktoré sa skladajú z vymeniteľných batériových modulov. Táto nová technológia má veľký potenciál ovplyvniť dojazd, modularitu, údržbu a recykláciu batérií elektrických vozidiel.

Spoločnosť Webasto nedávno uviedla na trh inovatívny "plug-and-play" batériový systém, ktorý ponúka 35 kWh energie na jednu batériovú súpravu. Tento systém si kladie za cieľ poskytnúť pre výrobcov cenovo efektívne riešenie elimináciou potreby samostatného batériového systému pre každé vozidlo. Batériové súpravy môžu byť pripojené paralelne pre dosiahnutie 400 V s piatimi modulmi, alebo sériovo pre dosiahnutie 800 V s desiatimi modulmi [24].

Jednou z hlavných výhod modulárnych batériových systémov je, že umožňujú jednoduchšiu výmenu jednotlivých batériových modulov, namiesto výmeny celého batériového packu. Ďalšou výhodou modulárnych batériových systémov je ich jednoduchá aktualizácia s príchodom nových batériových technológií. To umožňuje majiteľom elektrických vozidiel predĺžiť životnosť svojich vozidiel jednoduchou výmenou jednotlivých modulov, namiesto nutnosti zakúpiť úplne novú batériu.

Tieto systémy sa skladajú z vymeniteľných batériových modulov, ktoré je možné podľa potreby meniť a prispôbovať kapacitu batérie vozidla. Toto je obzvlášť užitočné pre väčšie vozidlá, ako sú napríklad autobusy. Keďže môžu niesť len toľko modulov batérie, koľko je potrebné pre konkrétnu jazdu, vedú si značne znížiť hmotnosť, a tým pádom zvýšiť efektívnosť využitia energie z batérie.

ZÁVER

Článok sa zaoberal prehľadom batériových systémov pre elektromobily. Hlavnou témou je aktívne balansovanie, ktoré by malo byť základom pre vývoj v oblasti modulárneho batériového systému. Prehľad rôznych balansovacích metód prináša dôležité aspekty pre výber metódy aktívneho balansovania pri návrhu a zostrojení modulárneho batériového systému.

POĎAKOVANIE

Tento príspevok vznikol s podporou Agentúry pre vedu a výskum Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky v rámci projektu APVV-18-0436.

IV. REFERENCIE

- [1] M. A. Hannan, M. M. Hoque, A. Hussain, Y. Yusof and P. J. Ker, "State-of-the-Art and Energy Management System of Lithium-Ion Batteries in Electric Vehicle Applications: Issues and Recommendations," in IEEE Access, vol. 6, pp. 19362-19378, 2018, doi: 10.1109/ACCESS.2018.2817655.
- [2] Halimah, Putri Nur, Samuel Rahardian, and Bentang Arief Budiman. "Battery cells for electric vehicles." International Journal of Sustainable Transportation Technology 2.2 (2019): 54-57.
- [3] Sun, Peiyi & Bisschop, Roeland & Niu, Huichang & Huang, Xinyan. (2020). A Review of Battery Fires in Electri

- c Vehicles. *Fire Technology*. 1-50. 10.1007/s10694-019-00944-3.
- [4] What Types of Batteries Are Used in Electric Cars? [Online]. Available: <https://dragonflyenergy.com/electric-car-batteries/>
- [5] Supply Chain Opportunities from Electrification of Vehicles. [Online]. Available: <https://northernautoalliance.com/wp-content/uploads/2019/05/Ready-for-ACES-Supply-Chain-Opportunities-from-Electrification-of-Vehicles.pdf>
- [6] What is a Battery Management System? [Online]. Available: <https://www.synopsys.com/glossary/what-is-a-battery-management-system.html>
- [7] Andrea, Davide. *Battery management systems for large lithium-ion battery packs*. Artech house, 2010.
- [8] N. Ghaeminezhad, Q. Ouyang, X. Hu, G. Xu and Z. Wang, "Active Cell Equalization Topologies Analysis for Battery Packs: A Systematic Review," in *IEEE Transactions on Power Electronics*, vol. 36, no. 8, pp. 9119-9135, Aug. 2021, doi: 10.1109/TPEL.2021.3052163.
- [9] Battery-Management-Systems. [Online]. Available: <https://www.vde.com/resource/blob/2105544/731ed20695f9fbf2f66423407554d1c5/fact-sheet-battery-management-system-pdf-file-data.pdf>
- [10] Yang, Shichun, et al. *Advanced Battery Management System for Electric Vehicles*. Springer Nature, 2022.
- [11] S. B and P. Pradeepa, "Review on Battery Management System in EV," 2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSPP), Hyderabad, India, 2022, pp. 1-4, doi: 10.1109/ICICCSPP53532.2022.9862367.
- [12] P. Shen, M. Ouyang, L. Lu, J. Li and X. Feng, "The Co-estimation of State of Charge, State of Health, and State of Function for Lithium-Ion Batteries in Electric Vehicles," in *IEEE Transactions on Vehicular Technology*, vol. 67, no. 1, pp. 92-103, Jan. 2018, doi: 10.1109/TVT.2017.2751613.
- [13] L. Valda and K. Kosturik, "Comparison of Li-ion active cell balancing methods replacing passive cell balancer," 2015 International Conference on Applied Electronics (AE), Pilsen, Czech Republic, 2015, pp. 267-270.
- [14] B. Jiang, Y. Liu, X. Huang and R. R. R. Prakash, "A New Battery Active Balancing Method with Supercapacitor Considering Regeneration Process," *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, Singapore, 2020, pp. 2364-2369, doi: 10.1109/IECON43393.2020.9254839.
- [15] Plett, Gregory L. *Battery management systems, Volume II: Equivalent-circuit methods*. Artech House, 2015.
- [16] Liu, Wei & Placke, Tobias & Chau, K.T.. (2021). *Overview of Batteries and Battery Management for Electric Vehicles*.
- [17] X. Wei and B. Zhu, "The research of vehicle power Li-ion battery pack balancing method," 2009 9th International Conference on Electronic Measurement & Instruments, Beijing, China, 2009, pp. 2-498-2-502, doi: 10.1109/ICEMI.2009.5274520.
- [18] R. Paul and S. S., "Electric Vehicle Cell Balancing using Single and Multi tiered Switched Capacitor," 2022 4th International Conference on Energy, Power and Environment (ICEPE), Shillong, India, 2022, pp. 1-6, doi: 10.1109/ICEPE55035.2022.9798231.
- [19] A. F. Moghaddam and A. Van Den Bossche, "An Active Cell Equalization Technique for Lithium Ion Batteries Based on Inductor Balancing," 2018 9th International Conference on Mechanical and Aerospace Engineering (ICMAE), Budapest, Hungary, 2018, pp. 274-278, doi: 10.1109/ICMAE.2018.8467685.
- [20] T. H. Phung, J. -C. Crebier, A. Chureau, A. Collet and V. Nguyen, "Optimized structure for next-to-next balancing of series-connected lithium-ion cells," 2011 Twenty-Sixth Annual IEEE Applied Power Electronics Conference and Exposition (APEC), Fort Worth, TX, USA, 2011, pp. 1374-1381, doi: 10.1109/APEC.2011.5744771.
- [21] Cui, Xiudong & Shen, Weixiang & Zhang, Yunlei & Hu, Cungang. (2017). A Novel Active Online State of Charge Based Balancing Approach for Lithium-Ion Battery Packs during Fast Charging Process in Electric Vehicles. *Energies*. 10. 1766. 10.3390/en10111766.
- [22] Y. Shang, B. Xia, C. Zhang, N. Cui, J. Yang and C. Mi, "A Modularization Method for Battery Equalizers Using Multiwinding Transformers," in *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, pp. 8710-8722, Oct. 2017, doi: 10.1109/TVT.2017.2702065.
- [23] C. Bonfiglio and W. Roessler, "A cost optimized battery management system with active cell balancing for lithium ion battery stacks," 2009 IEEE Vehicle Power and Propulsion Conference, Dearborn, MI, USA, 2009, pp. 304-309, doi: 10.1109/VPPC.2009.5289837.
- [24] A 'plug-and-play' electric vehicle battery solution? Inside Webasto's new modular CV battery system. [Online]. Available: <https://www.fleetequipmentmag.com/webasto-modular-cv-battery-system/>

Synchronizácia 3-motorového pohonu

¹Lukáš PANCURÁK, ²Karol KYSLAN

^{1,2}Department of Electrical Engineering and Mechatronics, Faculty of Electrical Engineering and Informatics Technical University of Košice, Slovak Republic

¹lukas.pancurak@tuke.sk, ²karol.kyslan@tuke.sk

Abstrakt – Tento článok sa zaoberá problematikou synchronizácie a koordinovaného riadenia motorov prostredníctvom priemyselných frekvenčných meničov. Práca bola realizovaná s meničmi SINAMICS a riadiacou jednotkou SIMOTION od firmy SIEMENS, ktorými je riadený troj-motorový pohon určený pre demonštráciu polohovej synchronizácie. Riadiaci algoritmus v jazyku MCC a ovládanie systému je navrhnuté v prostredí konfiguračnej aplikácie SIMOTION SCOUT. Správna činnosť systému bola experimentálne overená a zdokumentovaná v tejto práci.

Keywords – Polohové riadenie, Rýchlostné riadenie, Synchronizácia motorov, Servopohon, SIMOTION

I. ÚVOD

Problém synchronizovaného riadenia je aktuálny z praktického aj teoretického hľadiska. V praxi je synchronizácia osí motorov často riešeným problémom. Najmä v oblasti riadenia servomotorov, robotov a robotických ramien sa vyžaduje veľmi presné polohové a rýchlostne riadenie niekoľkých, súbežne pracujúcich synchronizovaných pohonov. Mnoho aplikácií vyžaduje synchronizáciu dvoch alebo viacerých motorových osí, vrátane portálových zariadení, papierenských a tlačiarenských liniek, dopravníkových systémov, robotov a množstva ďalších aplikácií. Z hľadiska výskumu sa hľadajú nové metódy synchronizácie, ktoré sú efektívnejšie, presnejšie a odolnejšie voči rušivým vplyvom a nerovnomernému zaťaženiu.

Priemyselné aplikácie s viacerými motormi a osami si často vyžadujú nejakú formu synchronizácie. Takéto systémy často využívajú pokročilé priemyselné meniče, s riadiacou jednotkou, ktorej vlastnosti uľahčujú proces synchronizácie. Priemyselné meniče sú súčasťou takmer každej väčšej výrobnéj prevádzky, kde je potrebná regulácia pohonov. Moderné priemyselné meniče nachádzajú využitie aj mimo priemyselných parkov, a to napríklad v cestnej a železničnej doprave alebo v rozsiahlych automatizovaných systémoch v mestách.

Cieľom tohto článku je navrhnúť riadenie dvoch vedľa seba umiestnených servomotorov, vybavených ozubenými kolesami na hriadelí. Riadením sa musí dosiahnuť aby bolo možné ozubené kolesá na hriadeloch prepojiť v ľubovoľnom okamihu, a pri ľubovoľných otáčkach, či smere otáčania. Pre bezpečné a pohodlné spájanie a rozpájanie ozubených kolies bude nutné zabezpečiť vhodnú konštrukciu pre osadenie a posun motorov.

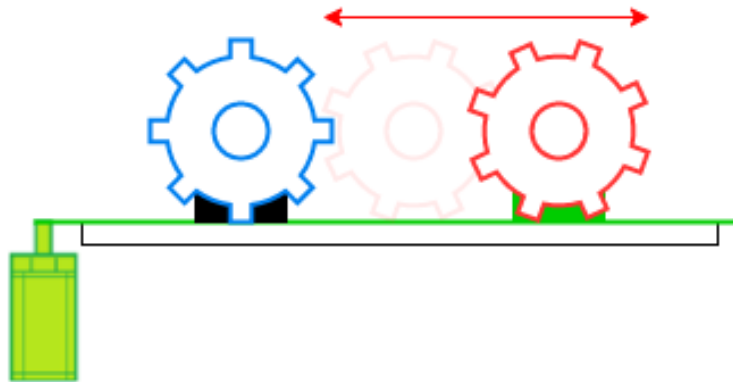
II. METÓDY SYNCHRONIZÁCIE

Metódy synchronizácie sa líšia podľa výstupnej veličiny alebo viacerých veličín motorov, ktoré chceme synchronizovať. Najčastejšie je potrebné synchronizovať motory podľa nasledujúcich veličín:

- Momentová synchronizácia,
- Rýchlostná synchronizácia,
- Polohová synchronizácia.

Momentová synchronizácia je v priemysle často využívaná najmä pre motory, ktoré sú mechanicky a momentovo prepojené. Avšak, motory v tomto systéme budú osadené iba ozubenými kolesami na hriadelí, tak ako je to znázornené na obrázku 1. Po prepojení sa bude jednať o bezmomentové spojenie. Závažový moment motorov bude preto minimálny a nebude dôležitým parametrom pri riadení a synchronizácii. Preto sa v nasledujúcej časti nebudeme zaoberať spôsobmi momentovej synchronizácie motorov. Je dobré poznamenať, že

metódy synchronizácie momentu sú zväčša len rozšírením alebo doplnením nižšie spomínaných metód. Viac o momentovej synchronizácii je možné nájsť v [1].



Obr. 1: Názorný náčrt riešenia

Metódy pre rýchlostnú a polohovú synchronizáciu sú v zásade veľmi podobné. Niektoré z metód sú vhodnejšie pre rýchlostnú synchronizáciu, iné zase pre polohovú. Pri polohovej synchronizácii sú riadiace štruktúry doplnené o regulátor polohy. Bolo vyvinutých niekoľko prístupov k efektívnej synchronizácii, ako napríklad:

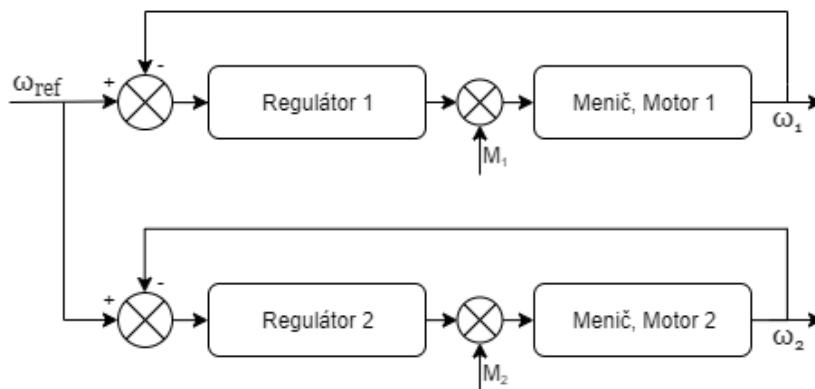
- Paralelná štruktúra,
- Master - Slave,
- Krížová väzba [2] [3],
- Virtuálna os [4].

Pri experimentálnej verifikácii bude synchronizácia realizovaná pomocou paralelnej štruktúry a master - slave štruktúry.

A. Paralelná štruktúra

V paralelnej riadiacej štruktúre, zdieľajú všetky motory rovnaký riadiaci vstup. Výstupné veličiny motorov budú rovnaké, za predpokladu, že parametre meničov a motorov sú rovnaké, a na systém nevyplýva žiadne rušenie. Za daných podmienok sú stabilita a dynamické vlastnosti synchronizovaných motorov výborné. Štruktúra a ovládanie takéhoto systému je jednoduché. Konkrétna ukážka paralelnej štruktúry pre 2 motory riadené v rýchlostnej regulačnej slučke, je znázornená na obrázku 2.

Zjavnou nevýhodou tejto štruktúry sú rušenia a iné vonkajšie vplyvy. Ak je rušenie v jednej z vetiev riadiacej štruktúry iné ako v ostatných, alebo sú motory nerovnomerne zaťažované, dôjde k zmene rýchlosti, polohy alebo momentu motora. Keďže regulátory motorov nie sú medzi sebou prepojené, nedôjde k vzájomnej regulácii a systém nebude pracovať synchronizovane. [5].



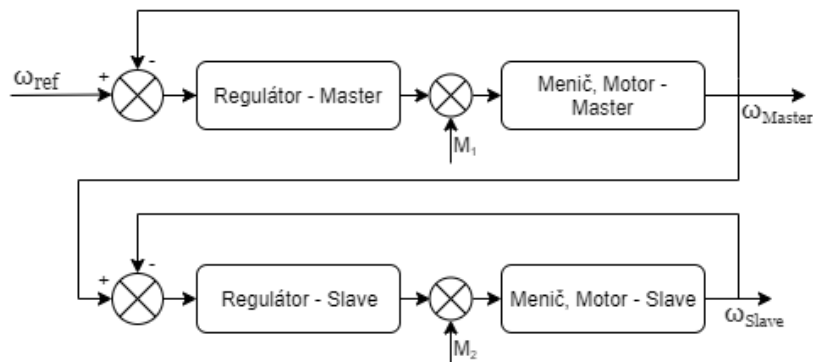
Obr. 2: Paralelná štruktúra

B. Master - Slave architektúra

Táto metóda bola po prvýkrát predstavená v [6]. Pri využití Master - Slave architektúry sú výstupné veličiny master motora použité ako referenčné hodnoty pre regulátory slave

motorov. Master motoru sú želané hodnoty zadané ľubovoľným spôsobom. Akékoľvek zmeny alebo rušenia na master motore sa prejavajú aj na slave motore. Avšak, zmeny a rušenie na slave motoroch nemajú vplyv na master motor, ani na ostatné slave motory, čo môže viesť k strate synchronizácie medzi motormi. Konkrétna ukážka master-slave architektúry pre 2 motory riadené v rýchlostnej regulačnej slučke, je znázornená na 3.

V priemyselných aplikáciách sa táto riadiaca štruktúra odporúča len v prípade, že sa pri prevádzke nebudú vyskytovať momentové rázy, pretože sa pri nich nedá s dostatočnou istotou zabezpečiť rýchlostná a polohová synchronizácia všetkých motorov.



Obr. 3: Master Slave architektúra s dvoma motormi

III. HARDVÉROVÁ ČASŤ NÁVRHU

Táto kapitola sa zaoberá technickým vybavením použitým pri riešení synchronizácie troj-motorového pohonu. Pre vytvorenie funkčného mechatronického systému budú využité nasledujúce zariadenia, rozdelené podľa miesta inštalácie:

- **Rozvodná skriňa**

- 1) Riadiaci systém SIMOTION D425-2 s integrovaným meničom SINAMICS S120,
- 2) Menič na strane siete - Smart-Line Module,
- 3) Menič na strane motora - Double Motor Module,
- 4) Menič na strane motora - Single Motor Module,
- 5) Dve snímačové moduly SMC 10.

