

Technická univerzita v Košiciach
Fakulta elektrotechniky a informatiky

QuoVadis Research @ FEI



ročník 6, č.2
September 2023

ISSN 2585-9587

QuoVadis Research @ FEI je vedecký časopis, ktorého obsahom sú prezentácie výsledkov vedeckého bádania v oblasti infor-
matických a elektrotechnických vied na Fakulte elektrotechniky a informatiky Technickej univerzity v Košiciach (FEI TUKE).
Časopis obsahuje plnotextové vedecké články, ktorých obsahom sú výsledky dizertačných prác doktorandov a ich školiteľov
obhájajúcich v aktuálnom období na Fakulte elektrotechniky a informatiky Technickej univerzity v Košiciach. Vedecké články
prispievajú k rozšíreniu vedeckých poznatkov v odboroch informatiky a elektrotechniky. Ide o podporu a rozvoj takých
odborov ako je výpočtová technika a informatika, počítačové siete, kyberbezpečnosť, mechatronika a automatizácia, ky-
bernetika a umelá inteligencia, multimediálna telekomunikačná technika, aplikovaná informatika, počítačové modelovanie,
vybrané oblasti matematiky, hospodárska informatika, elektronika, mikroelektronika a optoelektronika, spracovanie signálov,
teoretická elektrotechnika, elektroenergetika a technika vysokých napätí, silnoprúdová elektrotechnika, meracie systémy v
elektronike, riadenie elektrotechnickej výroby, materiály a technológie elektroniky, štruktúra a vlastnosti moderných ma-
teriálov. Prijaté rukopisy uvádzajú relevantnú medzinárodnú literatúru a opisujú nové objavy, aplikácie nových techník a
metodík, použitie existujúcej metodológie na nový problém, čím prispievajú k rozširovaniu základných vedeckých poznatkov
v danej vednej disciplíne.

Redakčná rada

Šéfredaktor časopisu:

prof. Ing. Liberios Vokorokos, PhD.

Výkonný redaktor:

doc. Ing. Karol Kyslan, PhD.

Členovia:

prof. Ing. Jaroslav Porubán, PhD.

odbor: informatika

prof. Ing. Liberios Vokorokos, PhD.

odbor: informatika, kyberbezpečnosť

prof. Ing. Peter Sinčák, CSc.

odbor: kybernetika, umelá inteligencia, inteligentné systémy, aplikovaná informatika

prof. Ing. Ján Paralič, PhD.

odbor: hospodárska informatika, informatika

prof. RNDr. Ján Plavka, CSc.

odbor: aplikovaná matematika, počítačové modelovanie, informatika

prof. Ing. Jozef Juhár, CSc.

odbor: multimediálne telekomunikácie, počítačové siete, informatika

prof. Ing. Michal Kolcun, PhD.

odbor: elektroenergetika a technika vysokých napätí, elektrotechnika

prof. Ing. Dobroslav Kováč, CSc.

odbor: teoretická elektrotechnika, automatizácia, elektrotechnika

prof. Ing. Daniela Perduková, PhD.

odbor: elektrotechnika, mechatronika, riadenie elektrotechnickej výroby

prof. Ing. Alena Pietriková, CSc.

odbor elektrotechnológie a materiály, elektrotechnika

prof. Ing. Ján Šaliga, PhD.

odbor: elektronika, spracovanie signálov, optoelektronika, meranie, elektrotechnika

prof. RNDr. Ján Ziman, CSc.

odbor: štruktúra a vlastnosti moderných materiálov, elektrotechnika

Adresa vydavateľa a redakcie: Redakcia časopisu QuoVadis Research @ FEI, Technická univerzita v Košiciach Fakulta
elektrotechniky a informatiky, Letná 9, 042 00 Košice, Slovenská republika, email:quovadis@fei.tuke.sk

Obsah

Analysis and evaluation of data using artificial intelligence for cyber security	
<i>Martin Hasin, Martin Chovanec</i>	4-10
Characterization of computer flow data based on machine learning to streamline network traffic	
<i>Martin Havrilla, Martin Chovanec</i>	11-17
Sémantický model pre popis dátovo-analytických procesov	
<i>Juliana Ivančáková, Peter Bednár</i>	18-26
Impact of magnetic and electric field on structure of magnetic fluids	
<i>Maksym Karpets, Milan Timko, Peter Kopčanský, Michal Rajňák</i>	27-33
Vysvetliteľnosť umelej inteligencie pri spracovaní videosnímkov z USG záznamov	
<i>Michal Kolárik, Ján Paralič, Martin Sarnovský</i>	34-40
Výskum supravodivosti v silne neusporiadaných systémoch	
<i>Marek Kuzmiak, Pavol Szabó</i>	41-47
Skúmanie obmedzení prenosu dát v hybridných FSO/RF systémoch	
<i>Maroš Lapčák, Ľuboš Ovseník</i>	48-58
Riadenie viacportového aktívneho mostíka	
<i>Adrián Marcinek, Marek Pástor</i>	59-79
Bezsúťažové riadenie synchronného motora s permanentnými magnetmi pre nízke rýchlosti	
<i>Viktor Petro, Karol Kyslan</i>	80-86
Design of Smart Network in an Industrial Zone	
<i>Rikin Tailor, Ľubomír Beňa</i>	87-102
Modifikácia elektrických a magnetických vlastností amorfných feromagnetických mikrodrôtov	
<i>Simeon Samuhel, Jozef Onufer</i>	103-110

Metódy a algoritmy akcelerácie/paralelizácie výpočtov fyzikálnych modelov

Michal Solanik, Jan Genčí, Pavol Bobík 111-119

Systém ochrany a manažment porúch pre mikrosiete a aktívne distribučné siete

Róbert Štefko, Zsolt Čonka 120-130

Skúmanie obmedzení prenosu dát v plne optických prenosových sieťach

Norbert Zdravecký, Ľuboš Ovseník 131-139

Analysis and evaluation of data using artificial intelligence for cyber security

Martin Hasin

Department of Computers and Informatics,
Faculty of Electrical Engineering and Informatics,
Technical University of Košice,
Letná 9, 042 00 Košice,
e-mail: martin.hasin@tuke.sk

Martin Chovanec

Department of Computers and Informatics,
Faculty of Electrical Engineering and Informatics,
Technical University of Košice,
Letná 9, 042 00 Košice
e-mail: martin.chovanec@tuke.sk

Abstract—Táto práca sa zaoberá analýzou HTTP útokov na aplikačnú infraštruktúru. Cieľom tejto práce je analyzovať možné typy DDoS útokov, ktoré môžu napadnúť server. Na základe týchto už známych útokov boli vyvinuté metódy zberu dát z aplikácie a sieťovej infraštruktúry. Pomocou nerelačnej databázy Elasticsearch a grafického rozhrania Kibana môžu vytvorené analytické grafy popisovať meranú infraštruktúru. Na meranie bolo použité meranie toku NetFlow a analýza LOG správ webového servera apache2 websphere. Meranie ukázalo centralizovaný útok na serverovú infraštruktúru z dvoch krajín, a to z Číny a USA, odkiaľ bol vytvorený pomalý útok outloňov na webové služby. Takýto útok ovplyvnil využitie aplikačných serverov využívajúcich služby HTTP. Pomocou upravenej infraštruktúry na vytvorenom datasete sa porovnávali rôzne typy modelov strojového učenia. V práci sú použité rôzne typy algoritmov, ktoré boli porovnané, pričom najlepšie výsledky dosiahli algoritmy RNN LSTM. Tieto algoritmy sa pri použití upravenej infraštruktúry považujú za vhodné na detekciu DDoS a iných útokov v sieťach pomocou nami vytvorených datasetov. Tieto algoritmy umožňujú rýchlu analýzu toku siete. Na základe týchto požiadaviek bola navrhnutá modifikovaná topológia pre analýzu a ukladanie výsledkov z prevádzky siete. Overenie navrhnutých modulov prinieslo potvrdenie ich distribuovaného riešenia schopného analyzovať sieťové pakety a následne ich ukladať do distribuovaných databázových systémov. Pomocou navrhovaného riešenia je možné v budúcnosti rýchlo rozvíjať a rozširovať vytvorený dataset o ďalšie doplnkové parametre, čo znamená, že model dokáže reagovať aj na nové hrozby, ktorých signatúry sa časom vyvíjajú.