- **Konštrukcia**

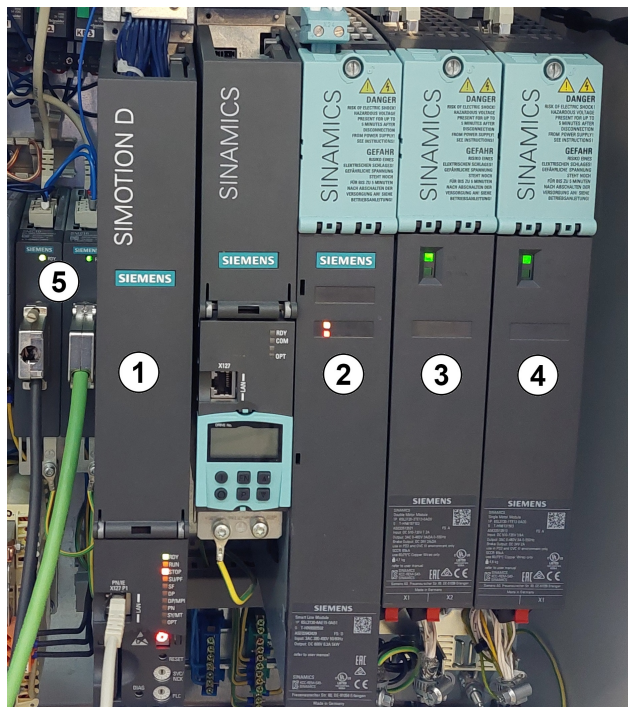
- 1) Master servomotor,
- 2) slave servomotor,
- 3) motor pre pohon posuvného vedenia,
- 4) indukčné snímače polohy.

A. Výkonové a riadiace zariadenia

Procesor D425-2 zabezpečuje chod Simotion Runtime System softvéru a Sinamics S120 Drive Control softvéru. Simotion Runtime System zabezpečuje komunikáciu, PLC a aritmetické funkcionality, riadenie polohy a koordináciu osí, užívateľské programy, a iné. Sinamics S120 Drive Control zabezpečuje vhodné riadenie meničov, čím dosahuje riadenie prúdu, otáčok, momentu a veličín medziobvodu v uzavretej slučke.

SINAMICS S120 je modulárny výkonový menič, určený pre aplikácie, ktoré vyžadujú precízne a vysoko dynamické riadenie. Vďaka modulárnemu prevedeniu, je možné menič prispôsobiť a implementovať na širokú škálu aplikácií. Systémy SINAMICS poskytujú viacero metód riadenia, ako napríklad skalárne a vektorové riadenie, ale aj tzv. servo riadenie, určené pre dynamicky náročné aplikácie. Menič disponuje usmerňovacím modulom Smart-Line Module a dvoma strieďačovými modulmi Double Motor Module a Single Motor Module.

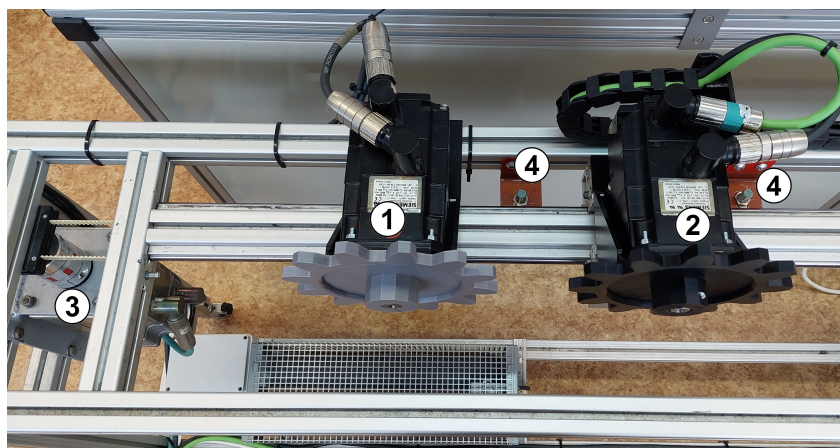
Moduly meniča sú inštalované v rozvodnej skrini, ktorá sa nachádza na obrázku 4. Číslovanie jednotlivých modulov zodpovedá číslovaniu v zozname na začiatku tejto kapitoly.



Obr. 4: Meniče SINAMICS s riadiacim systémom SIMOTION

B. Mechanické zariadenia

Pre demonštráciu polohovej a rýchlostnej synchronizácie osí motorov, budú využité dva identické synchronne motory s permanentnými magnetmi. Motory budú uchytené na konštrukcii s vlastným posuvným vedením. Jeden z motorov bude na konštrukcii pevne uchytený a druhý bude pripevnený na jazdcovi posuvného vedenia. Tretí, odlišný SMPM pre posuv jazdca bude umiestnený vertikálne, priamo pod posuvným vedením, ktoré bude poháňať. Zastavenie jazdca na požadovanej polohe budú zabezpečovať indukčné snímače osadené na konštrukcii. Konštrukcia sa nachádza na obrázku 5. Číslovanie jednotlivých zariadení zodpovedá číslovaní v zozname na začiatku tejto kapitoly.



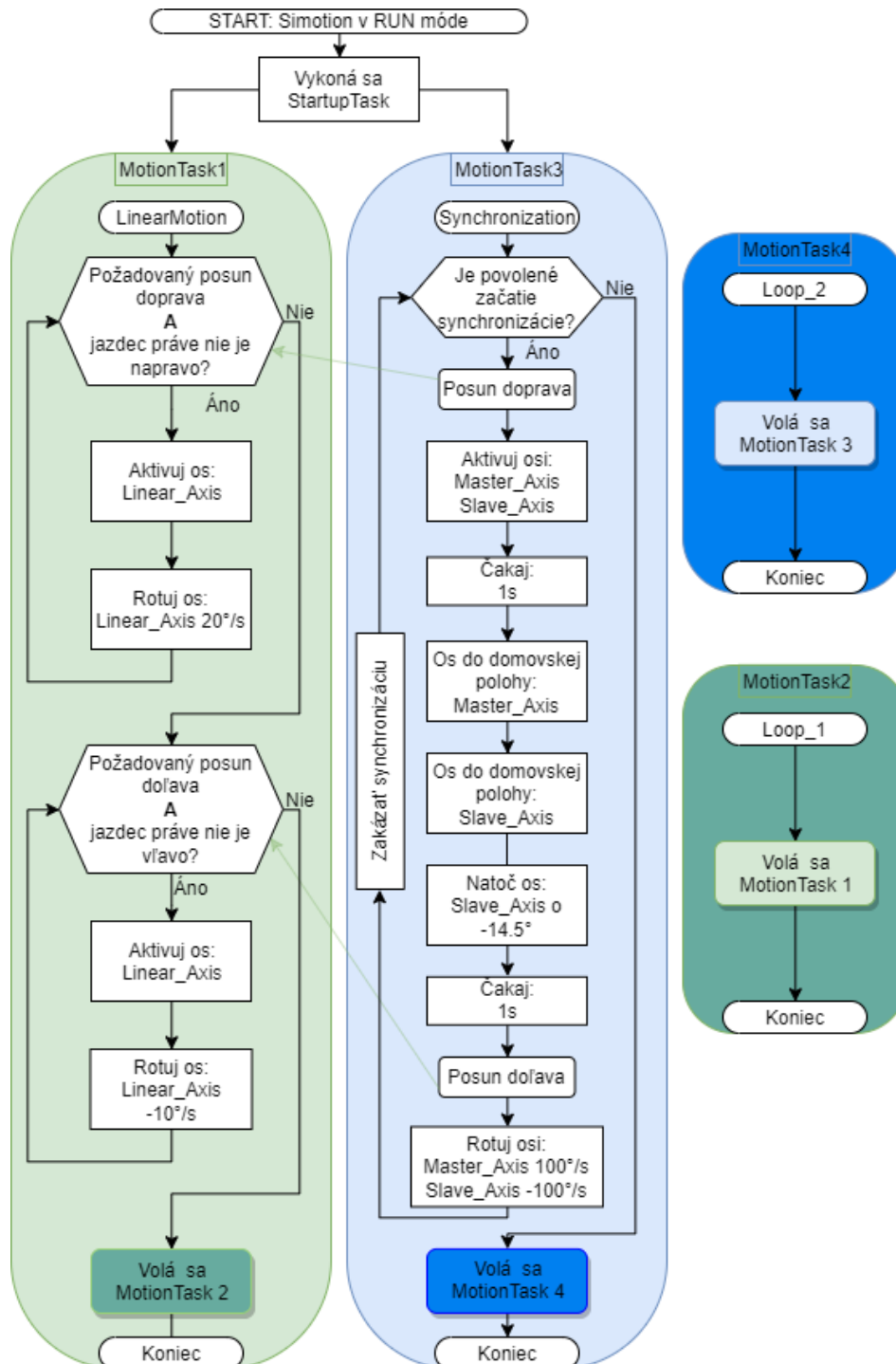
Obr. 5: Konštrukcia s osadenými motormi

IV. PROGRAMOVÁ ČASŤ NÁVRHU

Pri návrhu riadiaceho algoritmu bolo využité softvérová platforma pre parametrizovanie Simotion SCOUT. Algoritmus bol navrhnutý v programovacom jazyku *MotionControlChart* (MCC), s pomocou softvérového technologického objektu (TO) *TOAxis*. Vytvorené boli dve TO, *Master_Axis* a *Slave_Axis* pre master a slave motor.

Cieľom je zabezpečiť polohovú a rýchlostnú synchronizáciu osi *Slave_Axis* s osou *Master_Axis*, a simultánne zabezpečiť pohyb motora po posuvnom vedení tak, aby do seba zapadli zuby ozubených kolies osadených na hriadeloch. Vývojový diagram na obrázku

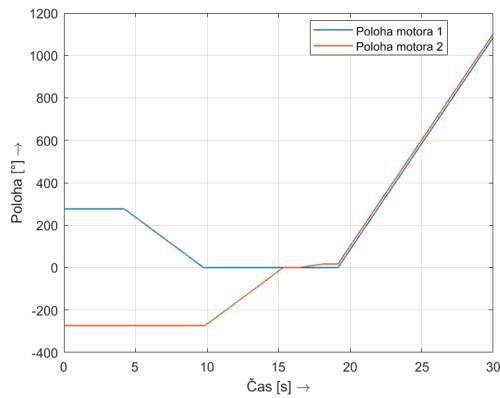
6 graficky znázorňuje chronologickú postupnosť jednotlivých príkazov programu, ktoré sa budú vykonávať tak, aby sa dosiahol vyššie spomínaný cieľ.



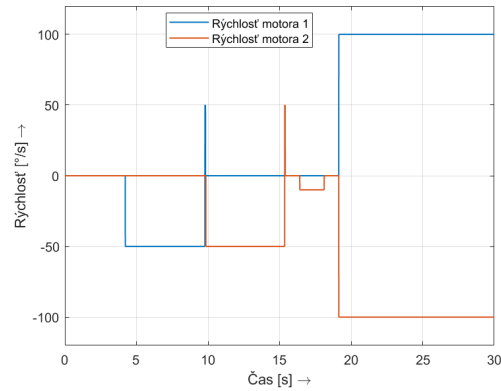
Obr. 6: Vývojový diagram

V. EXPERIMENTÁLNE VÝSLEDKY

Obrázky 7 a 8 sú časové priebehy synchronizácie pri využití paralelnej štruktúry. Rýchlosť otáčania synchronizovaných motorov je $100\text{ }^\circ/\text{s}$, čo predstavuje $16.66\text{ ot}/\text{min}$. Proces začína uvedením motora 1 a následne motora 2 do domovskej polohy. Po sekundovej pauze sa motor 2 natočí o -14.5° . Uhol natočenia vyplýva z geometrie samotných ozubených kolies a je meraný pomocou diferenciálneho snímača polohy na motore. Následne sa motory začnú otáčať rovnakou rýchlosťou, avšak do opačného smeru. Z tohto dôvodu je pre lepšiu názornosť priebeh polohy druhého motora invertovaný.



Obr. 7: Poloha synchronizovaných osí



Obr. 8: Rýchlosť synchronizovaných osí

Keďže sa jedná o bezmomentové spojenie, experimentálne výsledky pri Master-Slave konfigurácií sú identické. Mechatronický systém bol testovaný aj pre vyššie rýchlosti, kde je presnosť, dynamika a synchronizácia stále veľmi presná a spoľahlivá.

VI. ZÁVER

Polohová synchronizácia viacmotorového pohonu môže byť dosiahnutá rôznymi metódami, avšak kvalita synchronizácie sa môže dramaticky líšiť v závislosti od zvolenej metódy. Kvalita a funkčnosť elektronických synchronizačných metód, ktoré sú dnes už samozrejmosťou, boli demonštrované v tejto práci. Dosiahnutá kvalita synchronizácie bola pri všetkých metódach takmer identická a vysoko presná, pohony mali pomerne nízky záťažný moment a zotrvačnosť.

Táto práca predstavuje základ pre možné budúce využitie ako študijnú a výskumnú platformu.

ACKNOWLEDGMENT

This work was supported by Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovakia under Grant FEI-2023-98 *Experimentálne overenie prediktívnej regulácie synchronného motora s permanentnými magnetmi*.

LITERATÚRA

- [1] K. Kyslan and F. Ďurovský, “Elektromechanické výrobné systémy,” *Technická univerzita v Košiciach*.
- [2] Y. Koren, “Cross-Coupled Biaxial Computer Control for Manufacturing Systems,” *Journal of Dynamic Systems, Measurement, and Control*, vol. 102, no. 4, pp. 265–272, 12 1980. [Online]. Available: <https://doi.org/10.1115/1.3149612>
- [3] K. Ismail, T. Rachmildha, E. Rijanto, and Y. Haroen, “A hybrid cross coupling - master slave technique for speed synchronization of multi pmsms system using pid controller and on-line master selector,” *International Journal on Electrical Engineering and Informatics*, vol. 11, pp. 183–194, 03 2019.
- [4] R. Anderson, A. Meyer, M. Valenzuela, and R. Lorenz, “Web machine coordinated motion control via electronic line-shafting,” in *Conference Record of the 1999 IEEE Industry Applications Conference. Thirty-Forth IAS Annual Meeting (Cat. No.99CH36370)*, vol. 1, 1999, pp. 300–306 vol.1.
- [5] H. Huang, Q. Tu, C. Jiang, L. Ma, P. Li, and H. Zhang, “Dual motor drive vehicle speed synchronization and coordination control strategy,” in *Advances in Materials, Machinery, Electronics II*, ser. American Institute of Physics Conference Series, vol. 1955, Apr. 2018, p. 040005.
- [6] R. Lorenz and P. Schmidt, “Synchronized motion control for process automation,” in *Conference Record of the IEEE Industry Applications Society Annual Meeting*, 1989, pp. 1693–1698 vol.2.

Pretlakový plnič PET fliaš

Jaroslava ŽILKOVÁ, Ján KAŇUCH

Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,
Technická Univerzita v Košiciach, Slovensko

jaroslava.zilkova@tuke.sk, jan.kanuch@tuke.sk

Abstrakt — Článok rieši pohon pretlakového plniča PET fliaš, ktorý je súčasťou výrobných liniek. Riešenie zlepšuje výkonnosť procesu plnenia fliaš minerálnou vodou pri zavádzaní moderných technológií riadenia výrobných procesov. Celý návrh zohľadňuje požiadavky objednávateľa, ktorý sa zaoberá čerpaním a plnením minerálnej vody do spotrebiteľských balení. Z hľadiska funkčnosti a ovládania je pretlakový plnič pet fliaš fungujúcim zariadením celého výrobného procesu.

Kľúčové slová — dimenzovanie pohonov, elektrický pohon, plnič pet fliaš, program SEW WORKBENCH

I. ÚVOD

Výroba minerálnej vody je napriek použitiu automatizovaných zariadení veľmi zložitý proces, v ktorom sú obsiahnuté zložité elektrické zariadenia od čerpadiel, servopohonov, meničov, snímačov až po programovateľné automaty. Správnym návrhom a usporiadaním týchto zariadení vzniká linka na výrobu minerálnej vody.

Aby bola činnosť výrobného systému čo najefektívnejšia, musí byť výber motorov, prevodoviek a frekvenčných meničov čo najpresnejší. Dimenzovanie pohonu [1-3] ovplyvňuje rozhodujúcou mierou cenu pohonu a jeho prevádzkové vlastnosti. Chybné dimenzovanie môže byť príčinou škôd, vzniknutých pri prevádzke, prípadne môže znemožniť dosiahnutie požadovaných vlastností pohonu. Pri návrhu elektrického pohonu je potrebné zvoliť predovšetkým druh elektromotora, zdroje s príslušenstvom, riadiacu časť a zariadenie pre styk s obsluhou. Ďalším faktorom sú podmienky elektromagnetickej kompatibility (EMC) – je treba riešiť vplyv meničov na sieť a na motory, vplyvy rušenia ovládacích káblov súbežne idúcich so silovými káblami. Pri voľbe pohonu je dôležité prihliadať na technické hľadisko. Rozhodujúci je druh motora, požiadavky na riaditeľnosť, energetické požiadavky (napätie, veľkosť, spúšťanie), prostredie (teplota, vlhkosť, výbušnosť, hluk, vibrácie), konštrukčné požiadavky (motor s prevodovkou, rozmery elektrického zariadenia vzhľadom ku konštrukčným a stavebným možnostiam). Veľmi dôležité je aj ekonomické hľadisko. Do úvahy musíme brať produktivitu poháňaného zariadenia, výrobné náklady (cena zariadenia, stavebné úpravy, inštalácia a uvedenie do činnosti, servis, záruka), náklady na činnosť (spotrebovaná energia, údržba, opravy a náhradné diely).

Dimenzovanie elektrických pohonov [1-5] je v súčasnom trende rozvoja technologických liniek a rôznych zložitých výrobných zariadení zdĺhavá časť návrhu, či už ide o samotný výpočet síl, momentov a výkonov, alebo o kalkuláciu výslednej ceny. Práve cena sa stala najdôležitejšou položkou v celom procese výroby zariadení. Preto je veľmi dôležité, aby bol pohon pre zariadenie čo najpresnejšie navrhnutý. Musí sa však klásť veľký dôraz na to, aby výsledný pohon zodpovedal všetkým vopred zadaným požiadavkám objednávateľov. Predimenzovaný pohon by znamenal vo viacerých prípadoch prekročenie celkových cenových nákladov na výrobný proces, oteplenie izolácie motora, znemožnenie dosiahnutia požadovaných vlastností pohonu, ale aj zbytočné finančné straty pre firmy v rôznych odvetviach priemyslu [4-5].

V súčasnosti zdĺhavému prepočítavaniu pomáha moderný softvér, ktorý ponúka širokú škálu variácií výrobných systémov, a to napríklad návrh pohonu pre zdvíhanie bremien, rôzne dopravníkové pásy, rotačné stoly, páky, garážové brány alebo skladové zakladače a iné. Moderný software ponúka návrh aj doplnkových zariadení, ako sú frekvenčné meniče, prevodovky a rôzne komponenty nevyhnutné pre správnu činnosť pohonov.