Kľúčové slová: HTTP útok, DDoS detekcia, strojové učenie, nerelačná databáza, nfstream

Abstract—This work deals with the analysis of HTTP attack on application infrastructure. The aim of this work is to analyze the possible types of DDoS attacks that can attack the server. Based on these already known attacks, methods of data collection from the application and network infrastructure were developed. Using the non-relational Elasticsearch database and the graphical Kibana interface, the created analytical graphs can describe the measured infrastructure. NetFlow flow measurement and LOG message analysis of the apache2 websphere WEB server were used for the measurement. The measurement showed a centralized attack on the server infrastructure from two countries, namely China and the USA, from where a slowloris attack on web services was created. Such an attack affected the utilization of application servers using HTTP services. With the help of the modified infrastructure on the created dataset, different types of machine learning models were compared. Different types of algorithms are used in the work, which were compared, with

the RNN LSTM algorithms achieving the best results. These algorithms, when using a modified infrastructure, are assumed to be suitable for detecting DDoS and other attacks in networks using datasets created by us. These algorithms enable fast analysis of network flow. Based on these requirements, a modified topology was designed for the analysis and storage of results from network operation. Verification of the proposed modules brought confirmation of their distributed solution capable of analyzing network packets and subsequently storing them in distributed database systems. With the help of the proposed solution, it is possible to quickly develop and expand the created dataset with other additional parameters in the future, which means that the model can also respond to new threats, the signatures of which evolve over time.

Keywords: HTTP attack, DDoS detection, machine learning, non-relational database, nfstream

I. INTRODUCTION

At present, methods of accessing information over the Internet have moved to HTTP communication. This protocol provides intuitive access and graphical display of data shared by its owner with all participants connected to the Internet. The introduction of this method of communication also brings the emergence of new, sophisticated attacks which are designed to eliminate servers providing these services via HTTP web page. Such a limitation can currently cause not only a loss of information but also a financial loss. The malfunction of web services in the commercial sphere can produce a large financial loss, for example in online sales[2].

The security of web servers and their optimization and scaling of resources is often underestimated in organizations where the graphical design of the website is a priority. When creating security, it is necessary to implement a model of data collection and subsequent evaluation.

Figure 1 shows a model that provides an optimal system for protecting a web server from attacks. The model consists of 4 items where each item is an important part in detecting or blocking inappropriate network traffic. Detection of DDoS attacks performed by sophisticated botnets is the most complex problem, where the difference between a bot performing a server load and a regular user is not large. The biggest problem is to correctly identify the required patterns of behavior which

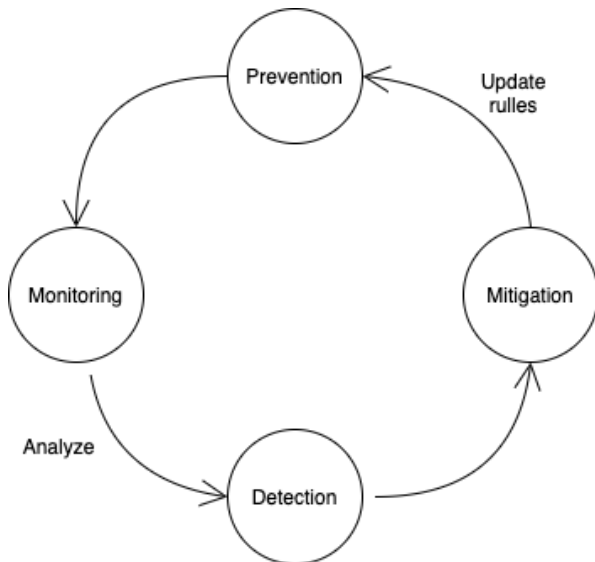


Fig. 1. HTTP DDoS defence life cycle

largely represent only inappropriate network traffic while achieving the least possible impact on real clients [3].

II. METHODS OF DDOS ATTACKS ON HTTP PROTOCOL

In environments where the HTTP protocol is widespread, a large number of network attacks are created. These attacks can be divided into a group of attacks aimed at compromising a given web server. Examples of such attacks are XSS attack, Injection attack, Zero-Day attack and others [9]. The task of this group of attacks is to use the method and compromise the given web server which spoils the data on the web page or steals sensitive information from internal databases with which the web site works. The second model of attacks on the HTTP protocol is to overload the server providing the web page. Achieving the correct load shuts down a webpage that is not available to users working on it [6]. The best known methods include DoS/DDoS attacks which can be divided into:

- Session Flooding Attack - This attack process is aimed at creating a large number of connections to the HTTP server which the server must handle properly. Connections are usually made when making GET/POST requests to a server. Dragging a sufficient number of such connections can cause the system resources of the HTTP server to be depleted, the response time of which will increase or the server will stop responding to requests [14]. To carry out this type of attack, an attacker needs to secure a sufficiently large botnet that can generate several dozen requests per second. To carry out this type of attack, an attacker needs to provide a sufficiently large botnet that can generate several dozen requests per second. Such a botnet consists of a large number of compromised computers connected to a network.
- Request Flooding Attack - This method is similar to the Session Flooding Attack method but with this type of

attack, a large number of requests to the WEB server are created within the HTTP protocol during one connection. In most cases these requests are larger and longer than the average user can create while visiting a website. Within the HTTP 1.1 protocol, it is possible to initiate multiple requests which are in the form of GET/POST [15]. These requests can be processed by the server multiple times in one session which can speed up the server but by sending specially modified requests, an attacker can overwhelm the server's resources. This type of attack occurs using a botnet that is controlled and the attacker can make special GET requests to the requested server. When the maximum limit is reached the system resources are depleted which causes the WEB page to malfunction.

- Asymmetric Attack - This type of attack combines server load with other methods. Most often the attacker attacks a server with HTTP requests which can overwhelm or slow down a database server from which the given web page retrieves data. Such an attack also slows down other applications linked to the database. Another option is to periodically download large files using the HTTP protocol stored on the server. This type of attack congests the network connection, making the server unavailable. Such an attack can also disable other applications running on the infrastructure.
- Slow Request/Response Attack - When using these types of attacks, an attacker sends a large number of HTTP requests to the server [16]. The server accepts these requests and opens new connections that the attacker does not close, depleting the server's system resources. The attack on the server usually takes place using a low limit of open connections to one of attacker's computers. This makes it difficult to distinguish the attack from real network traffic.