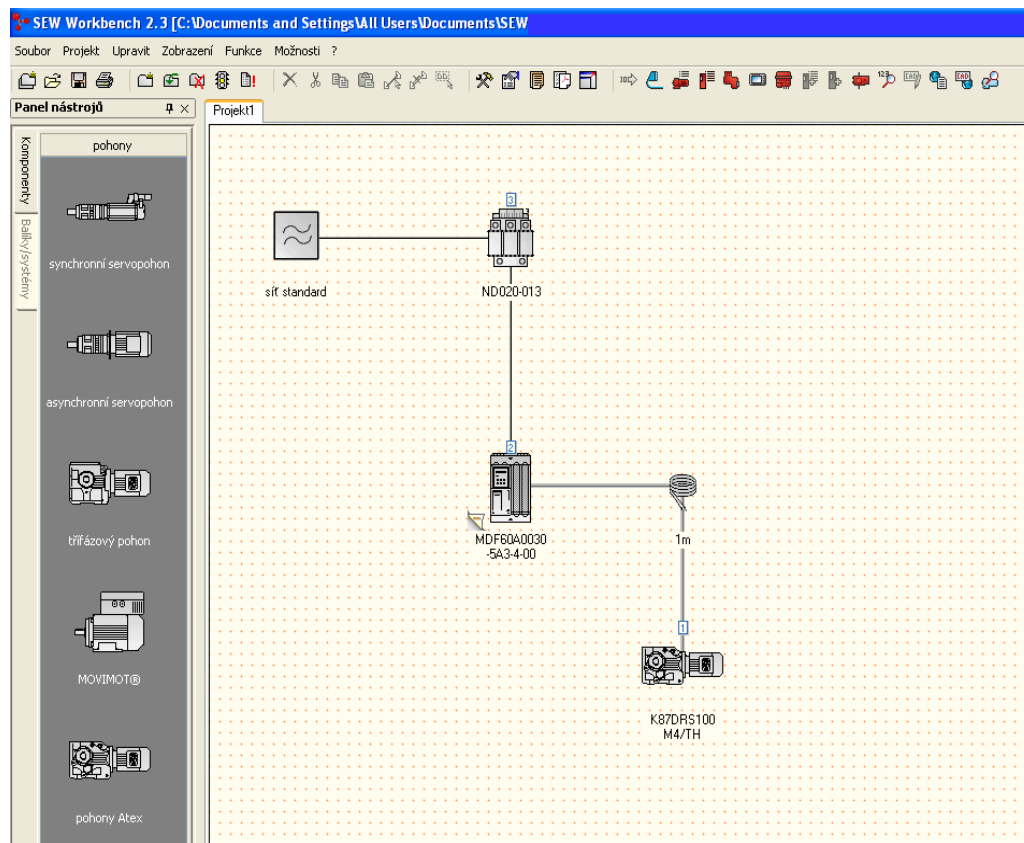
II. PROGRAM SEW WORKBENCH

Program SEW Workbench poskytuje užívateľovi centrálnu pracovnú plochu pre zostavenie a konfigurovanie kompletných systémov pohonov z jednotlivých produktov SEW-Eurodrive: pohonov, meničov, káblov, atď. Systémy sa zostavujú presúvaním objektov, pričom pre výber jednotlivých produktov sú k dispozícii rôzne funkcie. Každý produkt je na pracovnej ploche znázornený ako grafický objekt. Viac objektov je možné zoskupiť do systému pohonu.

Takto zostavené systémy pohonov následne program v niekoľkých stupňoch kontroluje a hlási prípadné chyby – porušenie kombinačných pravidiel. Grafická pracovná plocha odpovedá jednému projektu. Viac pracovných plôch môžeme paralelne začleniť do jednej "projektovvej zložky". Objekty je možné kopírovať, vkladať, vymazať a presúvať. Projekty môžeme uložiť ako súbory PDF alebo vytlačiť. Projektové zložky je umožnené uložiť aj odoslať. Program SEW Workbench môže jednotlivé produkty SEW alebo konfiguráciu ukladať do súboru v pamäti, ktorý si môžu užívatelia programu zostaviť. Výsledkom práce s programom SEW Workbench je zoznam produktov, ktorý je možné prevziať priamo do spracovania zákaziek SEW.

Na pracovnej ploche programu SEW Workbench, zobrazenom na Obr.1 sa nachádza:

- riadok menu s roletovým menu,
- riadok tlačidiel s volaním rôznych funkcií, pričom tlačidlá sú konfigurovateľné,
- panel nástrojov s jednotlivými produktmi SEW-Eurodrive s už nakonfigurovanými kombináciami produktov pre distribuovanú techniku,
- grafická pracovná plocha k navrhovaniu systémov pohonov,
- samostatné, informačné okná (zoznam produktov, popis produktov, zoznam dokumentácie),
- stavový riadok pre rôzne hlásenia programu.



Obr. 1 Pracovná plocha programu SEW Workbench

III. URČENIE POHONU PLNIČA

Pri riešení návrhu pretlakového plniča sme vychádzali zo známych technických a konštrukčných parametrov, ktoré sú uvedené v Tab. 1 - Tab. 4. Postup riešenia pohonu pozostával z určenia prevodu celej sústavy, určenia výslednej účinnosti, zistenia otáčok motora

uvedených v Tab. 5, ktoré sú potrebné pre zabezpečenie žiadanej hodinovej výroby a ďalších technických parametrov, ktoré sú uvedené v Tab. 6. Tieto výpočty boli prevedené na základe metód uvedených v[2-4].

Tabuľka 1
 Základné technické parametre pretlakového plniča

Parameter		Hodnota
Rozsah nastaviteľných výkonov	[fľaše/hod.]	15000 - 17000
Garantovaný výkon fľaša 0,33l	[fľaše/hod.]	16000
Garantovaný výkon fľaša 0,5l EURO	[fľaše/hod.]	12000
Počet plniacich ventilov	[ks.]	40
Počet uzatváracích hláv	[ks.]	8
Presnosť plnenia	[mm]	±3
Hlučnosť	[dB]	85
Maximálny plniaci pretlak	[MPa]	0,6
Hmotnosť	[kg]	5200

Tabuľka 2
 Prehľad konštrukčných parametrov potrebných pre dimenzovanie pohonu

Parameter	Symbol	Hodnota
mechanická účinnosť remenice	η_{rem}	97%
mechanická účinnosť prevodovky	η_{preV}	86%
maximálna uhlová rýchlosť plniča	ω_{max}	0,261 rad.s ⁻¹
minimálna uhlová rýchlosť plniča	ω_{min}	0,105 rad.s ⁻¹
uhlové zrýchlenie na max. rýchlosť plniča	ϵ_r	0,141 rad.s ⁻²
uhlové zrýchlenie na zabrzdzenie	ϵ_b	0,263 rad.s ⁻²
maximálna doba chodu plniča	t_c	3600 s

Tabuľka 3
 Zoznam ozubených kolies a ich konštrukčné parametre

Názov	Počet zubov	Rozstupový priemer [mm]	Modul ozubenia
1. plnič	0	1800	45
2. hnacie kolo plniča	120	1080	9
3. vložené koleso	24	216	9
4. vložené koleso	50	360	7
5. vložené koleso	40	288	7
6. hviezda vstupná	75	540	7
7. hviezda. prevádzacia – plnič	75	540	7
8. uzatváračka	50	360	7
9. hviezda výstupná	50	360	7
10. pastorek hlavného pohonu	25	180	7

Tabuľka 4
 Tabuľka prevodov plniča

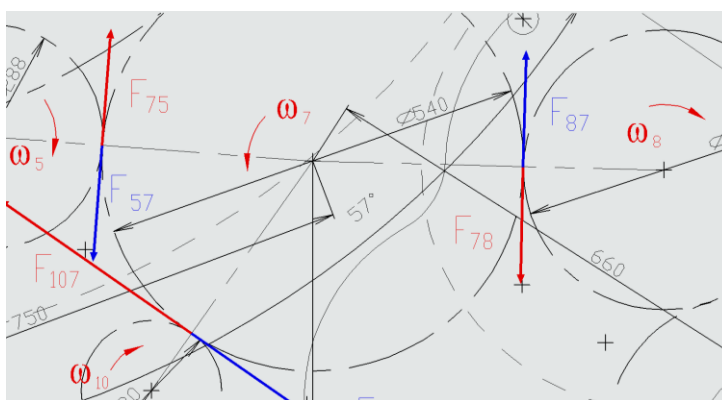
Označenie	Popis	Hodnota
i1	vložené koleso (3) → hnacie koleso (2)	5
i2	vložené koleso (4) → vložené koleso (3)	1
i3	vložené koleso (5) → vložené koleso (4)	1,25
i4	vložené koleso (5) → vstupná hviezda (6)	1,876
i5	prevádzacia hviezda (7) → vložené koleso (5)	0,533
i6	prevádzacia hviezda (7) → uzatváračka (8)	0,666
i7	uzatváračka (8) → výstupná hviezda (9)	1
i8	pastorek hl. pohonu (10) → prevádzacia hviezda (7)	3
i	prevodovka	57
i9	remenica → motor	2

Dôležitou časťou riešenia bol výpočet sily potrebnej k roztočeniu plniča, na základe ktorej bol určený potrebný moment motora. Na Obr. 2 sú zobrazené silové účinky na ozubeniach

vybraných kolies, ktoré boli potrebné pre výpočty prítlakových síl a sily potrebnej na roztočenie plniča.

Tabuľka 5
 Prehľad uhlových rýchlostí jednotlivých častí plniča pet fliaš

Názov	n [ot.min ⁻¹]	ω [rad.s ⁻¹]
1. plnič	2,5	0,261
2. hnacie kolo plniča	2,5	0,261
3. vložené koleso	12,5	1,308
4. vložené koleso	12,5	1,308
5. vložené koleso	15,6	1,633
6. hviezda vstupná	8,3	0,896
7. hviezda.prevážacia – plnič uzatváračka	8,3	0,869
8. uzatváračka	12,5	1,308
9. hviezda výstupná	12,5	1,308
10. pastorok hlavného pohonu	25	2,617



Obr.2 Silové účinky na ozubení kolies 7 a 8 a 10

Výpočtom ekvivalentného momentu M_{ekv} , podľa metód uvedených v [3 a 4], bolo overené, že vybraný motor s vlastným chladením nebude preťažovaný. Z katalógu 2-pólových motorov firmy ABB bol vybraný motor s vlastným chladením s typovým označením M2Q2 100 L2A. Vybraný motor bude pri rozbehu preťažovaný momentom $M_r' = 10,539 N.m$, čo predstavuje 1,05 násobok nominálneho momentu motora. Ďalším krokom výpočtu bol návrh meniča pre daný motor. Zabezpečiť ochranu motora proti preťaženiu, zmenu otáčok, možnosť nastavenia maximálnych a minimálnych otáčok, rýchlosť rozbehu a zastavenia a možnosť znižovať rýchlosť plniča voči zásobe fliaš pred plničom a za plničom bude frekvenčný menič s vektorovým riadením od firmy Danfoss. V Tab. 6 sú uvedené výsledné hodnoty získané výpočtom.

Tabuľka 6
 Tabuľka vypočítaných hodnôt

Parameter	Hodnota
Celková účinnosť	78,1 %
Výsledná sila potrebná na roztočenie plniča	8989,12 N
Uhlové zrýchlenie na hriadeľ motora pri rozbehu	160,63 rads ⁻¹
Uhlové zrýchlenie na hriadeľ motora pri brzdení	299,631 rads ⁻¹
Moment motora pri rozbehu	10,539 Nm
Moment motora pri brzdení	3,836 Nm
Čas rozbehu	1,85 s
Čas brzdenia	0,99 s
Čas chodu konštantnou rýchlosťou	3597,15 s
Výkon motora	3,133 kW
Ekvivalentný moment s uvážením doby kľudu	8,454 Nm

Určenie pohonu plniča pet fliaš bolo realizované aj programom SEW Workbench. Pre výpočet pohonu plniča fliaš pomocou programu SEW Workbench bol ako podkladový príklad zvolený rotačný stôl, pretože sa svojimi vlastnosťami, parametrami a systémom výpočtu najviac približoval existujúcemu plniču pet fliaš. Vzhľadom na to, že príklad rotačného stola nemá úplne totožné konštrukčné vyhotovenie ako plnič, museli sa dodatočne dopočítavať parametre potrebné pre najlepšie priblíženie sa k správne výsledku.

Prvou fázou výpočtu bolo nájdenie vhodného motora na to, aby daný výrobný systém zodpovedal požiadavkám objednávateľa. Štruktúra programu viedla k postupnému zadávaniu parametrov potrebných k realizácii výpočtu. Program umožňoval dokonalé nastavenie časov nábehu a dobehu respektíve nastavenie uhlového zrýchlenia pri rozbehu a uhlového zrýchlenia – spomalenia pri dobehu systému. Po zadaní všetkých potrebných konštánt program následne previedol celý výpočet pre výber konkrétneho motora z katalógu implementovaného priamo v programe. V ďalšej fáze boli realizované výpočty pre výber vhodnej prevodovky do systému tak, aby boli splnené požiadavky na hodinovú výrobu. Program vybral 3 kW motor no moment na výstupnom hriadelí prevodovky 741 N.m by nebol postačujúci oproti pôvodnému momentu. Táto odchýlka mohla vzniknúť kvôli vyššie spomenutým faktom, že plnič a rotačný stôl nemajú úplne totožnú konštrukciu a vstupné konštrukčné parametre. V projektovej správe program vygeneroval motor s kompletnými parametrami, a preto bolo možné vybrať vhodný frekvenčný menič spolu s ďalším príslušenstvom potrebným pre správnu činnosť celého výrobného systému.

Výsledok výpočtu spolu s projektovou správou vygenerovanou programom Workbench bol výstupom programu.

IV. ZÁVER

V predloženej príspevku je uvedený návrh pohonu pretlakového plniča minerálnej vody. Po oboznámení sa s technológiou plnenia, prevedením plniča a zistení súčasného stavu sme po analýze navrhli riešenie hlavného pohonu. Pohon plniča minerálnej vody sme dimenzovali, podľa dostupnej technickej dokumentácie vyššie uvedeným postupom, vypočítaním sily potrebnej k roztočeniu plniča, na základe ktorej sme určili ekvivalentný moment motora a následne vybrali vhodný motor a frekvenčný menič. Návrh pohonu sme riešili aj pomocou programu SEW Workbench, pričom výstupom programu bola projektová správa, ktorá obsahovala výsledky výpočtu a navrhovaný motor s uvedenými parametrami, pre ktorý bolo možné vybrať vhodný frekvenčný menič spolu s ďalším príslušenstvom z portfólia firmy SEW Eurodrive.

Pri konečnej realizácii pohonu plniča realizačná firma použila komponenty rôznych dodávateľov podľa predchádzajúceho vypočítaného riešenia. Motor dodávala firma ABB, frekvenčný menič firma Danfoss a prevodovku firma SEW-Eurodrive. Z hľadiska funkčnosti a ovládania je pretlakový plnič pet fliaš fungujúcim zariadením celého výrobného procesu.

POĎAKOVANIE

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23.

LITERATÚRA

- [1] C. Fantuzzi, F. Basissi, "Filling System Fuzzy Control of Packaging Machinery", Università di Ferrara, 2008.
- [2] V. Kus: "Elektrické pohony a výkonová elektronika", ZU v Plzni, 2005.
- [3] F. Poliak a kol.: "Elektrické pohony", Bratislava, Alfa, 1987.
- [4] J. Timko, J. Žilková, J. Dudrik, P. Girovský, "Elektrické akčné členy a pohony", FEI TU v Košiciach, 2009.
- [5] L. Zboray, F. Ďurovský, J. Tomko a kol., "Regulované pohony", Viena, 2000.
- [6] M. Pastierovic, "PLC integrované do frekvenčného meniča". AT&P Journal, .no5, 2006.

Meranie indukčností vinutí elektrických motorov s použitím samovyvažovacieho mostíka a mikrokontroléra

¹Daniel GORDAN, ²Tomáš BASARIK, ³Viktor ŠLAPÁK

^{1,2,3}Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Slovenská republika

¹daniel.gordan@student.tuke.sk, ²tomas.basarik@student.tuke.sk,
³viktor.slapak@tuke.sk

Abstrakt – Príspevok má za úlohu opísať meranie indukčnosti samovyvažovacím mostíkom ako jednu z dostupných metód merania indukčnosti, vhodnú pre implementáciu a vyhodnotenie mikrokontrolérom. Prípevok opisuje základný princíp metódy a spôsob vyhodnotenia nameraných údajov. Metóda je dokumentovaná experimentálnymi výsledkami.

Kľúčové slová – indukčnosť, meranie, mikrokontrolér, samovyvažovací mostík

I. ÚVOD

Správne určenie impedancie prvkov je dôležité v mnohých oblastiach návrhu elektrických obvodov a strojov. V prípade elektrických motorov je impedancia vinutí dôležitým parametrom pre návrh regulácie a pre určenie dynamiky prúdového obvodu. Návrh regulácie len na základe katalógových údajov motora nie je vhodný, keďže indukčnosť motora je udávaná s toleranciou 10 % alebo vyššou. Je preto veľmi dôležité určiť, resp. odmerať impedanciu vinutí každého vyrobeného motora.

Vo všeobecnosti môže byť meranie impedancie založené na rôznych fyzikálnych princípoch, využívajúc pri tom snímanie úbytku napätia na meranom prvku, či jeho frekvenčnú charakteristiku. Medzi najjednoduchšie metódy patrí meranie impedancie mostíkovou metódou, rezonančnou metódou a volt-ampérovou metódou. Stručný prehľad týchto metód meraní impedancie možno nájsť v [1]. V literatúre možno nájsť aj menej tradičné metódy pre presné určenie impedancií alebo indukčností, ako je napríklad meranie indukčnosti kapacitným mostíkom [2] či metódy merania zložiek indukčnosti v súradnicovom systéme dq , ktoré sú určené pre identifikáciu parametrov synchronných motorov s permanentými magnetmi (SMPM). Analýzu a prehľad rôznych metód pre určenie indukčností pre SMPM s magnetmi na povrchu rotora, vrátane porovnania meraní pomocou budenia sínusovým a PWM signálom, možno nájsť v [3]. Okrem offline merania je možné identifikovať indukčnosť SMPM aj pomocou online metód, napr. využitím injekcie [4], priamo z budenia pomocou PWM a snímania zmeny prúdu [5] alebo estimáciou počas riadenia [6]. Online metódy nájdu využitie najmä v bezsnímačovom riadení SMPM. Okrem rotačných SMPM sa taktiež skúma aj meranie indukčností pre lineárne SMPM [7].

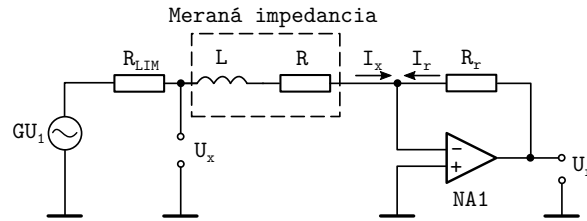
V tomto príspevku opíšeme metódu merania pomocou samovyvažovacieho mostíka, ktorá je vhodná pre svoju jednoduchú implementáciu a dostatočne vysokú presnosť merania. V nasledujúcej sekcii bude opísaný základný princíp tejto metódy a následne v tretej sekcii bude načrtnutá implementácia meracej metódy a opísaný spôsob vyhodnotenia snímaných signálov. Štvrtá sekcia je venovaná experimentálnym výsledkom.

II. SAMOVYVAŽOVACÍ MOSTÍK

Samovyvažovací mostík využíva pre meranie indukčnosti striedavé napätie, ktoré je vyvažované operačným zosilňovačom [1]. Operačný zosilňovač v zapojení podľa Obr. 1 vytvára na svojom výstupe také napätie, aby pre prúd referenčného rezistora v zápornej spätnej väzbe operačného zosilňovača platilo:

$$I_x = I_r, \quad (1)$$

kde I_x je prúd, pretekajúci meranou impedanciou, a I_r je prúd, pretekajúci referenčným rezistorom R_r . Rezistor R_{lim} je rezistor limitujúci prúd zo zdroja GU_1 , vytvárajúceho budiace napätie U_x .