Another option is to use the TCP/IP protocol to create attacks. These attacks can be considered DoS attacks where the attacker generates appropriately modified packets to the server that must then process them [4]. When using these attacks, identified vulnerabilities in the application protocols in use are exploited. Examples of DoS attacks:

- TCP SYN attacks (SYN Flooding) - Communication using the TCP protocol uses an ordering system that communicates with the server using a three-way handshake phase. These attacks can exploit errors in this connection initialization process by sending deceptive TCP SYN attributes. If the server receives a TCP SYN request, system time is allotted to monitor the current connection in the queue. An attacker can send the requests but never responds with the appropriate TCP SYN response. In this way the table securing the connection queue is exhausted and the service becomes unavailable [1].
- TCP RESET attacks - In this type of attack, an attacker monitors the network traffic of the server and waits for new requests [8]. When a new incoming request is made, the attacker sends a packet with the TCP RESET attribute

to the server whereby the server closes the requested connection before it is even established.

- ICMP attacks - Internet Control Message Protocol is a protocol used in TCP/IP networks to send information or control messages to a server where the best known message is ping. As this protocol does not allow any authorization, these messages can be spoofed. By changing the headers of these messages it is possible to get the server to respond to invalid responses.

III. MEASUREMENT SETUP

The network we are analyzing consists of a web server based on apache2. This web server is represented as a virtual machine on the VMware platform. This platform ensures the creation of multiple virtual machines on one physical device.

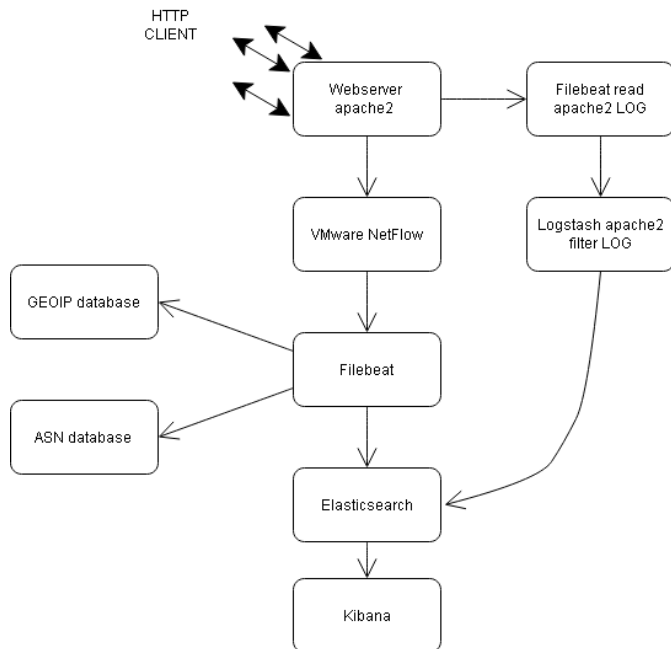


Fig. 2. Measurement topology

Figure 2 shows the topology of the web server connection and the topology used to measure the ongoing DDoS attack. The measurement consisted of obtaining two sources of data, namely:

- NetFlow protocol
- Analysis of apache2 LOG files for access and error LOG

NetFlow measurement consisted of the implementation of a NetFlow collector which allows you to analyze network traffic directly entering the measured server [12]. This data represents the current network connections whose information has been sent to the Filebeat server. This server checks the received applications and divides them into specialized columns. To better detect network attacks it is necessary to ensure the geolocation of the measured network traffic. The system automatically uses predefined filters to determine the current location of the source IP address that is accessing the server [10]. It performs such control using GEOIP databases

in which it is defined which location is represented by a given IP address. Subsequently from this data using the database to determine the ASN (Autonomous System), the current owner of the network range and the assigned ASN number are determined. This number represents one full range of the network assigned to one owner of a given IP subnet. For the completeness of the analysis, a system of data collection was also created from the application WEB server apache2 which can generate LOG messages from current accesses to the server[13]. LOG messages are divided into two types, namely access LOG and error LOG. Using access LOG it is possible to define the current clients that have successfully accessed the server. The processed data from these two sources are stored in a common non-relational Elasticsearch database which allows the Filebeat system to perform network analysis using machine learning and also to represent this data in graphical form as clear graphs [7], [11].

TABLE I
COMPUTER SPECIFICATION

Application	System	Configuration
Webserver	RedHat 6	4X CPU Intel Xeon e2670 V3, 10GB RAM, 200GB HDD
Database DB2	RedHat 6	2x CPU Intel Xeon e2670 V3, 10GB RAM, 200GB HDD
Measurement 3x Elasticsearch	Debian 10	10x CPU Intel Xeon e2670 V3, 20GB RAM, 4TB HDD
LogStash, Kibana and Filebeat	Debian 10	8x CPU Intel Xeon e2670 V3, 10GB RAM, 200GB HDD

Table I lists the currently used hardware used to run the WebServer and the servers performing the analysis. The application server consists of two servers which together form one unit. The first server is the apache2 web server itself which is connected to a database based on DB2 technology. The WebSphere server is used as the server for rendering the WEB which ensures the creation of a WEB that the user can read in a comprehensible form. The infrastructure that serves as a WEB server consists of two servers based on RedHat (4x CPU, 4GB RAM, 50GB HDD). The server that serves as a collector contains the installed Filebeat application. The server used for the non-relational Elasticsearch database was based on Debian 10 (8x CPU, 16GB RAM, 200GB HDD).

The measured data in the Elasticsearch database can be viewed using the Kibana WEB interface. This interface provides a graphical interpretation of the data stored in the database. Appropriate graphs can be created from the stored data. NetFlow stores a large amount of data about a given network flow in a database. Basic and important data include the source and destination address and the ports through which the communication took place.

IV. HTTP ANOMALY ANATOMY

The started data analysis consisted of the infrastructure connection as in Figure 2. ALOG files created by the apache2 web server were represented as input data. The monitored infrastructure was attacked by WEB services in the period from 23.9.2021 to 25.9.2021, as shown in Figure 4. The

attack itself consisted of three separate attacks between which short breaks were observed. This distribution of attack power may correspond to the possibility of no longer available web services that stopped responding when the maximum attack was reached. The bots' functionality thus responded to the web server's malfunction by shutting down the attack.

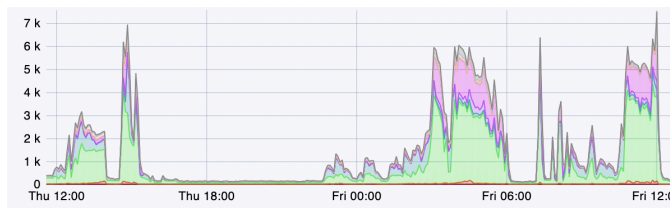


Fig. 3. TCP connections to HTTP server

In Figure 3 is a graph of TCP connections that were established during the attack on the WEB server. These connections represent requests to the web server that the server had to process. The connections increased 100-fold compared to normal operation which caused a load on the WEB server that could no longer process new requests to display the WEB. Connections to the server consisted of creating a large number of small connections with the GET request to the WEB server which resulted in the utilization of system resources where the server became inaccessible to the public. Within its capabilities the server was able to continue to process network communication.