Obr. 1 Principiálna schéma merania so samovyvažovacím mostíkom

Prúd I_r je následne možné určiť meraním výstupného napätia U_r operačného zosilňovača pomocou rovnice:

$$I_r = \frac{U_r}{R_r}, \quad (2)$$

kde R_r je odbor referenčného rezistora v spätnej väzbe operačného zosilňovača. Výslednú veľkosť impedancie je potom možné určiť ako:

$$|Z_x| = \frac{U_x}{I_x} = \frac{U_x}{I_r} = R_r \frac{U_x}{U_r}. \quad (3)$$

Jednotlivé prvky impedancie, t.j. odpor R a reaktanciu X_L , je možné potom určiť pomocou vzťahov:

$$R = |Z_x| \cos \phi, \quad (4)$$

$$X_L = |Z_x| \sin \phi, \quad (5)$$

kde ϕ je fázový posun medzi napätiami U_r a U_x , ktorý zodpovedá fázovému posunu medzi budiacim napätím na meranej impedancii a prúdu, ktorý ňou preteká. Indukčnosť sa následne vypočíta podľa nasledujúceho vzťahu:

$$L = \frac{X_L}{2\pi f}, \quad (6)$$

kde f je frekvencia budiaceho napätia U_x .

Výhodou opísaného zapojenia je, že nie je potrebné merať prúd pretekajúci meranou impedanciou, ale namiesto toho je možné merať napätie na výstupe operačného zosilňovača, čo je jednoduchšie z hľadiska ďalšieho spracovania signálov pomocou mikrokontroléra. Zmenou referenčného rezistora je zase možné jednoducho meniť rozsah merania, resp. je možné impedančne prispôsobiť merací obvod.

III. VÝHODNOTENIE MERANIA MIKROKONTROLÉROM

Pri praktickej realizácii je výhodné pre generovanie budiaceho napätia použiť integrovaný generátor signálov so zosilnením výstupného napätia a prúdu pomocou operačných zosilňovačov. Následne, ak sa na meranie použije mikrokontrolér a samovyvažovací mostík, tak je jednoduché snímať obe potrebné napätia pomocou AD prevodníka mikrokontroléra, a to napätie na meranej impedancii U_x (t.j. budiace napätie) a napätie na referenčnom rezistore U_r .

Vzorkovanie signálu môže byť teoreticky spustené napr. pri prechode budiaceho signálu nulou napätia a fázový posun by tak bolo možné určiť odmeraním času od tohto okamihu do prechodu nulou snímaným napätím U_r . Vzorkovanie týmto prístupom však nemusí dosahovať dostatočnú presnosť, keďže signály môžu byť ovplyvnené šumom. Vhodnejší prístup je preto súčasne navzorkovať jednu periódu oboch napätí a následne sa na základe týchto vzoriek aproximuje sínusový signál.

A. Aproximácia snímaných signálov

Aproximáciu snímaných signálov je možné vypočítať na základe trojparametrového algoritmu, uvedeného v [8]. Aproximácia vychádza z nasledujúceho matematického vyjadrenia sínusového signálu v čase t :

$$y(t) = A_0 \cos(2\pi ft) + B_0 \sin(2\pi ft) + C_0, \quad (7)$$

kde A_0 a B_0 sú amplitúdy kosínusovej a sínusovej zložky a C_0 je veľkosť jednosmernej zložky.

Ak je vektor nasnímaných vzoriek \mathbf{x} definovaný ako:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad (8)$$

kde x_1 až x_n sú nasnímané vzorky v časovom slede, potom koeficienty A_0 , B_0 a C_0 je možné určiť na základe vzťahu:

$$\begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} = (\mathbf{D}_0^T \mathbf{D}_0)^{-1} (\mathbf{D}_0^T x). \quad (9)$$

Matica D_0 v rovnici (9) je konštantná matica vytvorená len na základe známej frekvencie f budiaceho napätia a definovanej doby vzorkovania T_s nasledovne:

$$\mathbf{D}_0 = \begin{bmatrix} \cos(2\pi f 0 T_s) & \sin(2\pi f 0 T_s) & 1 \\ \cos(2\pi f 1 T_s) & \sin(2\pi f 1 T_s) & 1 \\ \cos(2\pi f 2 T_s) & \sin(2\pi f 2 T_s) & 1 \\ \vdots & \vdots & \vdots \\ \cos(2\pi f (n-1) T_s) & \sin(2\pi f (n-1) T_s) & 1 \end{bmatrix}, \quad (10)$$

kde n je počet zosnímaných vzoriek.

Pre určenie fázového posunu získanej aproximácie je potrebné rovnicu (7) prepísať do amplitúdovo-fázového tvaru:

$$y(t) = A \cos(2\pi f t + \theta), \quad (11)$$

kde θ je fázový posun funkcie od počiatku súradnicového systému a A je amplitúda funkcie. Amplitúdu A je možné vyjadriť pomocou už získaných koeficientov A_0 a B_0 nasledovne:

$$A = \sqrt{A_0^2 + B_0^2}. \quad (12)$$

Fázový posun θ je taktiež možné určiť z koeficientov A_0 a B_0 podľa vzťahu:

$$\theta = \arctan\left(\frac{B_0}{A_0}\right). \quad (13)$$

B. Výpočet odporu a indukčnosti

Pre výpočet odporu a indukčnosti sa použijú vzťahy (4) až (6), pre ktoré platí, že fázový posun ϕ určíme nasledovne:

$$\phi = \theta_x - \theta_r, \quad (14)$$

kde θ_x a θ_r sú fázové posuny napätí U_x a U_r voči počiatku vzorkovania, získané aproximáciou podľa (7) až (13).

IV. EXPERIMENTÁLNE VÝSLEDKY

Opísaná metóda bola experimentálne overená s využitím mikrokontroléra TMS320F28035. Ako zdroj budiaceho napätia slúžil generátor signálov AD9833, ktorého výstup bol zosilnený pomocou operačných zosilňovačov, tak aby jeho efektívna hodnota bola 1 V a max. prúd bol na úrovni 250 mA. Frekvencia f budiaceho napätia bola stanovená na 1 kHz a doba vzorkovania T_s bola na úrovni 2.5 μ s, čo odpovedalo 400 vzorkám na jednu periódu budiaceho napätia.

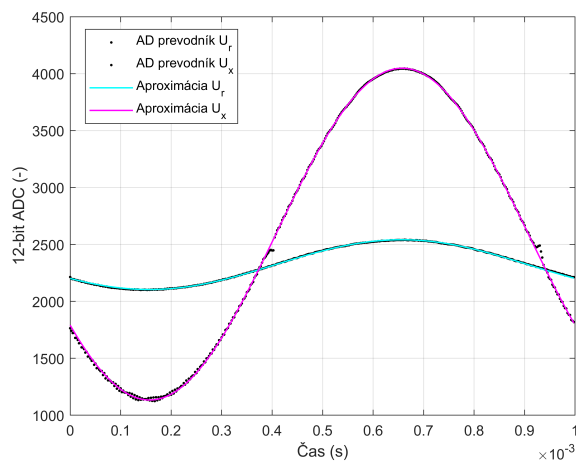
Mikrokontrolér disponuje 12-bitovým AD prevodníkom, ktorý umožňuje simultánne vzorkovanie 2 signálov. Ako už bolo spomenuté, v tomto prípade sa jednalo o napätia U_x a U_r , ktorých priebeh a aproximácia sú na Obr. 2, kde sú napätia reprezentované ako číselný výsledok AD prevodu. Ako je vidieť, uvedená aproximácia značne eliminovala šum, prítomný v meraní.

Opísaná metóda bola experimentálne verifikovaná meraniami indukčností jednotlivých fáz PMSM motorov. Experimentálne výsledky a ich porovnanie s výsledkom merania pomocou komerčného LCR metra HIOKI 3511-51 sú uvedené v tab. 1. Ako je z výsledkov vidieť, opísaná metóda merania a vyhodnotenia nameraných výsledkov dosahuje vysokú presnosť v jednotkách percent. Väčšia odchýlka je badateľná pri nízkej hodnote indukčnosti, čo je spôsobené veľmi malým fázovým posunom. Dosiahnutie vyššej presnosti by bolo možné rýchlejšími vzorkovaním snímaných napätí.

V. ZÁVER

Príspevok opísal jednu z dostupných metód merania indukčností vinutí elektrických motorov, ktorá je vhodná pre praktickú implementáciu s využitím mikrokontroléra. Ako bolo ukázané v experimentálnych výsledkoch, meranie pomocou zvolenej metódy dosahuje dostatočne vysokú presnosť v porovnaní s komerčným LCR metrom, pričom si zachováva jednoduchosť a relatívne nízku výpočtovú náročnosť.

Je potrebné spomenúť, že aj keď opísaná metóda umožňuje odmerať impedanciu, na základe ktorej je možné určiť aj odpor meranej impedancie, tak je pre výpočty regulácie či modelov elektrických motorov vhodnejšie odmerať odpor vinutí jednosmerným prúdom jednoduchou V-A metódou.



Obr. 2 Priebek snímaných napätí a ich aproximácia

Tabuľka 1
 Experimentálne výsledky

Referenčná hodnota [mH]	Odmeraná hodnota [mH]	Relatívna chyba [%]
0.380	0.411	8.351
1.200	1.22	1.633
16.110	16.035	0.463
32.430	33.575	3.531
52.540	53.929	3.234
60.250	61.315	1.767

PODAKOVANIE

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23.

LITERATÚRA

- [1] Keysight Technologies, "Impedance Measurement Handbook: A Guide to Measurement Technology and Techniques," *Application Note, Keysight Technologies, USA*, 6th Edition, July 2020.
- [2] J. Horská, J. Horský, "Precision inductance measurement on high precision capacitance bridge," *2008 Conference on Precision Electromagnetic Measurements Digest, Broomfield, CO, USA*, vol. 39, pp. 572–573, June 2008, doi: 10.1109/CPEM.2008.4574908.
- [3] H. B. Ertan and I. Sahin, "Inductance Measurement Methods for Surface-Mount Permanent Magnet Machines," in *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1-16, 2023, Art no. 2000116, doi: 10.1109/TIM.2022.3225048.
- [4] Yang Ge and Weizhang Song, "An asymmetric-square-wave voltage injection method for online inductance identification of PMSM under sensorless control," *Energy Reports*, vol. 9, supp. 3, pp. 168-174, 2023, ISSN 2352-4847, doi: 10.1016/j.egy.2022.12.127.
- [5] R. Raja, T. Sebastian and M. Wang, "Online Stator Inductance Estimation for Permanent Magnet Motors Using PWM Excitation," in *IEEE Transactions on Transportation Electrification* vol. 5, no. 1, pp. 107-117, March 2019, doi: 10.1109/TTE.2019.2891047.
- [6] S. Ye and X. Yao, "A Modified Flux Sliding-Mode Observer for the Sensorless Control of PMSMs With Online Stator Resistance and Inductance Estimation," in *IEEE Transactions on Power Electronics*, vol. 35, no. 8, pp. 8652-8662, Aug. 2020, doi: 10.1109/TPEL.2019.2963112.
- [7] Liyi Li, Junjie Hong, Hongxing Wu, Baoquan Kou, Rizhong Liu, "Direct and quadrature inductances measurement of the permanent magnetic linear synchronous machines," *Energy Conversion and Management*, vol. 52, issue 5, pp. 2282-2287, 2011, ISSN 0196-8904, doi: 10.1016/j.enconman.2010.12.020.
- [8] IEEE Standards Association, "IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters," in *IEEE Std 1241-2010 (Revision of IEEE Std 1241-2000)*, pp.1-139, 14 Jan. 2011, doi: 10.1109/IEEE-ESTD.2011.5692956.

Troj-statorový diskový synchronný motor s permanentnými magnetmi

¹Ján KAŇUCH

¹ Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach, Letná 1/9, 04200 Košice-Sever, Slovenská republika

¹jan.kanuch@tuke.sk

Abstrakt — V tomto príspevku je prezentovaný návrh viacdiskového synchronného motora s permanentnými magnetmi. V prvej časti sú popísané konštrukcie diskových motorov s obojstranným ťahom. V nasledujúcich častiach sú navrhnuté rozmery a je popísané konštrukčné usporiadanie elektromagnetického obvodu diskového synchronného motora, ktoré rešpektuje jeho technickú realizáciu v troj-statorovom prevedení. Výpočty motora sú robené v programe Matlab a model nakreslený v programe AutoCad.

Kľúčové slová — synchronný motor, permanentný magnet, diskový rotor.

I. ÚVOD

Vývoj motorov a pohonov, ktorý sa v súčasnosti realizuje a ktoré sa označujú ako „moderné“, tak ich princíp bol objavený a používaný už v 19. storočí. Prikladom jedného z takýchto motorov je aj synchronný stroj. Synchronný stroj je točivý elektrický stroj, ktorý premieňa elektrickú energiu na mechanickú alebo naopak. Je to najpoužívanejší stroj pri výrobe energie [1]. Je všeobecne známe, že z hľadiska premeny energie môže pracovať synchronný v dvoch prevádzkach: v generátorickej a aj motorickej.

Konštrukcia synchronného stroja je podobná asynchronnému motoru. Vinutie statora je zvyčajne trojfázové, ale v súčasnosti prebieha aj vývoj s iným počtom fáz [2]. V prípade rotora sú dva typy: hladký rotor a rotor s vyjadrenými pólmi. Budiace vinutie na rotore je napájané zo zdroja jednosmerného prúdu. Jednosmerný budiaci prúd sa privádza na rotor cez dva klzné kontakty (uhlíky), ktoré sa kľžu po mosadzných krúžkoch umiestnených na hriadeli rotora. Pri strojoch menšieho výkonu sa budiace vinutie nahradí permanentným magnetom. V takom prípade sú to synchronne stroje s permanentným magnetom a dostávajú sa do popredia posledné desaťročia vďaka ich výhodám, ktoré sú bežne popísané aj v staršej bežne dostupnej literatúre [3]. V súčasnosti sú bežne prezentované a prebieha aj vývoj konštrukcie a výskum riadenia viacfázových synchronných motorov s permanentnými magnetmi (PMSM) do výkonu až 200 kW, ktoré sú prevažne určené pre pohony elektrických vozidiel [4], [5].

Moderné regulačné systémy vyžadujú rýchle, dynamické pohony s presne definovanou polohou. Pri malých výkonoch strojov s vysokou dynamikou je možné konštruovať motory s diskovým rotorom a magnetizáciou v axiálnom smere. V súčasnosti sú známe takéto jednosmerné a asynchronne motory, BLDC a BLAC motory, motory s permanentnými magnetmi s diskovým rotorom, pričom všetky majú veľmi dobré dynamické vlastnosti v dôsledku malého momentu zotrvačnosti rotora stroja. Z toho vyplýva aj konštrukcia diskového synchronného motora s permanentnými magnetmi s axiálnou vzduchovou medzerou, ktorý by zlučoval výhodné vlastnosti synchronného motora s permanentnými magnetmi valcového typu s dynamickými vlastnosťami motorov s diskovým rotorom.

Pre potreby integrácie pohonných motorov s planétovými prevodmi v moderných elektromechanických pohonných systémoch sa často používajú diskové synchronne motory s permanentnými magnetmi, z čoho vyplýva aj nutnosť návrhu a výskumu týchto viacrotorových strojov.

II. PERMANENTNÉ MAGNETY PRE BEZKEFOVÉ SYNCHRONNE MOTORY

Posledné desaťročia sa do popredia dostávajú synchronne motory s budením z permanentných magnetov. Pri návrhu takéhoto bezkefového motora je jednou z požiadaviek správny výber permanentných magnetov. Základnými kritériami na výber magnetov sú [6]:

1. cena. Pre pohony, ktoré sa vyrábajú čo najlacnejšie, sa môžu použiť ferity, ak sú

vlastnosťami dostačujúce. Najdrahšie magnety sú zo vzácnych zemín (napr. magnet typu Sm-Co). Cena permanentných magnetov posledné roky neustále rastie a je ovplyvňovaná okrem okrem obchodných záujmov aj politickou situáciou. Dominantná krajina v exporte magnetov zo vzácnych zemín je Čína.

2. merný moment motora. Magnety dokážu pre určitý objem motora vyprodukovať obmedzený moment motora a pri niektorých typoch magnetov ten moment nemusí byť dostačujúci.

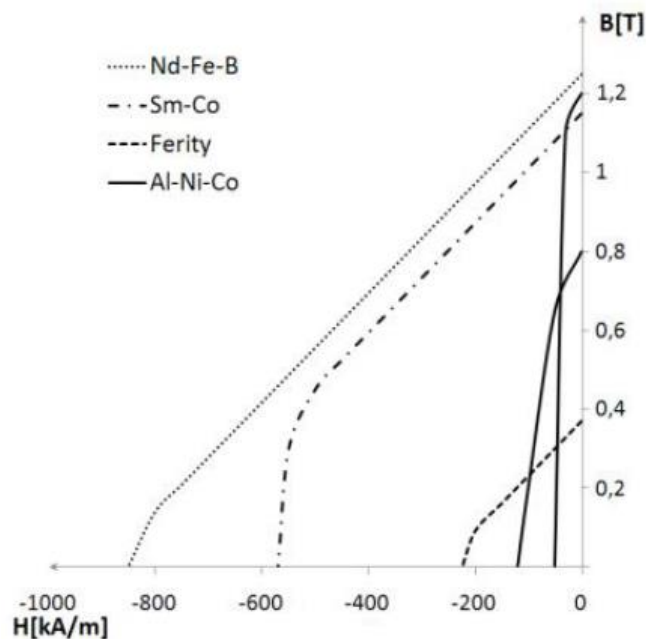
3. vysoká korektivita. Vysoká koercitivita (HC) znamená odolnosť magnetu voči demagnetizácii. Koercitivita je hraničná hodnota intenzity magnetického poľa, ktorá ak sa prekročí, magnet sa odmagnetuje.

4. rozsah prevádzkovej teploty. Magnety zo vzácnych zemín majú nízku Curieho teplotu, preto pri vyšších zaťaženiach motora je veľmi dôležitý systém chladenia rotora.

5. remanencia. Najsilnejšie magnety nemusia byť zárukou jednoduchosti a kvality. Magnety s vysokou remanenciou môžu vytvárať prehnane vysoké magnetické indukcie v motore a vysoká indukcia znamená nasýtenie kritických častí magnetického obvodu, čo predstavuje vyššie straty.

6. rozmery magnetu. Na vybudenie malého magnetického napätia na určitej ploche postačujú ferity. Matematicky sa dá také isté napätie dosiahnuť neodýmovými magnetmi (NdFe). Prakticky ale môže vyjsť tak malá šírka magnetu, že by bolo náročne takýto magnet vyrobiť a použiť. Taktiež použitie magnetov skupiny AlNiCo môže byť obmedzené pre náročné rozmery a tvar magnetu, lebo magnety tejto skupiny sa ťažšie obrábajú (sú tvrdé a krehké).

Pre praktické využitie sa udáva energetický súčin BH_{max} , ktorý charakterizuje kvalitu permanentného magnetu a vypovedá o miere energie uschovanej v magnetu (Obr. 1).



Obr. 1 BH krivky permanentných magnetov

III. MOTORY S DISKOVÝM ROTOROM

Magnetický tok valcového motora sa uzatvára v radiálnom smere a magnetický tok diskového motora sa uzatvára v axiálnom smere na os motora.

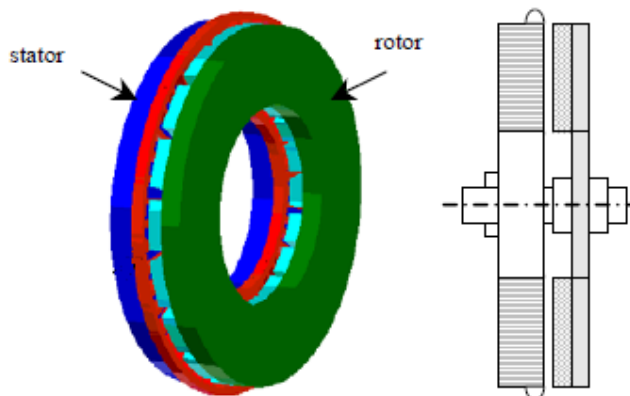
Prvýkrát sa diskové motory s PM skonštruovali koncom 70-tych a začiatkom 80-tych rokov minulého storočia [7], [8]. Konštruktérmi boli Campbell (1975), Leung a Chan (1980), Weh a kol. (1984). Záujem o nich v posledné desaťročia rástol hlavne pre ich vysoký pomer momentu k hmotnosti motora a vysokú účinnosť. Oproti motorom s radiálnym tokom sa používajú tenšie magnety. Tiež hluk a vibrácie, ktoré diskové motory produkujú, sú menšie ako u bežných strojov.

Napredovanie pri diskových motoroch nastalo s pokrokom vo výkonovej elektronike. Diskové motory sa môžu vyrábať o výkonoch až do stoviek kW. Diskový motor môže mať v porovnaní s valcovým motorom rovnakého objemu a použitého materiálu až o 78% vyšší výkon [9]. Možné konštrukčné riešenia diskových motorov sú uvedené nižšie.

A. Diskový motor s jedným rotorom a jedným statorom

Takýto typ motora má jednoduchú konštrukciu. Rotor a stator sú paralelne vedľa seba. Na rotore vo vnútornej strane sú permanentné magnety, na statore vinutie v radiálnom smere a medzi

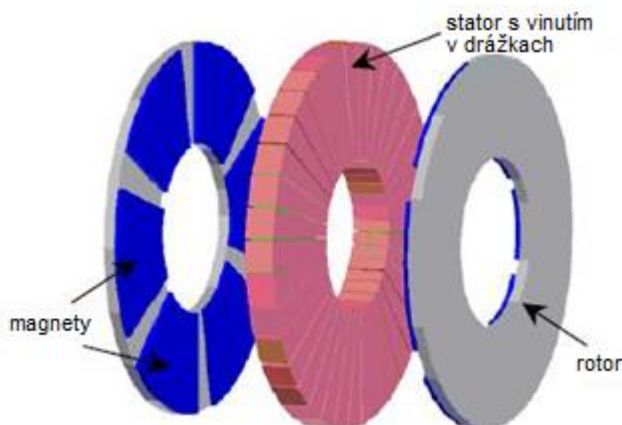
magnetmi a pólmi statora je vzduchová medzera (Obr. 2).



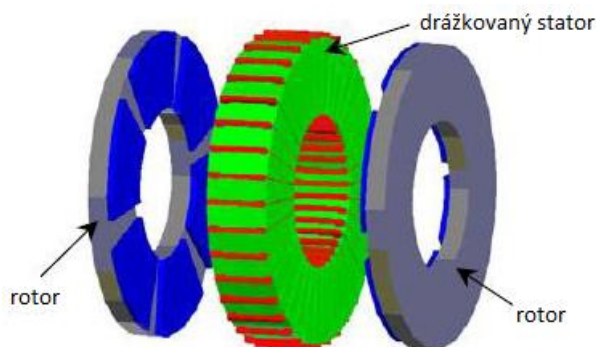
Obr. 2 Diskový motor s jedným rotorom, jedným statorom a axiálnym vinutím statora.

B. Diskový motor s jedným statorom a dvomi rotormi

Prvýkrát bola táto konštrukcia vyhotovená až koncom 80-tych rokov. V strede je stator a na kraji sú rotory. Táto konštrukcia rieši problémy axiálnych síl pôsobiacich na ložisko, keďže sa anulujú. Podľa prevedenia vinutia statora môže byť stator s drážkami (Obr. 3). alebo bezdrážkový (Obr. 4). kde jednotlivé vinutia sú oddelené vzduchovou medzerou vyplnenou epoxidovou živicom, ktorá dobre odvádza teplo.

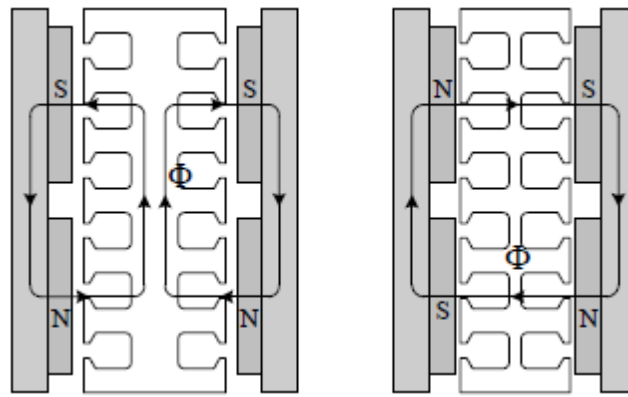


Obr. 3 Diskový motor TORUS-S s bezdrážkovým statorom.



Obr. 4 Diskový motor TORUS-S s drážkovaným statorom

Magnety na diskoch oproti sebe môžu byť nesúhlasne polarizované (rovnaké póly magnetov sú oproti sebe – NN) alebo súhlasne polarizované (opačné póly sú oproti sebe - NS), ako je vidieť na Obr. 5.



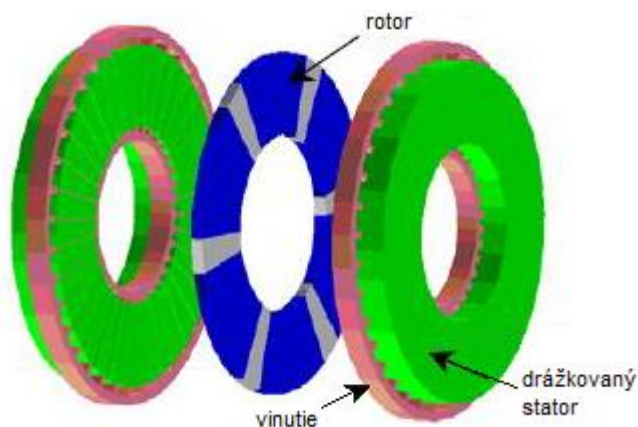
Obr. 5 Magnetický tok motormi TORUS s rotormi typu NN a NS.

Pri usporiadaní magnetov NN (SS) magnetický tok z jedného rotora neprechádza na druhý rotor, ale sa uzatvára cez jarmo statora (Obr. 4 vľavo). Pri NS usporiadaní magnetov sa hlavný magnetický tok pôsobením magnetov neuzatvára cez jarmo statora, ale prechádza oboma rotormi (Obr. 4 vpravo). Stator môže mať teda menšiu hrúbku (nepotrebujeme jarmo statora), čo vedie k vyššej účinnosti motora a zníženiu jeho hmotnosti. Podstatný rozdiel medzi statorom s drážkami a statorom bez drážok je aj ten, že v prípade bezdrážkového statora sa použije menej medi, čoho dôsledok sú menšie straty. Okrem toho je tu lepšie odvádzané aj teplo zo statora.

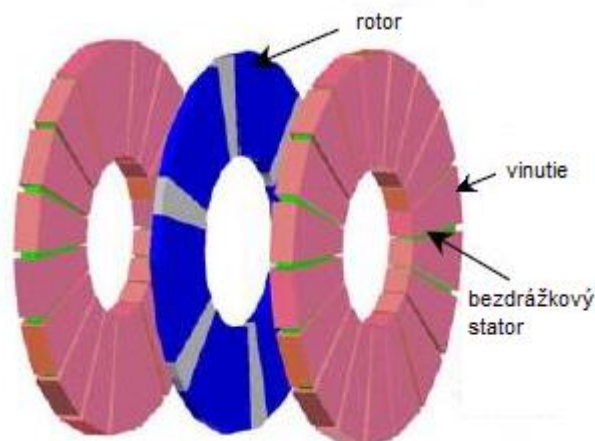
Ďalšími výhodami motora typu TORUS je aj menšia vzájomná indukčnosť, eliminovanie vyšších harmonických z priebehu momentu motora, menšie vysokofrekvenčné straty v rotore a v zuboch statora.

C. Diskový motor s dvomi statormi a jedným rotorom (AFIR)

Tento typ motora rieši problém pôsobenia axiálnych síl na ložiská rotora, keďže axiálne sily sa rozdelia a pôsobia vyrovnane oproti sebe. Stator môže byť drážkovaný alebo bez drážok. Priestor medzi drážkami sa vyplní živicom, ktorá dobre odvádza teplo. j v tomto prípade platí, že pri motore s drážkami sa použije viac medi ako v prípade motora bez drážok. Rotor je odlišný od typu TORUS. Permanentné magnety sú na rotorovom disku alebo v ňom a vyplňajú celú hrúbku. Magnetický tok má prechádzať rotorom iba axiálnym smerom. To sa dosiahne, ak sa rotorový disk vyrobí z magneticky nevodivého materiálu, ktorý zabráni pretokaniu magnetického toku v radiálnom smere rotora a uzatvárať sa iba cez jeden stator a jarmo rotora. Keďže rotorom radiálne neprechádza magnetický tok, môže sa použiť tenší disk. Zámena ocele za ľahší magneticky nevodivý materiál a zároveň tenší rotor znamená menší moment zotrvačnosti a lepšiu dynamiku. Preto sú vhodné na použitie v aplikáciách, kde sa vyžaduje veľká dynamika motora. Aj pri tomto motore podľa prevedenia vinutia statora môže byť stator s drážkami (Obr. 6). alebo bezdrážkový (Obr. 7).



Obr. 6 Diskový motor AFIR s drážkovaným statorom



Obr. 6 Diskový motor AFIR s bezdrážkovým statorom

Výhody jednodiskového axiálneho motora s rotorom uprostred sú:

1. Použitím dvoch vzduchových medzier sily medzi rotorom a statorom sú v rovnováhe. Na rotor menej pôsobia náhle ťahové sily, ktoré ničia ložiská.
2. Vďaka tvaru motora sa stator lepšie chladí.
3. Magnety majú plochý tvar, netreba ich brúsiť ako v prípade radiálnych motorov.
4. Vzduchová medzera je konštrukčne nastaviteľná, čo je v prípade radiálnych motorov takmer nemožné.
5. Stator je plochý a otvorený – ľahko sa navíja.

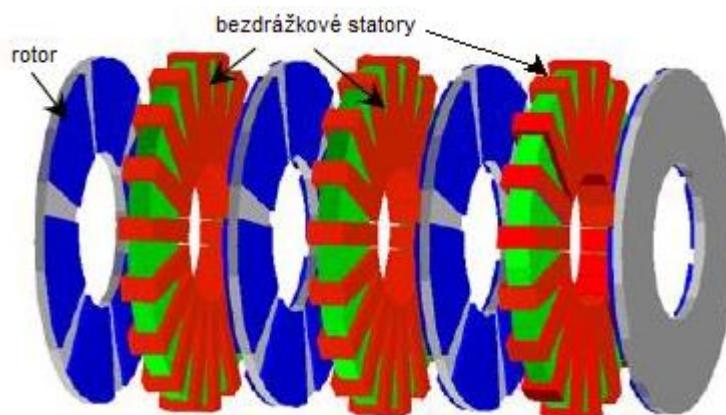
Nevýhody jednodiskového axiálneho motora s rotorom uprostred sú:

1. Ak je veľký rozdiel medzi vonkajším a vnútorným priemerom statora alebo je veľký počet magnetických pólov, je potreba použiť dlhší vodič na vinutie a tým je neefektívne využitý.
2. Závity vo vnútornom polomere statora majú obmedzený objem.
3. Krútiaci moment nie je sínusový, ale ozubený.
4. Statorové lamely musia mať špirálovitý tvar [10].

D. Multidiskový motor

V niektorých aplikáciách je potrebný aj malý priemer diskového motora. Moment jednoduchého diskového motora je funkciou priemeru rotora a malý priemer znamená aj malý moment. V prípade potreby sa moment dá zväčšiť viacdiskovým motorom. Viacdiskový motor je diskový motor rozšírený o ďalšie stupne. Motory sú umiestnené pri k sebe na jeden hriadeľ. Spravidla má byť o jeden statorový alebo rotorový disk viac, no zvyčajne sa používa druhá voľba. V porovnaní s viacerými valcovými motormi spojenými cez spoločný hriadeľ sa diskové motory jednoduchšie konštruujú.

Ako v predošlých motoroch, viacdiskový motor môže mať radiálne alebo axiálne statorové vinutie, s drážkami alebo bez drážok a uloženie magnetov môže byť NN alebo NS. Stator je spravidla tvorený oceľovými plechmi. Magnetický tok sa uzatvára podľa toho, či sú rotory typu NN alebo NS.

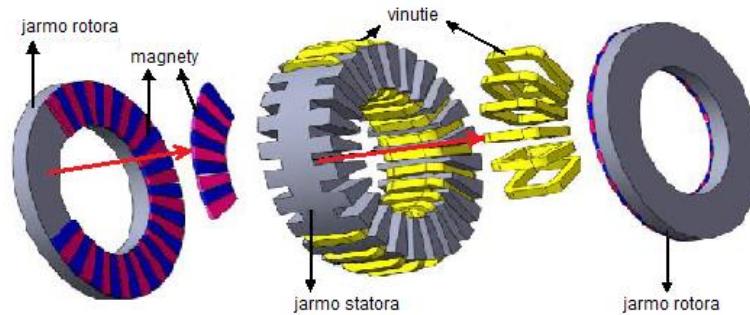


Obr. 7 Axiálne vinutý diskový motor MULTI-NS s bezdrážkovými statormi

IV. NÁVRH TROJ-STATOROVÉHO DISKOVÉHO SYNCHRÓNNEHO MOTORA S PERMANENTNÝMI MAGNETMI

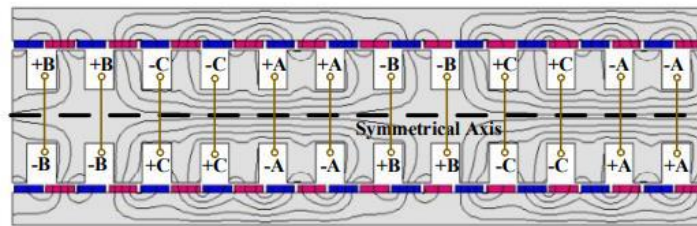
Navrhovanie diskových motorov je pomerne náročná problematika, ktorá nie je ucelene spracovaná v dostupnej literatúre. Väčšinou sú návrhy neúplné a venujú sa len určitej problematike a časti motora.

Troj-statorový motor sme jednoducho získali spojením troch diskových motorov jedným spoločným hriadeľom, pričom motory sú napájané paralelne z jedného spoločného zdroja. Dôležité je, že motory sú navzájom pootočené o 30° a sú oddelené nemagnetickou oceľou, aby krajné motory svojím magnetickým tokom neovplyvňovali prostredný motor. Ďalej je popísaný návrh jedného motora. Motor je typu TORUS s typom rotora NN (Obr. 8).



Obr. 8 Motor typu TORUS s magnetmi na rotore NN

Znázornenie magnetického toku motora je na Obr. 9. Obrázok zobrazuje časový okamih pri rôznej veľkosti napätí vo fázach. Toky majú symetrický tvar vzhľadom na os statora.



Obr. 8 Znázornenie magnetického toku v motore typu TORUS s magnetmi na rotore NN

A. Návrh základných parametrov motora

Pre navrhovaný troj-statorový synchronný motor sa zvolil celkový zdanlivý výkon 3 kVA. Tretina zo zdanlivého výkonu je zdanlivý výkon jedného diskového motora, teda 1kVA. Taktiež bolo určené, že motor bude pracovať s účinníkom 0,8 a podľa teórie z odborných článkov a porovnaní s inými návrhmi motorov s permanentnými magnetmi by sme chceli dosiahnuť účinnosť motora $\eta = 0,87$. K určeniu prúdu motora je potrebné poznať aj veľkosť fázového nominálneho napájacieho napätia. Podľa katalógových údajov iných diskových motorov je zvolená nominálna hodnota fázového napätia 50 V.

Následne boli určené ďalšie základné parametre navrhovaného diskového synchronného motora, ktoré sú uvedené v tabuľke 1.

Tabuľka 1
Ďalšie parametre navrhovaného diskového synchronného motora

Parameter	Veľičina	Hodnota
Činný výkon	P_{IN} [W]	800
Združené napätie	U_S [V]	86,6
Počet fáz motora	m [-]	3
Počet pólov	$2p$ [-]	4
Počet cievok na fázu	n_f [-]	2
Počet drážok statora	n_d [-]	24
Počet magnetov na rotore	n_m [-]	4

B. Návrh základných rozmerov motora a parametrov elektromagnetického obvodu

Na základe doteraz známych parametrov a preštudovaných odborných článkov bol navrhnutý vonkajší priemer statorového disku 0,3 m a šírka medzikružia disku 0,05 m. Z toho je vypočítaný

vnútorný priemer statora 0,2 m. Rovnaké priemery majú aj rotorové disky. Celkový počet drážok na stator a teda aj pólových zubov je 24. Drážky majú jednoduchý tvar a drážky aj teda aj pólové zuby majú rovnakú veľkosť uhla v medzikruží, ktorý je určený ako $360^\circ/24/2 = 7,5^\circ$. Na obr. 9 je zobrazený 3D model drážkovaného disku statora.



Obr. 9 Model drážkovaného disku statora

Veľkosť indukcie vo vzduchovej medzere je zvolená na 0,8 T, aká je pri takýchto typoch motorov bežná. Veľkosť vzduchovej medzery je 0,3 mm. Potom s využitím vytvorených vlastných funkcií v Matlabe s použitím základných rovníc, ktoré sa požívajú pri navrhovaní a výpočtoch parametrov elektromagnetického obvodu motora sa postupne vypočítala magnetická vodivosť, intenzita magnetického poľa, magnetické napätie a magnetický tok vo vzduchovej medzere. Vypočítané hodnoty sa potom postupne optimalizovali zmenou parametrov magnetického obvodu do konečnej podoby.