TABLE II
HTTP CONNECTION TYPE IN MAXIMUM STATE OF ATTACK

TCP SYN parameter	Incidence MAX
FIN_WAIT2	1.28K
FIN_WAIT1	314
CLOSING	81
CLOSE_WAIT	1.71K
LAST_ACK	485
TIME_WAIT	2.5K
ESTABLISHED	2.27K
SYN_RECV	177
SYN_SENT	1
LISTEN	20

Table II shows the maximum number of connections measured on the WEB server during the highest peak of the attack. The most used connections are ESTABLISHED, i.e. successfully established connections. The next part consists of connections that are waiting to be processed by the WEB server. The number of pending connections exceeded the value of already established connections which is shown in Figure 3.

From the measured values given in Table I III is clear that the most frequent call to the server is of the GET type which represents up to 99.1 percent of all received requests. A small part of the calls is the POST and HEAD method. Using the GET method the content of the WEB page is requested whereby the page loads and sends the data of the

TABLE III
HTTP SERVER REQUEST METHOD

Request method	Incidence [%]
GET	99.1
HEAD	0.5
POST	0.4

page body. Such transmission and page creation takes up network bandwidth as well as the number of connections to the server, as the communication took place exclusively using the HTTP 1.1 protocol which does not allow the number of connections to be optimized compared to HTTP2.

Figure 5 shows the number of bytes that were transferred from the WEB server for clients. It can be seen that at the time of the attack the file transfer on a minute basis exceeded the level of 400MB/min while during normal operation the transfer takes a maximum of 40MB/min. Such a rapid increase in clients caused a load on the system which had to generate the contents of the WEB page stored in the DB2 database. Such generation takes a large amount of processor time which was reflected in the increasing load of processors that exceeded the value of 60. With a high processor load, a sequence of processing requests to the WEB server was created which made the WEB inaccessible to some clients or its loading was lengthy.

Figure 6 shows a list of agents that have connected to the web server. Among the most common agents used by users to connect to the web server are:

Mozilla/5.0 (Windows NT 10.0; Win64; x64)
AppleWebKit/537.36 (KHTML, like Gecko)
Chrome/68.0.3440.75 Safari/537.36

This agent represents the three most used devices that are used to communicate with the WEB server. These are the three most used web browsers through which users access the WEB server. The signatures of these agents are also used by various applications that communicate using the HTTP protocol. The second most used client is:

Mozilla/5.0 (compatible; SemrushBot/7~bl; +http://www.semrush.com/bot.html)

This client represents an automatic BOT server that is used as a web robot, web spider or web crawler. This agent is used by the "semrush" software. As a result the bot that attacked the web server was able to successfully disguise the agent it used to access the server as an agent that is used globally on the network.

Table IV shows the percentage of clients from which countries they join. The most frequented clients connect from the country SLOVAKIA as the given website is created for this country. The second most frequented country is CHINA, followed by the USA. In total these two countries generated more than 40% of the network traffic on the server which was illegitimate. As the server is built for access from local networks, such growing network traffic evokes unwanted access

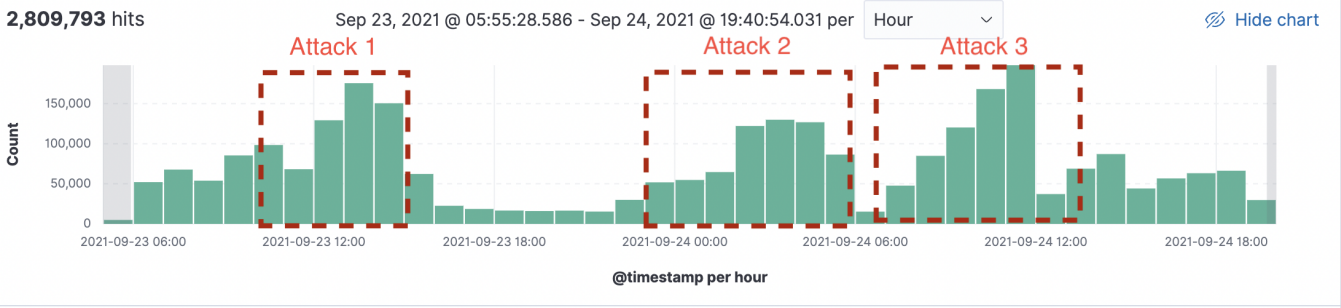


Fig. 4. Running HTTP attack, detected from HTTP LOG

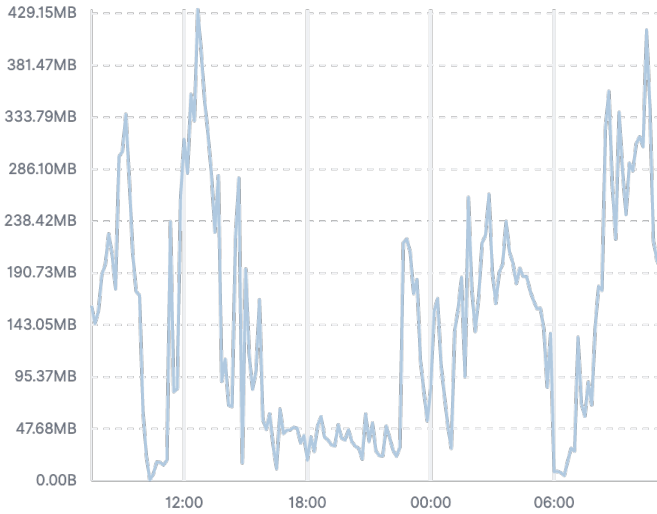


Fig. 5. HTTP server response bytes SUM

TABLE IV
HTTP TOP LOCATION IDENTIFIED BY GEOIP

HTTP GEOIP location	Incidence [%]
SK	43.31
CN	26.46
US	17.83
NL	2.13
DE	0.93
OTHER	9.35

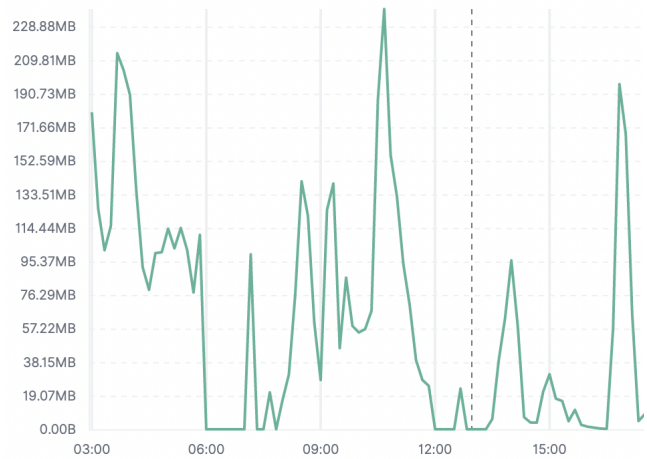


Fig. 7. HTTP agent, connections to WEB server

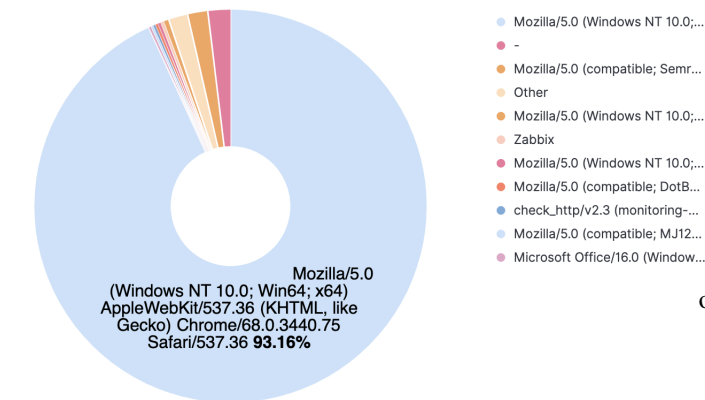


Fig. 6. HTTP agent, connections to WEB server

to the server, i.e. the activation of a network botnet that has mostly attacked from China and US networks.

Figure 7 shows network traffic to the server from networks operating in China. These networks were able to create up to 200MB/min of network traffic on the WEB server. This network traffic is considered undesirable on the network as it generated a large number of GET requests to the server,

overloading the server by loading the web page.

V. MACHINE LEARNING FOR DETECTING HTTP ATTACK

Comparing machine learning algorithms is a key step in the process of developing machine learning models for DDoS attack detection. There are many machine learning algorithms, and each has its own strengths and limitations that can affect its performance in different tasks. The main factors usually considered when comparing machine learning algorithms are accuracy, training speed, robustness and interpretability. These parameters were described in previous chapters. Based on these parameters, it is necessary to evaluate and select the correct machine learning model, which will then be implementable in the following experiments.

Machine Learning Model	Accuracy	Training Time (s)
Decision Trees	94.45%	3965
Random Forests	94.98%	6153
Logistic Regression	84.37%	5321
Neural Networks RNN	98.59%	3847
SVM	96.59%	4429
Gradient Boosting	96.05%	5523
Naive Bayes	92.07%	4457

TABLE V
TESTING MACHINE LEARNING MODELS

When choosing machine learning models, it is important to take care to achieve as much accuracy as possible while achieving a relatively short time to train the model. Based on the analysis described in chapter 2, some machine learning methods were selected and tested. Tab. V describes the result of an experiment with training different machine learning models. When creating this experiment, global settings were used:

- Number of epoch repetitions = 100.
- Dataset size = 4095865 records describing real and attack network traffic.
- The distribution of the dataset in the ratio 70:30 where 70% represented the attack samples and 30% of the samples represented the verification test data.

The function of the selected parameters was to create a homogeneous environment where it would be possible to compare the accuracy and speed of machine learning algorithms. During the measurement, 7 experiments were created, where each experiment represented one selected machine learning algorithm. Subsequently, the experiment was repeated for each algorithm. The aim of this experiment was to determine the accuracy and training speed of the machine learning model that would be most suitable for the classification of network attacks and DDoS attacks.

As described in tab. V the results of measurement and comparison of machine learning models reached accuracy from 92.07% to 98.59%. The model that achieved the highest accuracy was able to train the model at 100 epochs in 3847 seconds. In the next work, we are subsequently considering the use of RNN neural networks, which in comparison of accuracy were able to detect DDoS attacks in the measured dataset the most successfully. By analyzing the RNN network described in chapter 2, it is possible to divide RNN networks into three selected models:

- Basic RNN (Recurrent Neural Network) - is a type of neural network where connections between nodes are made to form a sequence of data. This allows RNNs to process data such as text or time-dependent data. A problem with classical RNNs is the so-called vanishing gradient problem, where the model loses its ability to learn from a data sample that is too long. This model is not suitable for classification of a large amount of data.
- LSTM (Long Short-Term Memory) - is a special kind of RNN that is designed to be able to learn from long time sequences, thus solving the gradient vanishing problem

that occurs in classic RNN networks. These models are widely used in areas such as speech processing, pattern and anomaly recognition, and many other areas where sequential data processing is important.

- GRU (Gated Recurrent Unit) - is a type of recurrent neural network (RNN) suitable for processing sequential data. This type is similar to LSTM networks where the vanishing gradient problem is solved, but the algorithm allows you to choose and set fewer parameters compared to LSTM. The advantage of this model is its lower complexity for calculations, which, however, compensates for the disadvantage of the closedness of the given model without the possibility of intervention or finding out why the model's error rate occurs.

Based on the comparison of these machine learning methods that use RNN neural networks, three experiments using these methods for classification of anomalies were designed and implemented. Tab. VI describes a comparison of three algorithms, namely classical neural networks, LSTM networks and GRU networks. All three algorithms achieve model accuracy of 94% or more. The accuracy of these models consequently depends on the amount of data they process. The basic RNN machine learning algorithm is unsatisfactory when analyzing a large amount of data, as it is not capable of processing a large set of data. The other two algorithms optimize this disadvantage, but when comparing the two methods, we chose LSTM, which achieves an accuracy of 97.87%, compared to the GRU model, which achieves an accuracy of 96.21%. The difference between these models was the time it took them to learn and recognize anomalies in the network. In this category, the LSTM network needed a longer time interval, but it allowed a better model setting and thus its optimization for classification [5].

Type of RNN network	Accuracy	Training Time (s)
Basic RNN	94.52%	4842
LSTM	97.87%	5144
GRU	96.21%	4871

TABLE VI
COMPARISON OF RNN ALGORITHMS

VI. CONCLUSION AND FUTURE WORK

Using technology to collect data from application servers has its own potential. By measuring the network traffic and application logs of the apache2 web server, it was possible to analyze the LOG messages from access messages and messages containing errors on the page. Using Elasticsearch technology with graphical interpretation in Kibana it is possible to easily and quickly search for data collected from application servers and the network layer. By creating custom modules that filter messages and then categorize them for easier retrieval and using external databases, data can be categorized by geographic location or by IP network owner according to the AS network list. The most important part of the application messages were the access logs which represented the current accesses to the web server. Using these approaches along with

GEOIP databases it was possible to create a map of the states from which the server was most accessed. The analysis found that the largest number of attacks on the server came from China and the USA. These two networks were able to utilize access to the server using large access of GET requests to the HTTP server. It is known from the measured data that the largest number came from the Chinese operator "CHINA UNICOM China169 Backbone", this operator owns several networks from which the attack took place. The attack took place by sending a small number of requests from various IP addresses registered under the ownership of this registrar. The application of network protection was blocking access from these networks to the server infrastructure. With this data only networks that generate unwanted traffic could be easily blocked. In the future work it is possible to create an automated system for detecting and searching for unwanted attacks which search the database for anomalies caused by the attack. Such an autonomous system makes it then possible to quickly intervene in the current security infrastructure and block the already given network traffic at the beginning without limiting the operated services. The result of the work is the implementation of a modified architecture, which consists of the classic architecture supplemented by additional blocks processing network traffic. An additional analysis of network flows was carried out as additional blocks. This modified architecture enables the detection of attacks already in the process of analyzing network flows and LOG reports. It also enables the implementation of parallel calculations since the implemented architecture consists of separate blocks that can be autonomously distributed to multiple devices at the same time. The result of this work is the implementation of a modified architecture where the accuracy of the machine learning model was verified using RNN LSTM machine learning.