Pre magnetický obvod statora bola zvolená elektrotechnická oceľ. typu M - 4/120 Carlite. Má nízke straty spôsobené hysteréziou a vírivými prúdmi. Vyrába sa valcovaním za studena. Z magnetizačnej charakteristiky [11] sa odčítali hodnoty intenzity magnetického poľa. pre jednotlivé indukcie v jednotlivých častiach statorového magnetického obvodu a z týchto sa určili príslušné magnetické napätia. Potom sa zasa s využitím vytvorených vlastných funkcií v Matlabe vypočítali magnetické napätia v jarmách obidvoch rotorov a magnetické napätia v permanentných magnetoch. Z celkového magnetického napätia potrebného na vybudenie sa určí počet závitov vo fáze. Na obr. 10 je zobrazený 3D model statora s vinutím a vypočítané parametre elektromagnetického obvodu motora sú uvedené v tabuľke2.



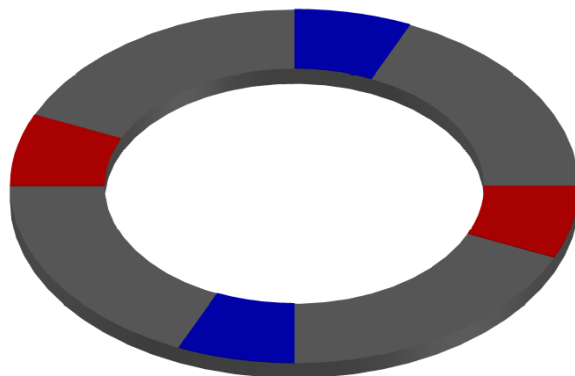
Obr. 10 Model diskového statora s axiálnym vinutím

Tabuľka 2
 Parametre elektromagnetického obvodu motora

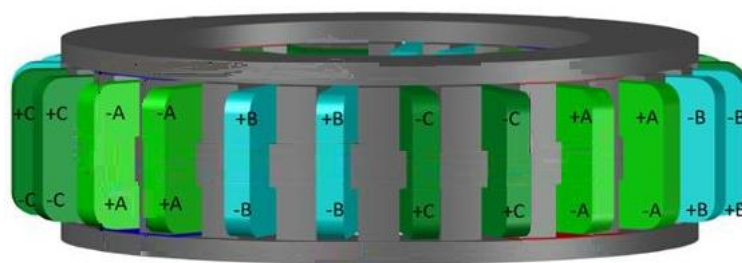
Parameter	Hodnota
Magnetický tok v jarme statora	0,0026 Wb
Magnetická indukcia v jarme statora	1,31 T
Magnetické napätie v jarme statora	4,91 A
Magnetická indukcia v zube statora	1,61 T
Magnetické napätie v zube statora	4,52 A
Magnetická indukcia v jarme rotora	1,28 T
Magnetické napätie v jarme rotora	3,98 A
Magnetické napätie v magnetu rotora	65,2 A
Celkové magnetické napätie	1032,3 A
Fázový prúd	7,65 A
Počet závitov vo fáze	67

Keďže sa jedná o synchronný motor, tak počet magnetov rotora sa volí tak, aby mechanická uhlová rýchlosť rotora bola rovnaká ako elektrická uhlová rýchlosť statora. Z tohto dôvodu na

rotore sú 4 magnety, teda dva kladné magnety voči statoru a dva záporné magnety. Pretože stator má jeden pól tvorený dvomi cievkami, magnet musí plochou pokryť obe cievky, ale nesmie sa prekryvať zároveň aj s inými fázami, tak uhol jedného magnetu je 24° . Umiestnenie magnetov na rotore a ich polarizácia vzhľadom na to, že bola zvolená konfigurácia typu NN je uvedené na obr. 11. Model elektromagnetického obvodu motora s vinutím je zobrazený na obr. 12.

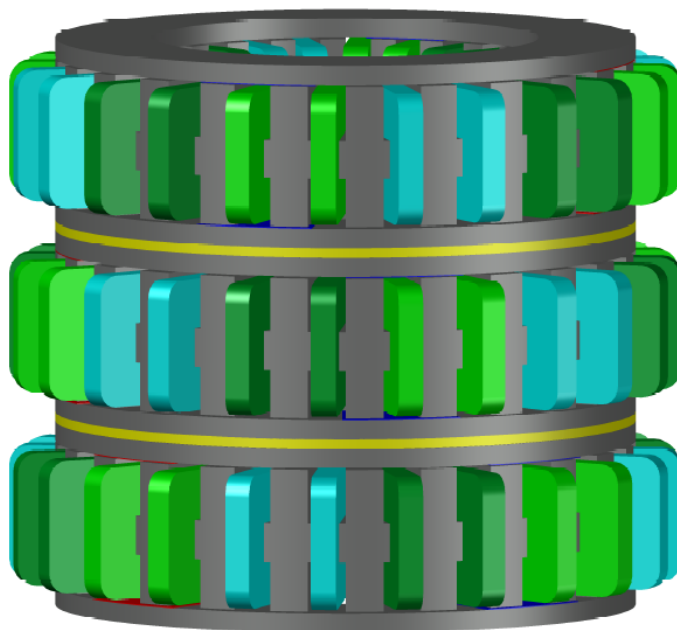


Obr. 11 Model diskového rotora s PM



Obr. 12 Model elektromagnetického obvodu motora s vinutím

Troj-statorový motor dostaneme jednoduchým umiestnením troch jednodiskových motorov na jeden spoločný hriadeľ. Jednotlivé motory budú napájané paralelne. Motory sú ale vzájomne posunuté o 30° , čím sa dosiahne zmenšenie zvlnenia celkového momentu motora. Jednotlivé motory navzájom oddelené nemagnetickou oceľou triedy 316. Tá je na obr. 13 zvýraznená žltou farbou a má hrúbku 5 mm.



Obr. 13 Model troj-statorového elektromagnetického obvodu motora

C. Výpočet strát motora a účinnosti

Pri informatívnom a zjednodušenom výpočte celkových strát v motore (s využitím vytvorených vlastných funkcií v Matlabe) sa počítali Joulove straty vo vinutí, straty v elektromagnetickom obvode motora, mechanické straty a prídavné straty. Prídavné straty, ktoré vznikajú pri zaťažení motora a ich príčinou sú vírivé prúdy s vyššími harmonickými, je pomerne ťažko presne vypočítať, tak sa udávajú odhadom, čo je približne 0,5% z nominálneho výkonu motora. Vypočítané hodnoty strát a účinnosť sú uvedené v tabuľke 3.

Tabuľka 3
Straty motora a účinnosť

Parameter	Hodnota
Joulove straty vo vinutí	18,43 W
Straty v jarme statora	6,02 W
Straty v zuboch statora	12,61 W
Straty v jarme rotora	3,01 W
Mechanické straty	17,42 W
Prídavné straty	4,01 W
Účinnosť motora	0,91

Účinnosť motora je pomerne vysoká, ale celková účinnosť pohonu s týmto motorom bude nižšia, pretože motor bude potrebovať vlastný menič, ktorého straty tu nie sú zahrnuté.

V. ZÁVER

Navrhovanie diskových motorov je pomerne náročná problematika, ktorá nie je ucelene spracovaná v dostupnej literatúre. K dispozícii sú len neúplné návrhy. Pre overenie správnosti vyššie uvedeného návrhu sa v budúcnosti plánuje doplniť ešte aj výpočet momentovej charakteristiky motora a jednoduchý tepelný výpočet.

Na základe v tomto článku prezentovaného návrhu motora je možné urobiť overenie analytického výpočtu a optimalizáciu vypočítaných rozmerov a parametrov motora pomocou elektromagnetickej simulácie vo vhodnom simulačnom programe. Po tejto optimalizácii je možné spracovať aj celkovú konštrukčnú technickú dokumentáciu potrebnú na výrobu motora a po úspešnej výrobe prototypu urobiť overenie jeho funkčnosti a parametrov meraním.

POĎAKOVANIE

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23.

Tento príspevok bol podporený Agentúrou na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0210.

LITERATÚRA

- [1] Poliak, F., Fedák, V., Zboray, L.; Elektrických pohony. Bratislava, 1987, ALFA – Vydavateľstvo technickej a ekonomickej literatúry.
- [2] Rossa, R., Bogatyrev, P.: Wpływ podwojenia liczby faz uzwojenia na wybrane parametry 3-fazowego silnika synchronicznego z magnesami trwałymi na przykładzie silnika napędu eKIT dla pojazdów elektrycznych, *Maszyny Elektryczne – Zeszyty Problemowe*, nr 1, str. 221 -227, 2022.
- [3] Reinboth, H.: Vlastnosti a použití magnetických materiálů. Praha, 1975, SNTL – Nakladatelství technické literatury
- [4] Świerczyński, D., Książek, W., Moradewicz, A. J.: Napęd sześciofazowy z silnikiem PMSM dla pojazdów hybrydowych, *Maszyny Elektryczne – Zeszyty Problemowe*, nr 2, str. 65 -69, 2021.
- [5] El Hajji, T., Hlioui, S., Louf, F. et al.: Optimal Design of High-Speed Electric Machines for Electric Vehicles: A Case Study of 100 kW V-Shaped Interior PMSM. *Machines*, 2023, 11 (1), pp.57. DOI: 10.3390/machines11010057
- [6] Hrabovcová, V. a kol.; Moderné elektrické stroje. Žilina, 2001, EDIS – vydavateľstvo ŽU, IS N 80-7100-809-5
- [7] Axial-flux permanent-magnet machine modeling, design, simulation and analysis, dostupné na: https://www.researchgate.net/publication/228847506_Axial-flux-permanent-magnet-machine-modeling-design-simulation-and-analysis
- [8] Axial flux permanent magnet disc machines: A review, dostupné na: https://www.researchgate.net/publication/228449891_Axial_flux_permanent_magnet_disc_machines_A_review
- [9] Kaňuch, J., Ferková, Ž.: Design and simulation of disk stepper motor with permanent magnets
- [10] Hanselman; D.C.: Brushless Permanent-Magnet Motor Design. USA, 1994, ISBN 0-07-026025-7

Priame momentové riadenie AM

Jaroslava ŽILKOVÁ, Marek PÁSTOR

Katedra elektrotechniky a mechatroniky, Fakulta elektrotechniky a informatiky,
Technická Univerzita v Košiciach, Slovensko

jaroslava.zilkova@tuke.sk, marek.pastor@tuke.sk

Abstrakt — V predložennom príspevku sú charakterizované dva spôsoby priameho momentového riadenia asynchrónneho motora. V závere práce sú uvedené a zhodnotené simulačné výsledky oboch metód získané v programe MATLAB.

Príučové slová — asynchrónny motor, priame momentové riadenie, MATLAB

I. ÚVOD

Pre riadenie dynamicky náročných striedavých pohonov sa dnes používajú riadiace štruktúry vektorového alebo priameho momentového riadenia. K najnovším spôsobom riadenia elektrických pohonov patrí priame momentové riadenie (DTC riadenie) [1-4]. Pri priamom momentovom riadení sa snažíme priblížiť riadeniu jednosmerného motora s cudzím budením, teda oddelene riadiť moment a magnetický tok motora. Na rozdiel od vektorového riadenia [1-4], kde sa riadi moment a magnetický tok prostredníctvom momentotvornej a tokotvornej zložky statorového prúdu, sa pri priamom momentovom riadení riadi moment a magnetický tok motora priamo. Regulácia je vykonávaná pomocou hysteréznych regulátorov. Metóda DTC je orientovaná na statorový tok [1,2].

Poznáme niekoľko metód DTC [1,2,3,4,6], ktoré sa navzájom líšia trajektóriou koncového bodu vektora statorového magnetického toku. Pôvodná Depenbrockova metóda, pri ktorej sa tento koncový bod pohybuje po šesťuholníku bola postupne doplnená o jej ďalšie modifikácie ako sú Takahashiho metóda ako aj rôzne ďalšie vylepšené Takahashiho metódy [3,6].

V predložennom príspevku sa zaoberáme porovnaním oboch typov DTC riadenia asynchrónneho motora s kotvou nakrátko v programe MATLAB a porovnaním dosiahnutých simulačných výsledkov.

II. PRIAME MOMENTOVÉ RIADENIE

Priame riadenie toku a momentu predstavuje najnovšiu metódu riadenia elektrických pohonov, ktorú po prvý raz publikoval Prof. Manfred Depenbrock z Univerzity v Bochumi, v roku 1985. Túto metódu navrhol špeciálne pre vysoko výkonné trakčné pohony, pretože dokáže pracovať s nízkou spínacou frekvenciou striedača.

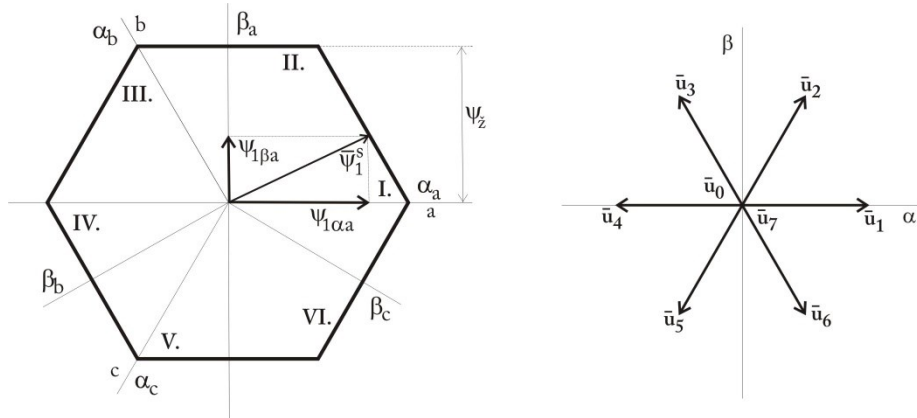
Priame momentové riadenie podľa Prof. Depenbrocka sa stalo základom pre rozličné úpravy a vylepšenia [4]. Pravdepodobne najrozšírenejšou je metóda, ktorú v roku 1986 publikovala dvojica japonských výskumníkov Isao Takahashi a Toshihiko Noguchi. Táto metóda zároveň zachováva všetky ostatné výhody pôvodného priameho momentového riadenia, predovšetkým jeho jednoduchosť a výbornú momentovú dynamiku [2]. Takahashiho metóda priameho momentového riadenia pracuje s vysokou spínacou frekvenciou striedača, preto je vhodná skôr pre priemyselné pohony a preto jej v ďalších častiach príspevku budeme venovať väčšiu pozornosť. Okrem týchto dvoch základných metód existujú aj rôzne ďalšie modifikácie Takahashiho metódy [2]. V nasledujúcich častiach budú stručne charakterizované obe spomínané metódy DTC.

A. Depenbrockova metóda DTC

Depenbrockova metóda pracuje s transformáciou súradníc do osí α a β . Pri tejto metóde sa nereguluje vektor statorového prúdu, ale priamo moment vo vybranom tolerančnom pásme a priebeh priestorového vektora magnetického toku. Tieto veličiny sú pomocou matematického modelu AM na základe meraných veličín vyhodnocované. U DTC riadenia je regulácia

orientovaná na satorový magnetický tok. Zmena satorového magnetického toku sa pri zanedbaní vplyvu satorového odporu rovná vektoru satorového napätia. Preto je možné trajektóriu koncového bodu vektora satorového toku určovať vhodnou postupnosťou vektora napätia striedača [1-4].

Depenbrockova metóda DTC je charakteristická tým, že točivé magnetické pole statora je vytvárané riadením vektora magnetického toku statora tak, že sa koncový bod vektora magnetického toku pohybuje po šesťuholníku. Spôsob ako dosiahnuť požadovaný pohyb vektora magnetického toku za pomoci šiestich vektorov napätia u_1 až u_6 je znázornený na Obr. 1. Sú tu zakreslené tri komplexné roviny $[\alpha, \beta]$, ktorých reálne osi sú totožné s osami a, b, c, ktoré znázorňujú priestorové rozloženie fázových vinutí v statore asynchrónneho motora [3].



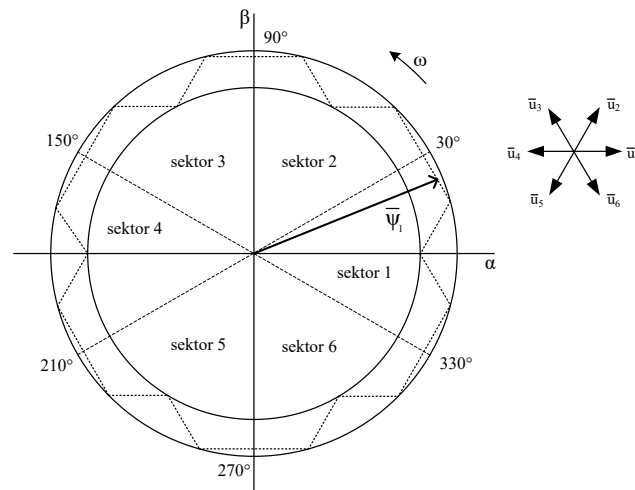
Obr. 1 Trajektória koncového bodu vektora satorového magnetického toku podľa Depenbrockovej metódy DTC.

Na Obr.1 je uvedených 6 sektorov označených číslicami I až VI, ktorých hranice tvoria pomyselné polpriamky prechádzajúce vrcholmi a končiace v strede šesťuholníka. Šesťuholník predstavuje trajektóriu koncového bodu satorového magnetického toku.

Táto metóda, ktorá bola špeciálne navrhnutá pre vysoko výkonné trakčné pohony, keďže dokáže pracovať s nízkou spínacou frekvenciou sa stala základom pre ďalšie úpravy a vylepšenia [1,2,4].

B. Takahashiho metóda DTC

Takahashiho metóda priameho riadenia momentu je charakteristická pohybom koncového bodu vektora satorového toku v medzikruží, pričom v zjednodušenom prípade sa trajektória podobá kružnici, ktorá je zobrazená na Obr.2. Táto regulácia je dvojhodnotová, určovaná dvomi kružnicami, v ktorých sa má veľkosť satorového toku pohybovať. Ak vektor satorového toku dosiahol vnútornú kružnicu, je zvolené napätie, ktoré zabezpečí zväčšenie satorového toku, ak dosiahne vonkajšiu kružnicu, bude vybrané napätie, ktoré zmenší satorový tok. [2].

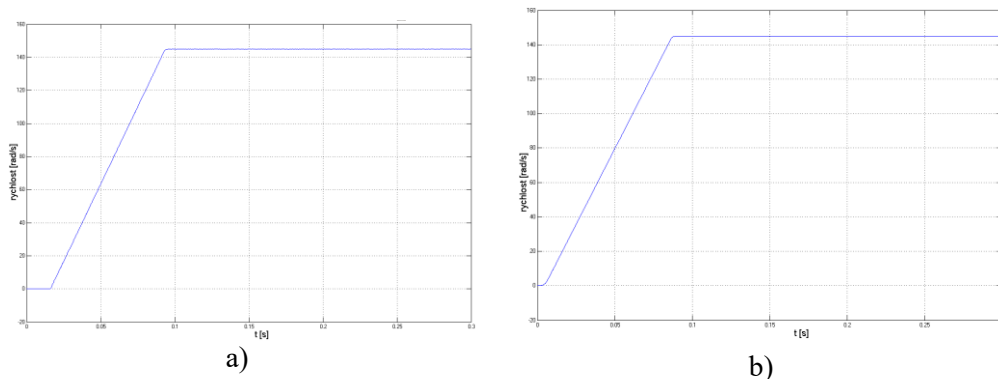


Obr. 2 Trajektória koncového bodu vektora satorového magnetického toku podľa Takahashiho metódy DTC.