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. Intelligent systems for UAV real-time operation and data processing, code ITMS2014+: 313011V422 and co-financed by the European Regional Development Found

REFERENCES

- [1] Ç. B. Aslan, R. B. Sağlam, and S. Li, "Automatic detection of cyber security related accounts on online social networks: Twitter as an example," in *Proceedings of the 9th International Conference on Social Media and Society*, 2018, pp. 236–240.
- [2] J. Bakker, B. Ng, W. K. Seah, and A. Pekar, "Traffic classification with machine learning in a live network," in *2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*. IEEE, 2019, pp. 488–493.
- [3] N. K. Chawla and A. Singh, "Review of cyber crime in india: An overview," *Asian Journal of Technology & Management Research*, vol. 6, no. 1.
- [4] M. Chovanec, E. Chovancová, and M. Dufala, "Dids based on hybrid detection," in *2014 IEEE 12th IEEE International Conference on Emerging eLearning Technologies and Applications (ICETA)*. IEEE, 2014, pp. 79–83.
- [5] M. Chovanec, M. Hasin, M. Havrilla, and E. Chovancová, "Detection of http ddos attacks using nfstream and tensorflow," *Applied Sciences*, vol. 13, no. 11, p. 6671, 2023.
- [6] D. Edwards, S. Simmons, and N. Wilde, "Prevention, detection and recovery from cyber-attacks using a multilevel agent architecture," in *2007 IEEE International Conference on System of Systems Engineering*. IEEE, 2007, pp. 1–6.
- [7] P. Fanfara, M. Dufala, and E. Chovancova, "Usage of proposed autonomous hybrid honeypot for distributed heterogeneous computer systems in education process," in *IEEE 11th International Conference on Emerging eLearning Technologies and Applications (ICETA)*, Oct 2013, pp. 83–88.
- [8] J. Hurtuk, M. Chovanec, and N. Ádam, "The arduino platform connected to education process," in *2017 IEEE 21st International Conference on Intelligent Engineering Systems (INES)*. IEEE, 2017, pp. 000071–000076.
- [9] G. A. Jaafar, S. M. Abdullah, and S. Ismail, "Review of recent detection methods for http ddos attack," *Journal of Computer Networks and Communications*, vol. 2019, 2019.
- [10] F. Jakab, J. Giertl, R. Jakab, and M. Kaščák, "Improving efficiency and manageability in ipfix network monitoring platform," in *Proc. International Network Conference (INC'06)*, Plymouth, United Kingdom, 2006, pp. 81–88.
- [11] S. Maeda, A. Kanai, S. Tanimoto, T. Hatashima, and K. Ohkubo, "A botnet detection method on sdn using deep learning," in *2019 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, 2019, pp. 1–6.
- [12] A. Pekár, J. Giertl, M. Révész, and P. Fecíľak, "Simplification of the real-time network traffic monitoring," in *Proc. Electrical Engineering and Informatics II, (EEI II'11)*, Košice, Slovakia, 2011, pp. 272–277.
- [13] L. Vokorokos, M. Ennert, J. Radušovský *et al.*, "A survey of parallel intrusion detection on graphical processors," *Open Computer Science*, vol. 4, no. 4, pp. 222–230, 2014.
- [14] L. Vokorokos, A. Pekár, N. Ádám, and P. Darányi, "Yet another attempt in user authentication," *Acta Polytechnica Hungarica*, vol. 10, no. 3, pp. 37–50, 2013.
- [15] T. Yatagai, T. Isohara, and I. Sasase, "Detection of http-get flood attack based on analysis of page access behavior," in *2007 IEEE Pacific rim conference on communications, computers and signal processing*. IEEE, 2007, pp. 232–235.
- [16] T. Zseby, E. Boschi, N. Brownlee, and B. Claise, "IP Flow Information Export (IPFIX) Applicability," RFC 5472 (Informational), Internet Engineering Task Force, 2009.

Characterization of computer flow data based on machine learning to streamline network traffic

Martin HAVRILLA

Department of Computers and Informatics
Faculty of Electrical Engineering and Informatics
Technical University of Košice
Letná 9, 042 00 Košice, Slovakia
Email: martin.havrilla@tuke.sk

Martin CHOVANEC

Department of Computers and Informatics
Faculty of Electrical Engineering and Informatics
Technical University of Košice
Letná 9, 042 00 Košice, Slovakia
Email: martin.chovanec@tuke.sk

Abstract—This work deals with the issue of classification of network flows using machine learning methods. One of the goals is to create a suitable model for the classification of various anomalies in network traffic, such as DDoS attacks, port scanning or data exfiltration. An analysis of selected machine learning methods is created with the aim of choosing suitable methods for the classification of network flows. An experimental architecture for the collection and storage of network flows from real network traffic is created in the work, and the created dataset is also a benefit, which, together with the available datasets, is a source for creating models for detecting anomalies in network traffic. Available classifiers include SVM, Naive Bayes classifiers, K-Means algorithm and Decision Trees. Using these algorithms, models for classification of anomalies are created. The benefit is the experimental part of the work, which is dedicated to improving the classification results, as well as the modified K-means algorithm. This part also deals with the optimization of selected algorithms by using trained models on worse classifiable data sets.

I. INTRODUCTION

In today's modern era, we encounter various elements of computer networks in practically every field of industry. Over the years, network infrastructure has created a significant environment for development, modernization, development of new methods, technologies, etc. Nowadays, there is probably no area where computer networks do not play at least a small role. The modern generation can no longer imagine not doing online banking transactions or not having access to the Internet anywhere and anytime. It is the internet network that has created space for opening and discovering new possibilities over the past decades. With increasing emphasis and the need to use these networks, the need to create new, more efficient and less demanding technologies or protocols that can be used through network communication often grows. Ever since the beginnings of networks, even today, more and more emphasis is placed on monitoring network traffic. This field of computer networks has grown over the years from the initial, less sophisticated methods of monitoring the functionality of network traffic, to simultaneous methods of prediction, analysis and monitoring of computer networks. Companies that focus their activities on computer networks often use prediction and analysis methods to estimate possible future problems or anomalies in their networks. In a large

number of cases, these methods are also used in the estimation of investments or the creation of various business strategies.

Therefore, it is possible to evaluate that the development in the field of computer networks is still on course and there are still many areas that require our attention. Due to the increasing load of Internet networks, this work also focuses on research in the field of network technologies, network communication processing and the search for new methods in the field of streamlining network operations based on machine learning. It uses many available methods that are adapted to the needs of streamlining network operations, such as the collection of data flows based on the available NetFlow, NFSstream or cFlow protocols. The subsequent analysis of data flows based on machine learning has the task of detecting anomalies and predicting network traffic.

The goal is to create a set of measures based on machine learning analysis, which will be autonomously reflected in the configuration of network devices in order to make network operations more efficient and secure. Last but not least, the work deals with data processing in the non-relational ElasticSearch database.

II. NETWORK FLOWS

Network flows are a key tool used to understand the nature of network traffic. A network flow can be characterized as a set of packets that go from the source to the destination of network communication. Monitoring network flows is an essential part of understanding the monitored network for several reasons. Network flows can provide important information about possible malicious activity such as intrusion attempts, various forms of flooding attacks by identifying unusual patterns of network messages. These features help to potentially increase security. Working with network flows can solve issues of efficiency and performance of computer networks, e.g. By identifying the overloaded part of the infrastructure[1]. Likewise, understanding network flows is closely related to planning and developing infrastructure for future growth. By the term flow, it is also possible to understand the sum of individual data flows. Subsequently, the record of this flow is defined as a set of information that describes the record of which hosts communicated with other hosts in the data

network, when the given communication took place and how the traffic was transmitted between the participants of the communication[1][2][3].

In general, a system of three components is used most often for the analysis of data flows, which cover the collection, storage and reporting of data flows. These components are shown in the Figure. Most often, they can also be found in the literature as[4]:

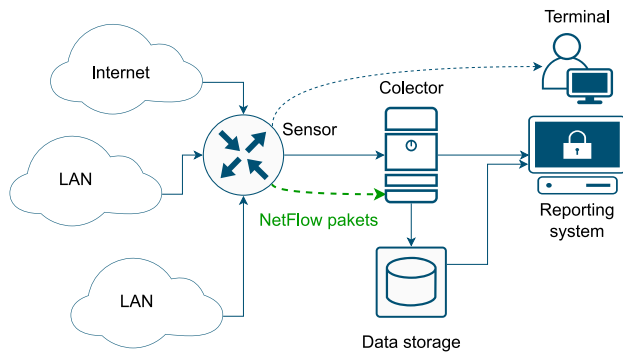


Fig. 1. Základná architektúra na zber dátových tokov

A sensor, or also a probe, is a component that covers the capture of the network and subsequently captures data about its operation from it. By default, this functionality can be covered by routers, switches, firewalls, etc. This component therefore monitors network connections and after their termination or when the connection has reached a time limit, it forwards the data[5].

The collector is most often software that expects data from the sensor and writes this data to memory. This is a very critical part of the entire architecture of data flow characterization, or possible control of network traffic based on data flows. Currently, there is no universally accepted format that covers the storage of flow records. This fact considerably limits and complicates the analysis[4]. The reporting system, after loading the data from the collector, can, on the basis of predefined rules or with the help of artificial intelligence, submit reports about the network in a humanly acceptable form. Currently, it is possible to find a considerable number of different implementations of individual components used for the analysis and management of network traffic based on flows[5]. Every manufacturer of industrial network elements already implements a data flow sensor. Similarly, many other people or companies create or implement their own sensors that meet their specific requirements.

A. NetFlow protocol

This protocol is directly intended for the analysis and monitoring of network traffic. Its direct deployment is found in routers, switches, firewalls, etc. The protocol will make it possible to extract basic information about source and destination IP addresses, ports and protocols, frequencies, etc. from network flows. The protocol works on the principle of collecting network flows from measured network traffic using network devices. These flows are subsequently exported to

the collector. This component transforms measured data to generate reports and possibly create visualizations[4][5].

NetFlow v9 brings significant improvements over older versions of NetFlow. The main innovation is a flexible and customizable architecture that works with the concept of "templates", allowing to define which information should be collected and exported. These templates make NetFlow v9 much more flexible and can include a wide range of technical parameters[4].

Basic technical parameters include source IP address, destination IP address, source port, destination port, transport layer protocol, class of service, and router interface, which are common to all versions of NetFlow[6].

However, NetFlow v9 may include many other technical parameters. For networks that use the IPv6 protocol, it can collect IPv6 addresses and ports. It provides timestamps for the start and end of the flow in the form of the time of the first and last packet. It shows how much data has passed through a given stream, through byte and packet count information. It helps to track which autonomous systems the data passes through, through information about the autonomous systems. It enables the identification of virtual LAN segments through which flows pass through VLAN tags. It helps identify specific states and behaviors of TCP flows using TCP Flags. And it provides information about BGP routes through BGP next-hop datas[2][4][5].

III. SELECTED MACHINE LEARNING METHODS

Machine learning is a branch of artificial intelligence focused on the development and application of algorithms that enable computers to learn and make predictions or decisions without explicit programming. These algorithms improve their performance based on experience gained from the data. The following machine learning methods were chosen based on a comparative analysis as the most suitable for the classification of network flows.

A. Decision Trees

Decision trees are machine learning algorithms used for classification and regression. They work on the principle of creating rules from data attributes for dividing data into subsets. Each internal node of the tree represents an attribute, a branch represents the result of a decision, and a leaf represents the final result[7].

The whole process starts from the root node, which contains the entire training dataset. The algorithm then searches for the best attribute to partition the data according to the target attribute, which is measured using impurity criteria such as Gini impurity or entropy. This process is repeated recursively for each subset until all data in a node is from one class or until all attributes are exhausted[7][8].

The resulting tree is then used to classify new examples. Due to the risk of overtraining, methods such as tree pruning, random forest or gradient boosting are often used to improve model robustness.

B. SVM

It works on the principle of constructing an N-dimensional plane, which has the task of evenly classifying data according to various parameters. For the case of binary classification, the SVM classifies the data into two categories, given the source set of instances of (x,y) pairs. In the data classification process, SVM maximizes the gap between two classes so that the best performance in the classification process is achieved[9]. SVM belongs to the category of classifiers that work above the data plane and is based on linear separability. If we take into account the assumption that we have N training data points $(x_1, y_1), \dots$ and $x_i \in R^d$ in d dimensional space and $y_i \in \{+1, -1\}$ [9]. We also consider that there is a hypersurface which is defined as (w,b), where the parameter b represents the deviation and the parameter w is the weight vector, then we can define a new object using the following relationship[9][10]:

$$f(x) = \text{sing} \left(\sum_i^n a_i y_i (x_i x) + b \right) \quad (1)$$

For all training values then there is a Lagrange multiplier α_i . The value of this multiplier then defines the severity of each data point. After finding the maximum boundary value of the hypersurface, there are points that are closest to the hypersurface. These points take on the value $\alpha_i > 0$ and are called support vectors[10].

C. K-Means

This algorithm works on the principle of dividing the studied set of data objects into several overlapping or non-overlapping parts, so that each studied object is always only in one cluster. The methods of dividing data into clusters can be divided into two categories based on the type, namely, centroid-based and medoid-based[11]. In partial clustering, which is based on the center of gravity, each cluster is represented by means of the center of gravity of the instances. In the case of clustering based on the medoid basis, each cluster is defined by the instance closest to the mean. Figure 2 shows data instances that were before clustering (on the left) and after clustering (on the right)[11].

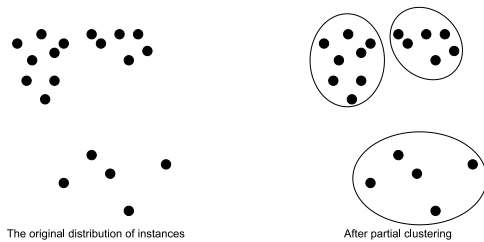


Fig. 2. Data distributed in a non-overlapping cluster

The k-means algorithm is one of the most well-known centroid algorithms. It divides the data set into several subsets so that all points in the examined subset are centered around the same center. The algorithm randomly selects the number

of instances that will represent the center of the clusters. Based on these instances, other instances are assigned to the nearest cluster center. The algorithm then determines new cluster centers by calculating the average of all points that belong to a given cluster. The process of averaging and finding a new cluster center is repeated until the convergence criterion is met[11][12].