III. VÝSLEDKY POČÍTAČOVÝCH SIMULÁCIÍ

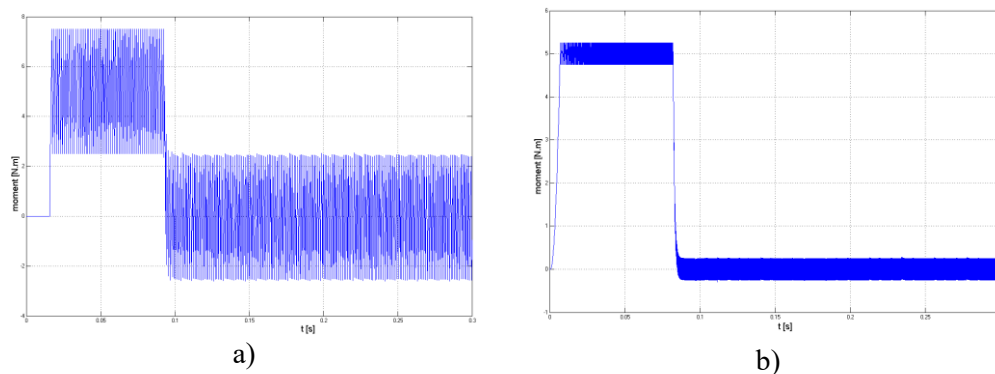
V nasledujúcej časti sú uvedené výsledky simulácií priameho momentového riadenia asynchrónneho motora pomocou Depenbrockovej a Takahashiho metódy.

Na Obr. 5 sú uvedené výsledky získané pri rozbehu AM z nulovej na nominálnu uhlovú rýchlosť. V časti a) je uvedený priebeh podľa Depenbrockovej metódy a v časti b) podľa Takahashiho metódy.

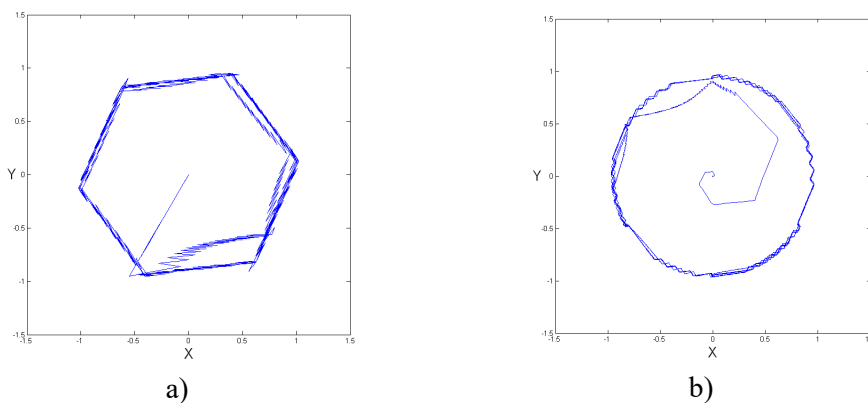


Obr. 5 Priebeh uhlovej rýchlosti AM s DTC riadením pri Depenbrockovej (a) a Takahashiho metóde (b).

Výraznejšie rozdiely medzi metódami predstavuje priebeh momentu. Na Obr. 6 a) je uvedený priebeh momentu pri Depenbrockovej metóde a na Obr.6 b) Takahashiho metóde DTC. Z priebehov je zrejmé, že moment pri Depenbrockovej metóde je výrazne zakmitanejší. Rýchlosť s akou nastala zmena momentu potvrdzuje výbornú dynamiku riadenia momentu pri oboch metódach.



Obr.6 Priebeh momentu AM s DTC riadením pri Depenbrockovej (a) a Takahashiho metóde (b).



Obr. 7 Priebeh koncového bodu vektora magnetického statorového toku v súradniciach X – Y pri Depenbrockovej (a) a Takahashiho metóde (b).

Z priebehov koncového bodu vektora statorového toku uvedených na Obr.7 je zrejmé, že potvrdzujú teoretické predpoklady.

IV. ZÁVER

Pre riadenie dynamicky náročných priemyselných striedavých pohonov sa dnes používajú riadiace štruktúry vektorového alebo priameho momentového riadenia. Metóda priameho riadenia momentu motora, ktorej je venovaný tento príspevok, vychádza podobne ako metóda vektorového riadenia z nezávislého riadenia magnetického toku a momentu motora. Vhodným výberom napäťových vektorov je možné zabezpečiť požadovanú trajektóriu statorového toku, čiže veľkosť a polohu statorového toku a dosiahnuť požadovaný moment motora. V predloženom príspevku je charakterizovaná Depenbrockova a Takahashiho metóda priameho riadenia momentu. Získané simulačné výsledky uvedené v príspevku potvrdili teoretické predpoklady.

Obe metódy DTC poskytujú veľmi rýchlu časovú ozvu, nižšie spínacie straty a straty od vyšších harmonických v porovnaní s klasickým vektorovým riadením využívajúcim PWM. Priebeh momentu pri Depenbrockovej metóde bol výraznejšie zakmitaný, pričom obe metódy preukázali výbornú dynamiku riadenia momentu.

Nakoľko Takahashiho metóda je vhodná pre nasadenie v oblasti priemyselných pohonov, jej opisu bola venovaná väčšia pozornosť.

POĎAKOVANIE

Tento príspevok vznikol s podporou Vedeckej grantovej agentúry Ministerstva školstva, vedy, výskumu a športu Slovenskej republiky a Slovenskej akadémie vied v rámci projektu VEGA 1/0363/23.

LITERATÚRA

- [1] K. Zeman, Z. Peroutka, M. Janda, M.,": Automatická regulace pohonů s asynchronními motory". ZČU Plzeň, 2004.
- [2] E. Jozefák. "DTC riadenie AM s predikčným filtrom." Doktorandská dizertačná práca, Žilina, 1998
- [3] P. Brandšteter.,: "Střídavé regulačné pohony." VŠB TU Ostrava, 1999.
- [4] L. Zboray,.. F. Ďurovský, J.Tomko., "Regulované pohony". Košice : Viena, 2000. .
- [5] P. Tiitinen, P. Pohjalainen, J. Lalu, Príme řízení momentu. Nová generace v řízení asynchronních motoru . Elektro, 1996, 4, 128 – 132.
- [6] J. Javurek. "Regulace moderních elektrických pohonu". Praha, Grada, 2003.

Hodnotenie dôveryhodnosti dát pohybových aktivít študentov

¹Daniel Kolibár, ²Filip Gurbál

^{1,2}Katedra počítačov a informatiky, Fakulta elektrotechniky a informatiky Technickej univerzity v Košiciach, Slovensko

¹daniel.kolibar@student.tuke.sk, ²filip.gurbal@tuke.sk

Abstrakt – Článok opisuje implementáciu systému pre spracovanie dát pohybových aktivít študentov. Súčasťou systému je posudzovanie dôveryhodnosti záznamov pohybových aktivít za účelom zvýšenia objektivity hodnotenia. Posúdenie dôveryhodnosti – alebo detekcia anomálií – je realizovaná využitím modelu umelej neurónovej siete typu autoenkoder. Systém dokáže odhaliť očividne upravené GPX záznamy, avšak neodhalí záznamy vygenerované sofistikovaným nástrojom ako je gpx.studio.

Kľúčové slová – spracovanie dát, detekcia anomálií, dištančná telesná výchova, strojové učenie

I. ÚVOD

Príchod pandémie ochorenia COVID-19 v roku 2020 priniesol do našich životov množstvo zmien. Pandémia okrem iných oblastí nášho života značne zasiahla aj oblasť školstva. Z dôvodu rôznych epidemiologických opatrení, ktorých cieľom bolo zabránenie prenosu a šírenia vírusu, sa väčšina vzdelávacích procesov musela upraviť do dištančnej podoby, napríklad presunutím kompletne do online priestoru. Jedným z predmetov, ktorého sa tieto opatrenia dotkli je aj telesná výchova, ďalej TV.

Riešením pre vedenie dištančnej formy TV mohlo byť použitie rôznych aplikácií pre záznam športových aktivít. Tie nie sú žiadnou novinkou a tak ako využívame iné aplikácie v našom profesijnom, či osobnom živote, tieto pomocou rôznych *smart* doplnkov a GPS nám dokážu zaznamenať priebeh našej športovej aktivity. Ich výsledkom sú dáta, ktoré vedú poskytnúť pohľad do športového vyžitia jednotlivca. Avšak vedenie TV použitím týchto aplikácií nie je dokonalé a prináša niekoľko problémov.

Hlavným problémom, ktorému sa v tomto článku venujeme a na ktorý netreba pri dištančnej forme TV zabudnúť je objektivita hodnotenia. Ak hodnotíme študenta len podľa záznamov aktivít, nijako nezabránilme tomu, aby tieto aktivity neboli „falošné”. Väčšina súborov obsahujúcich záznam športovej aktivity je vo formáte GPX¹, resp. TCX². Tento formát patrí do skupiny značkovacieho jazyka XML³, ktorý dokáže používateľ jednoducho upraviť pomocou akéhokoľvek textového editora. S problémom „falošných” aktivít sa stretli aj na OAS TUKE. Študenti nahrávali upravované, alebo cudzie záznamy aktivít.

II. DETEKCIA ANOMALIÍ V DÁTACH POHYBOVÝCH AKTIVÍT ŠTUDENTOV

Pre zvýšenie objektivity hodnotenia dokážeme využiť techniky strojového učenia pre odhalenie aktivít, ktoré budú predstavovať nejakú odchýlku od zvyšných a môžu vyučujúceho upozorniť, aby konkrétnemu záznamu aktivity venoval pozornosť. Konkrétne nás bude zaujímať skupina techník slúžiacich na detekciu anomálií.

Záznamy pohybových aktivít študentov sú ukladané vo formáte GPX, resp. TCX, pričom ide o typ jazyka XML. Samotná trasa záznamu pohybovej aktivity je uložená pod značkou `trk` ako zoznam prvkov značky `trkpt`, pričom každý tento prvok obsahuje zemepisné súradnice, nadmorskú výšku a časovú značku vytvorenia (viď výpis 1). Pri pohľade na štruktúru formátu GPX si môžeme všimnúť, že upraviť záznam pohybovej aktivity nebude vyžadovať žiadne špeciálne nástroje či schopnosti. Dokonca nie je problém takýto záznam aj kompletne vytvoriť. K úprave

¹GPX – GPS Exchange Format

²TCX – Training Center XML

³XML – eXtensible Markup Language

```
<gpx version="1.1" ... >
  <trk>
    ...
    <trkseg>
      <trkpt lat="48.7159190" lon="21.2378510">
        <ele>244.9</ele>
        <time>2022-03-27T11:49:41Z</time>
      </trkpt>
      <trkpt lat="48.7159150" lon="21.2378530">
        <ele>244.9</ele>
        <time>2022-03-27T11:49:47Z</time>
      </trkpt>
    ...
  </trkseg>
</trk>
</gpx>
```

Listing 1 Ukážka obsahu záznamu aktivity vo formáte GPX

nám môže pomôcť aj množstvo rôznych nástrojov, ako napríklad webová aplikácia s príznačným názvom: *gpx.studio*⁴, či aplikácia *Viking*⁵.

Pre zistenie, či záznam pohybovej aktivity nebol študentom manuálne upravený, môžeme využiť metódy hlbokého strojového učenia [1], pri ktorých sa často hovorí o umelých neurónových sieťach. Manuálny zásah v záznamoch môžeme označiť ako anomáliu, ktorá predstavuje nejakú odchýlku od normy. Poznáme tri základne prístupy detekcie anomálií. Prvým je detekcia na základe vzdialenosti, druhým na základe hustoty a posledným, tretím princípom je určenie na základe nejakého hodnotenia [2].

Práca [3] sa venuje detekcii falošnej pohybovej aktivity pomocou neurónových sietí určených pre klasifikáciu zo softvérovej knižnice *TensorFlow*⁶. Práca sa ale venuje primárne cvičeniam ako sú drepy, či skok a na záznam týchto aktivít používajú mobilnú aplikáciu, ktorá zaznamenáva dáta využitím akcelerometra a gyroskopu.

Ďalšia súvisiaca práca [4] sa venuje klasifikácii druhu pohybových aktivít na základe GPS záznamov trás. Pre fungovanie ich klasifikátora boli najpodstatnejšími atribútmi čas, rýchlosť a priestorové vzťahy medzi GPS bodmi. Ich výsledkom bolo niekoľko modelov. Prvý model využíval jednoduché pravidlá, kde hlavným identifikátorom cesty automobilom bola logicky vyššia rýchlosť, konkrétne išlo o medián rýchlosti vyšší ako 5 km/h . Druhý model, pre nás zaujímavejší, využíval strojové učenie, konkrétne klasifikáciu pomocou techniky *Random Forest*.

Ak nazrieme do oblasti umelých neurónových sietí, naskytne sa nám hneď jedna z možných implementácií detektorov anomálií, a to použitie *Autoenkódera*. Autoenkóder je druh umelej neurónovej siete, so schopnosťou znižovania dimenzionality dát. Pri znižovaní dimenzionality predpokladáme, že všetky dáta v konkrétnej dátovej sade budú medzi sebou nejakou korelovať. Toto nám zaručí, že po redukcii dimenzionality budú dáta, ktoré sú normálne, a dáta, ktoré predstavujú odchýlku, viditeľne odlíšené.

Pri tréňovaní Autoenkóderov vieme využiť metódu učenia bez učiteľa, pričom našim hlavným cieľom tréňovania bude zabezpečenie úspešnej reprodukcie vstupných dát na výstupe [5]. Jednoducho povedané, cieľom Autoenkódera je predikcia hodnoty vlastného vstupu. Po dokončení tréňovania by sme mali vedieť detegovať anomálie tak, že ak dáta na vstupe budú predstavovať anomáliu, Autoenkóder nebude schopný tieto dáta úspešne reprodukovať na výstupe. Toto zvýši hodnotu reprodukčnej chyby, ktorú vieme napríklad určiť pomocou štatistickej metódy reziduálneho súčtu štvorcov (Error Sum of Squares):

$$SSE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (1)$$

Pričom Y_i je i -tá hodnota vstupných dát a \hat{Y}_i výstupná hodnota [6].

Prvotnou podmienkou pre zaručenie vysokej presnosti výstupov nášho modelu implementujúceho Autoenkóder, či prakticky akýkoľvek iný druh neurónovej siete, je potrebné ešte pred samotným tréňovaním pripraviť kvalitnú dátovú sadu. V prípade detekcie anomálií a učenia bez učiteľa, by mala naša dátová sada reprezentovať vzorku dát, ktorých väčšina nebude svojimi hodnotami patriť medzi anomálie. Takto budeme mať istotu, že sa náš model naučí "črty" hodnôt nereprezentujúcich odchýlku od normy. Okrem toho je ďalším dôležitým faktorom množstvo dát v dátovej sade.

⁴<https://gpx.studio/>

⁵<https://sourceforge.net/projects/viking/>

⁶<https://www.tensorflow.org/>

Aj keď vieme označiť umelé neuronové siete ako univerzálny nástroj schopný predikovať akékoľvek merateľné hodnoty, pričom na rozdiel od štandardných regresných modelov si dokážu umelé neuronové siete udržať schopnosť nájdenia vzťahov medzi jednotlivými hodnotami dát, je pre ich tréning potrebná bohatá, resp. obrovská tréningová sada [7]. Získanie takejto dátovej sady môže byť v niektorých témach problém. Ak sa tentokrát zamyslíme, čo by predstavovalo dátovú sadu pre riešenie detekcie podozrivých pohybových aktivít v našom systéme, uvedomíme si, že by to mali byť samotné pohybové aktivity, resp. ich GPS záznam. Tento fakt môže predstavovať pre úspešné riešenie tejto funkcionality riziko, a to z dôvodu problematického získavania takéhoto druhu dátovej sady. Pri prehľadávaní verejných databáz za pomoci služby *Dataset Search*⁷ zistíme, že väčšina dátových sád je primárna, alebo je vytvorená jednotlivcami pre jeden konkrétny projekt. Najzaujímavejším výsledkom vyhľadávania bola dátová sada, ktorá je súčasťou technickej správy [8].

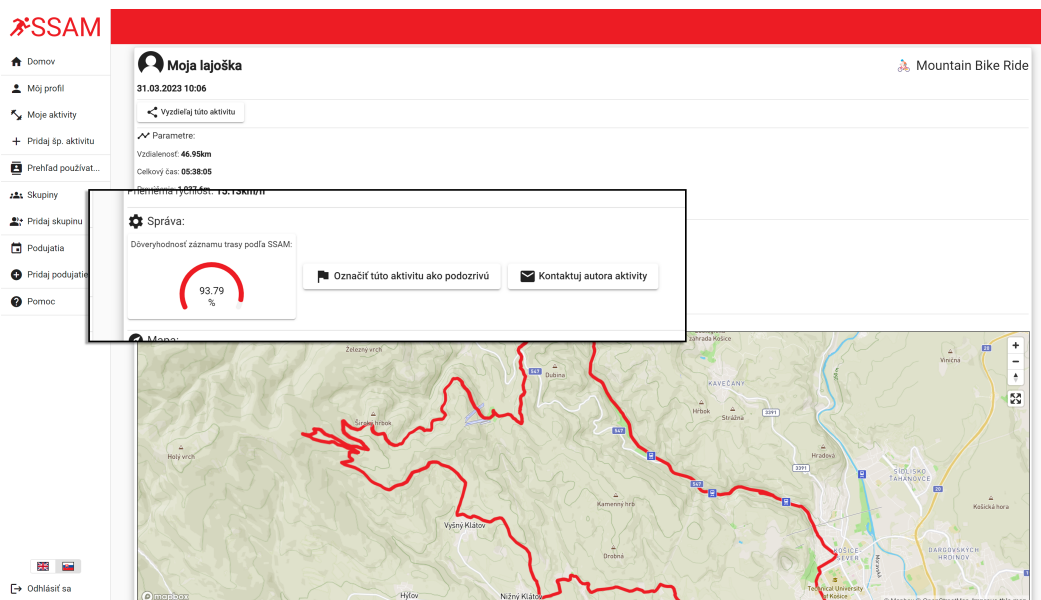
III. SYSTÉM PRE SPRACOVANIE DÁT POHYBOVÝCH AKTIVÍT ŠTUDENTOV

Nami navrhnutý systém – SSAM – sa snaží vyriešiť problémy dištančnej formy TV a tak zjednoduší prácu vyučujúcim a aj študentom. Pomocou ďalších doplnkových funkcionalít tak predstavuje riešenie, ktoré by mohlo dovoliť zaviesť dištančnú formu TV aj do bežnej praxe. Používatelia, čiže študenti alebo vyučujúci, so systémom pracujú pomocou grafického používateľského rozhrania vo forme webovej aplikácie.

Systém ponúka možnosť automatického zberu dát z aplikácie Strava, ale taktiež aj manuálne pridávanie aktivít pomocou GPX záznamov pre používateľov. Systém počíta odtlačok tohto záznamu aby predišiel opakovanému nahraniu rovnakej aktivity. Navyše, systém ponúka správčovský systém pre vyučujúcich, kde si môžu vytvárať skupiny študentov pre lepší prehľad ich hodnotenia.

Systém zabezpečuje výpočet rôznych štatistík. Počíta napríklad priemerné parametre aktivít jednotlivých používateľov za určité časové obdobie. Ponúka taktiež štatistiky o druhu vykonávaných športov alebo štatistiky pre jednotlivé skupiny, ktoré vyučujúci vytvorili. Vďaka týmto štatistikám majú potom vyučujúci jasný prehľad o športovom vyžití študentov.

Systém rieši problém objektivity hodnotenia počas dištančnej formy TV. Je realizovaná výpočtom dôveryhodnosti záznamu pohybovej aktivity, ktorý je postavený na umelej neurónovej sieti implementujúcej autoenkóder. Výsledkom je skóre, na základe ktorého sa vyučujúci môže rozhodnúť aktivitu označiť za podozrivú. Takto označená aktivita sa potom nepočíta do štatistík študenta. Ukážka rozhrania je zobrazená na obrázku 1.

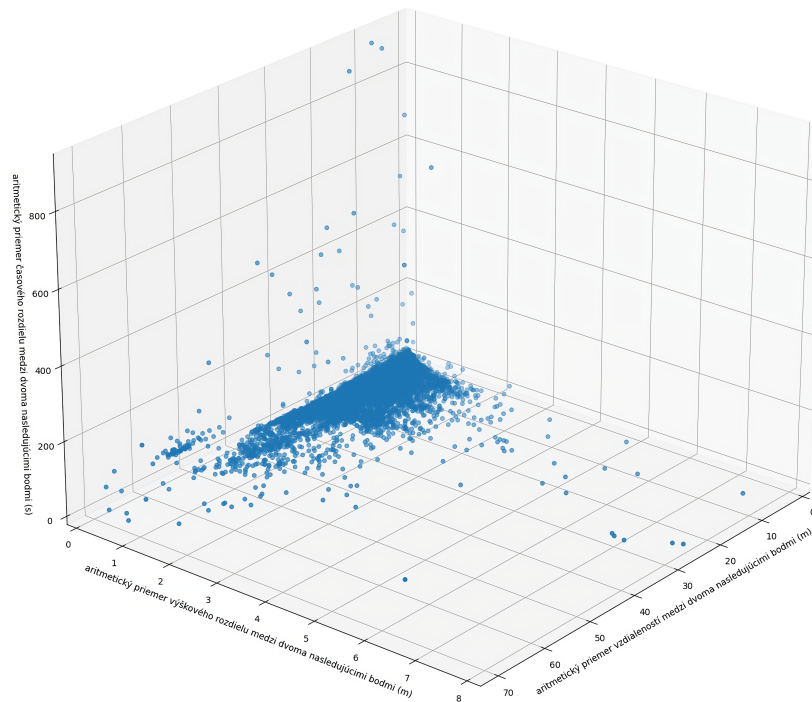


Obr. 1 Hodnota dôveryhodnosti záznamu pohybovej aktivity a ďalšie funkcie na obrazovke detailu aktivity v rozhraní vyučujúceho

IV. MODEL PRE OVERENIE DÔVERYHODNOSTI ZÁZNAMU GPX

Autoenkóder využíva model, ktorý bol natrénovaný pomocou dátovej sady vytvorenej z dát [8] a verejne dostupných dát aktivít používateľov aplikácie *Strava*. Výsledkom nášho snaženia bolo

⁷<https://datasetsearch.research.google.com/>



Obr. 2 Graf zobrazujúci hodnoty troch skúmaných vlastností

celkovo 10.76GB dát vo forme GPX súborov obsahujúcich záznamy pohybových aktivít pričom sme predpokladali, že väčšina záznamov bude predstavovať vzorku nepatriacu medzi anomálie.

Pre určenie dôveryhodnosti daného záznamu sme sa rozhodli neanalyzovať výkonnostne parametre autora pohybovej aktivity, ale samostatný záznam trasy. Skúmané vlastnosti, tzv. *features*, ktoré sme zvolili ako vstup do nášho modelu sú nasledovné:

- 1) aritmetický priemer vzdialeností medzi dvoma za sebou nasledujúcimi hodnotami,
- 2) aritmetický priemer výškového rozdielu medzi dvoma za sebou nasledujúcimi hodnotami,
- 3) aritmetický priemer časového rozdielu medzi dvoma za sebou nasledujúcimi hodnotami,
- 4) štandardná odchýlka vzdialeností medzi dvoma za sebou nasledujúcimi hodnotami,
- 5) štandardná odchýlka výškového rozdielu medzi dvoma za sebou nasledujúcimi hodnotami,
- 6) štandardná odchýlka časového rozdielu medzi dvoma za sebou nasledujúcimi hodnotami,
- 7) štandardná odchýlka rýchlosti medzi dvoma za sebou nasledujúcimi hodnotami, a
- 8) odtlačok hodnoty atribútu `creator` z GPX súboru

Na grafe 2 sme si manuálne zobrazili hodnoty troch z týchto ôsmich skúmaných vlastností. Graf pekne potvrdzuje náš predpoklad o množstve odlišných dát reprezentujúcich odchýlku.

Po normalizovaní sme mali k dispozícii 17837 vzoriek pre natrénovanie modelu umelej neurónovej siete. Model sa skladá zo 4 vrstiev, funkcií pre lineárnu transformáciu: $y = xA^T + b$, kde x sú vstupné dáta, y transformované výstupné dáta, A váhy prepojenia a b prahová hodnota. Dve z vrstiev slúžia na enkódovanie vstupných dát a zvyšné dve na dekódovanie výstupu. Výstup prvých troch vrstiev je zároveň ešte vstupom do aktivačnej funkcie ReLU (viď vzorec 2) a výstup poslednej vrstvy vstupuje do aktivačnej nelineárnej funkcie sigmoid (viď vzorec 3). Pôvodne sme v našej sieti používali funkciu hyperbolického tangensu (viď vzorec 4), ktorú sme nahradili ReLU z dôvodu rýchlejšej konvergencie [9] a to bez straty presnosti nášho modelu. Funkcia sigmoid v poslednej vrstve nám zabezpečí výstupnú hodnotu v intervale $[0, 1]$. Pri procese tréningu využívame optimalizačný algoritmus Adam⁸.

$$\text{Relu}(x) = \max(0, x) \quad (2)$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (4)$$

⁸<https://pytorch.org/docs/stable/generated/torch.optim.Adam.html>

Tabuľka 1
Výsledné hodnoty SSE testovacích záznamov

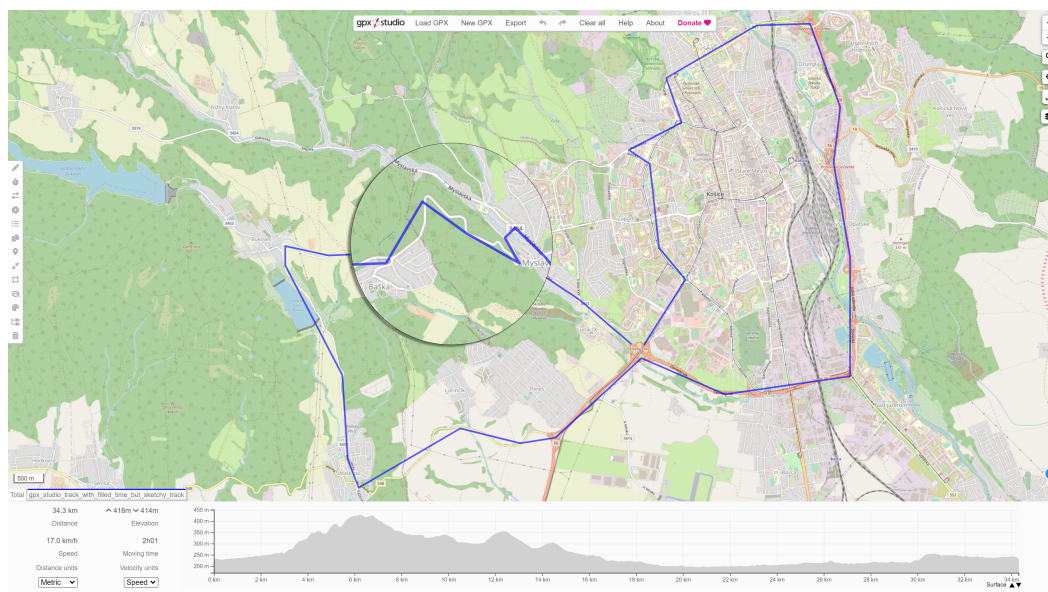
#	Opis záznamu	Výsledok SSE
1	Štandardný záznam z mikroslužby pre import dát zo <i>Stravy</i>	0.0004
2	Záznam vytvorený online nástrojom <i>gpx.studio</i> s nereálnymi zmenami prevýšenia a s umelo vypočítanou konštantnou rýchlosťou	1.2042
3	Záznam vytvorený online nástrojom <i>gpx.studio</i> s normálnou trasou a s umelo vypočítanou konštantnou rýchlosťou	0.0057
4	Záznam vytvorený online nástrojom <i>gpx.studio</i> s normálnou trasou	100.0000
5	Záznam vytvorený online nástrojom <i>gpx.studio</i> s prerušovanou trasou a s umelo vypočítanou konštantnou rýchlosťou	0.0372
6	Záznam vytvorený spojením reálneho záznamu s trasou vytvorenou online nástrojom <i>gpx.studio</i> s umelo vypočítanou konštantnou rýchlosťou	0.0001
7	Záznam vytvorený online nástrojom <i>gpx.studio</i> s normálnou trasou a s privysokou umelo vypočítanou konštantnou rýchlosťou	0.0012
8	Záznam vytvorený online nástrojom <i>mapy.cz</i>	100.0000

V. VYHODNOTENIE

Pre overenie presnosti s akou táto funkcia resp. náš natrénovaný model dôveryhodnosť určuje sme vytvorili 8 validačných testov, ktoré do tohto modelu vkladali nami umelo vytvorené záznamy reprezentujúce hlavne vzorku nedôveryhodných záznamov.

Testovacie dáta, na ktorých sme model otestovali sa skladali zo siedmych trás, ktoré sme umelo vytvorili cez online nástroje *mapy.cz*⁹ a *gpx.studio*¹⁰. Ôsma trasa bola vytvorená našou mikroslužbou pre import dát zo *Stravy*. V tabuľke 1 vidíme hodnoty SSE pre danú testovaciu trasu.

Hodnotenie SSE je pre každú trasu uvedené v tabuľke 1. Výsledky jasne naznačujú, že náš model vie bez problémov rozoznať ručne vytvorené záznamy, ktoré sú bez špeciálnych úprav. Skvelé taktiež je, že pre záznam číslo 2, ktorý obsahuje informácie o rýchlosti, ale parameter prevýšenia je nereálny, model určil taktiež vysokú hodnotu SSE. Vysoké SSE model určil aj pri zázname číslo 5, kde sme vytvorili záznam tak, že sme v nástroji *gpx.studio* nepoužili automatické trasovanie, čiže výsledná trasa bola prerušovaná. Pre lepšiu predstavu sa pozrime na obrázok 3, na ktorom môžeme vidieť že takáto trasa neobsahuje plynulé prechody.



Obr. 3 Prerušovaná trasa v online nástroji *gpx.studio*

Bohužiaľ, pri záznamoch 3, 6 a 7 sme narazili na nedostatky nášho modelu. Problémom je, že všetky tieto tri záznamy obsahujú rovnaké dáta ako štandardný záznam. Z toho dôvodu náš model nedokáže detegovať žiadne odlišnosti v dátach. Nastáva ale otázka, či je technicky vôbec

⁹<https://mapy.cz>

¹⁰<https://gpx.studio>

možné vytvoriť model, ktorý by dokázal detegovať tieto „dokonalé“ falošné záznamy. Aktuálne náš model rozoznáva len 8 vlastností záznamu. Napríklad, ak by sme model rozšírili o vlastnosti reprezentujúce parametre samotnej aktivity (priemerná rýchlosť, prejdená vzdialenosť a pod.) a taktiež rozlišovali druh aktivity, dokázal by náš model pravdepodobne určiť vysoké SSE aj pre testovací záznam číslo 7. Zväčšenie tréningovej sady by s presnosťou modelu pomohlo tiež.

VI. ZÁVER

Výsledkom tejto práce je tak systém, ktorý dokáže zjednodušiť priebeh dištančnej formy a organizovanie virtuálnych športových podujatí. OAŠ TUKE prejavilo o tento systém záujem a vie si predstaviť jeho využitie v praxi. V tomto článku sme opísali funkcionalitu posudzovania dôveryhodnosti dát športových aktivít nahrávaných študentami. Posúdenie dôveryhodnosti bolo realizované využitím modelu umelej neurónovej siete typu autoenkoder. Systém hodnotí každú športovú aktivitu vo formáte GPX hodnotiacim číslom SSE, podľa ktorého je možné určiť, či ide o očividne upravený GPX záznam a možno ho označiť za podozrivý. Systém však neodhalí záznamy vygenerované sofistikovaným nástrojom ako je gpx.studio.

LITERATÚRA

- [1] J. Sevilla, L. Heim, A. Ho, T. Besiroglu, M. Hobbhahn, and P. Villalobos, “Compute trends across three eras of machine learning,” *arXiv preprint arXiv:2202.05924*, 2022.
- [2] K. G. Mehrotra, C. K. Mohan, and H. Huang, *Anomaly detection principles and algorithms*. Springer, 2017, vol. 1.
- [3] E. Kock and Y. Sarwari, “How can machine learning help identify cheating behaviours in physical activity-based mobile applications?” 2020.
- [4] J. Wu, C. Jiang, D. Houston, D. Baker, and R. Delfino, “Automated time activity classification based on global positioning system (gps) tracking data,” *Environmental Health*, vol. 10, no. 1, pp. 1–13, 2011.
- [5] M. Sakurada and T. Yairi, “Anomaly detection using autoencoders with nonlinear dimensionality reduction,” in *Proceedings of the MLSDA 2014 2nd workshop on machine learning for sensory data analysis*, 2014, pp. 4–11.
- [6] S. Nesvadbová, “Odhady parametrů regresní přímky,” Ph.D. dissertation, Masarykova univerzita, Přírodovědecká fakulta, 2017.
- [7] T. Shaikhina and N. A. Khovanova, “Handling limited datasets with neural networks in medical applications: A small-data approach,” *Artificial intelligence in medicine*, vol. 75, pp. 51–63, 2017.
- [8] I. F. Jr., S. Rauter, D. Fister, and I. Fister, “A collection of sport activity datasets for data analysis and data mining 2017a,” *Technical report 2017a*.
- [9] B. Ding, H. Qian, and J. Zhou, “Activation functions and their characteristics in deep neural networks,” in *2018 Chinese control and decision conference (CCDC)*. IEEE, 2018, pp. 1836–1841.

Author's index

- A**
Adam Tomáš 102, 107
Andraš Miroslav 248
Andrejčík Samuel 57, 62, 79, 85, 149, 153
- B**
Babič František 102, 107
Baláž Anton 221, 229, 237, 272
Basarik Tomáš 460
Baumgartner Maroš 317, 323, 329
Bačík Ján 314
Belej Marek 272
Biceková Anna 134, 145
Bodnár Dávid 278
Bujňák Ján 62
- C**
Cymbalák Dávid 7, 12, 17, 22
- Č**
Čubiňáková Alena 145
- D**
Dimun Sebastián 124
Dobranská Lenka 102
Docí Jozef 107
Dzurovčin Jakub 418
Džmura Jaroslav 140
- Ď**
Ďurča Daniel 356
Ďurovský František 278, 305
- F**
Fedor Marek 259
Fedor Pavol 298, 314
- G**
Gans Šimon 340
Genčí Ján 89
Giňovský Ján 79
Girovský Peter 282, 288, 293, 362, 379, 464
Gladišová Iveta 112
Gordan Daniel 460
Gurbál Filip 478
- H**
Haborák Kristián 182
Haluška Renát 384, 389, 394, 408
Harahus Maroš 399, 414
Hládek Daniel 414
Horváth Mário 26
Hreha Daniel 423
Humeník Jozef 140
- Ch**
Chavková Katarína 215
- J**
Jadlovský Ján 170, 267, 367
Jakab František 7, 12, 17, 22
Janda Peter 394, 408
Jankurová Petra 89
Jano Milan 293
Janok Michal 40
Jansky Roland 221
Jevčák Ján 96
Jevčák Juraj 373
Juhár Jozef 399
Juríni Filip 210
- K**
Kaňuch Ján 455, 464
Kmec Tomáš 164
Kolibár Daniel 478
Kopecký Martin 267
Korečko Štefan 118, 124
Kormanský Daniel 282
Kovalčín Branislav 432
Kováčová Lívia 193
Kriššák Martin 188
Kurkina Natalia 346, 351, 356
Kyslan Karol 449
- L**
Lacko Milan 444
Lahunov Dmytro 35
Levendovský Stanislav 158
Lohaj Oliver 96
- M**
Madoš Branislav 193, 215, 221, 253, 272
Magyar Ján 418, 423, 428
Mach Marián 30, 35, 204, 403
Machová Kristína 26, 40, 129, 335, 432
Majkut Jozef 229
Marcin Daniel 444
Marcinek Adrián 438
Marchevský Stanislav 394, 408
Mattová Miriama 118, 124
Matulová Soňa 118
Medveď Dušan 182, 188, 210, 242, 248
Molčanyi Dominik 389
Molnár Ján 340
Molnár Patrik 362
Móricová Nikola 12, 22
Murin Miroslav 17
- N**
Nemergut Peter 288
Nguyen Martin 68
- O**
Oklamčák Samuel 170
- Oravec Jakub 57, 62, 79, 85, 149, 153
Ovseník Luboš 57, 62, 79, 85, 149, 153
- P**
Pancurák Lukáš 449
Papaj Ján 317, 323, 329, 346, 351, 356
Paralič Ján 96
Pástor Marek 373, 438, 473
Pejková Alexandra 335
Perduková Daniela 259, 298, 314
Pietriková Emília 229
Pilár Slavomír 394, 408
Plavec Miroslav 129
Pleva Matúš 384, 389, 414
Pribula Richard 85
- R**
Radváni Peter 298
Rusnák Lubomír 204
- S**
Sabol Matúš 323
Sárosi Peter 329
Sedlačko Martin 253
Schlachta Martin 57
Skybjak Ján 30
Smoleň Pavol 305
Smolko Matúš 51
Sobota Branislav 118, 124
Sobota Timotej 7
Sokolová Zuzana 399, 414
Staš Ján 399
Strigáč Silvester 367
Szomosi Vladimír 242
- Š**
Šimon Kristián 351
Šlapák Viktor 460
Švenk Erik 384
- T**
Titko Samuel 45
Tkáčik Tomáš 170, 267, 367
Tóth Marek 428
- U**
Ujhelyi Dávid 134
- V**
Vaščák Ján 45, 51, 73, 176, 199
Vápeník Roman 7, 12, 17, 22
Vrábel Michal 158
Vysokai Matej 237
- Ž**
Žilková Jaroslava 288, 293, 373, 379, 455, 473

Electrical Engineering and Informatics 14

**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, 486 pages
Number of CD Proceedings: 50 pieces

Editors: Karol Kyslan
Emília Pietríková
Lukáš Pancurák

ISBN 978-80-553-4407-2