D. Naive Bayes

A naive Bayesian model is generally a simplified Bayesian probabilistic model. In this model, it is assumed that there is a probability of the final outcome based on the associated variables. This probability is considered to be independent of the probability that the evidence variable acquires with respect to another evidence variable. During training, a set of Naive Bayes classifier classes is observed. This set then assigns the observed data to the class that has the highest probability. Subsequently, the probability of the result of a specific attribute is calculated. This process is repeated for all attributes in the dataset. Classification is then performed with an emphasis on the parent node, which is hidden[13]. Based on the calculations, the examined attribute is then assigned to the class to which it belongs. Only the conditional probability is calculated in the scope of the investigated set of attributes because the structure of the attributes is unique. After the calculation of the network, it is subsequently possible to classify any new object using the parameters of its attributes by the Bayes rule. Because the Naïve Bayes model assumes that the attributes are independent and therefore the combined probability can be obtained as[13][14]:

$$P(c_i|A) = \frac{P(a_1|c_i) \cdot P(a_2|c_i) \cdot \dots \cdot P(a_n|c_i)}{P(A)}, \quad (2)$$

where P(A) is determined based on the normalization condition. The time required to calculate the probability of a given class is directly proportional to the number of investigated attributes[14].

IV. MODIFIED ARCHITECTURE FOR NETWORK FLOW COLLECTION

The basic architecture consists of four parts, which are shown in Fig. 3. The first part is the exporter in the form of a network device that will serve as a data source. The data is then imported into the Elastic Beats module, which contains the logstash tool, which transforms the data into the required form for the ElasticSearch database. Filtering and possible data deduplication will then be implemented in this module. The data will then be stored for processing and analysis in the ElasticSearch database in the form of indexes. The Kibana tool will be used to visualize measured data in the form of graphs, datagrams, maps, etc.

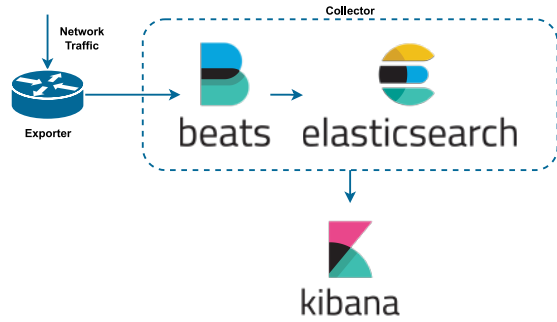


Fig. 3. Architecture for data storage based on ELK stack technologies

This infrastructure consisted of two identical virtual servers in a stack with the following parameters:

TABLE I
HARDWARE RESOURCES FOR NETWORK STREAM STORAGE ARCHITECTURE.

Component	Parameter
CPU	16 cores - Intel® Xeon® Gold 5218 Processor
RAM	32GB DDR4
HDD	60GB system 2TB ELK data
OS	Debian 11

After starting the measurement from the network device, basic data such as source and destination IP address and ports, packet frequency and bit size were imported into the ELK stack, similar to the previous chapter. The results are represented in the form of graphs and maps in the Kibana environment.

On the following fig. 4 shows a selection of the ten most requested IP addresses within the experimental network.

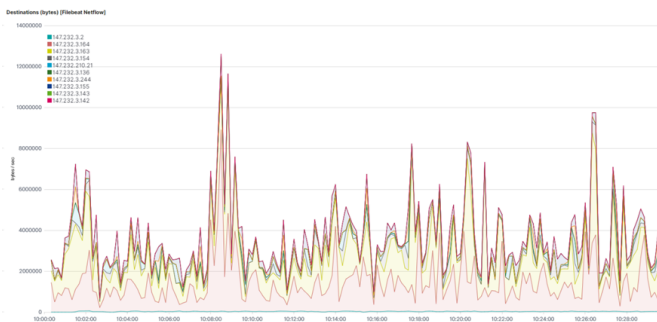


Fig. 4. The most requested target IP addresses during the experimental measurement

A. Dataset creation

An overview of the network topology is shown in Fig. 5. The topology is divided into several parts. The first part is access to the Internet or to local LAN networks within the overall network topology. The second essential part of the topology is the network flow collector, which is implemented using a Cisco Catalyst 3850 switch. Several user devices are connected to this switch, which are represented by the

designation PCX, as well as a server covering the storage of records from the export of network flows. The last part of the topology is a trio of anomaly generators, namely a data exfiltration generator, a port scanning generator and a DDoS attack generator.

The NetFlow v9 protocol will be used for data collection and export on a network with real traffic. From the point of view of the network topology, the measurement was carried out as follows:

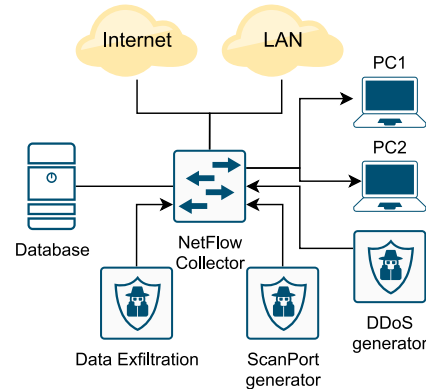


Fig. 5. Overview network topology for dataset collection

Anomaly generators were included in the measurement to create the dataset. Technically, these generators were implemented as follows:

- *Data exfiltration generator* - Data leakage or exfiltration is a relatively complex problem in the case of various companies. Exfiltration of data is most often realized in two ways, in the form of a botnet or malware, or in a physical form. Often, this process is implemented by copying data to various available data repositories, such as OneDrive, Dropbox, etc. In this work, data exfiltration was simulated by moving a large amount of test data to the OneDrive cloud service.
- *Port Scan Generator* - Port scanning is a complex problem due to the unwanted detection of the nature of the available infrastructure and services therein, as well as the detection of possible vulnerabilities. The tool for generating port scans was the Network Mapper tool in short. Nmap.
- *DDoS attack generator* - DDoS attacks constitute a significant number of attacks in the network with the aim of reducing the quality of service, or shutting down the service completely, using the creation of a large number of requests to the server. For these purposes, the XOIC tool was used, which enables the generation of TCP, UDP, HTTP and ICMP DDoS attacks.

Before starting the measurement of the experimental network with the aim of creating a dataset, it is necessary to create a list of required attributes so that subsequent analysis is possible. These attributes are shown in Tab II.

TABLE II
LIST OF NETFLOW ATTRIBUTES IN THE DATASET

Property	Description
IPV4_SRC_ADDR	Source IPv4 address
IPV4_DSC_ADDR	Destination IPv4 address
L4_SRC_PORT	Source IPv4 port
L4_DST_PORT	Destination IPv4 port
PROTOCOL	Protocol identifier
TCP_FLAGS	Cumulative value of TCP flags
IN_BYTES	Incoming number of bits
OUT_BYTES	Outgoing number of bits
IN_PKTS	Incoming number of pakets
OUT_PKTS	Outgoing number of pakets
FLOW_DUR_MS	Flow duration in ms

The ratio of benign traffic to attacked traffic was 8.9 to 1.1 in favor of normal network traffic. The distribution of benign and anomalous traffic is shown in Tab. III.

TABLE III
DISTRIBUTION OF NETWORK TRAFFIC CLASSES IN THE MEASURED DATASET

Class	No	Description
Benign	1620167	Normal network operation without anomalies
Data Exfiltration	61544	Operation during which data exfiltration was carried out
ScanPort	21396	Network scanning to get an overview of available services
DDoS	97852	A class representing a DDoS attack

V. MODEL CLASSIFICATION

The model evaluation methodology will be based on the confusion matrix. Based on the individual cells of the confusion matrix, it is possible to calculate important metrics for evaluating the model. The parameters of the confusion matrix can be defined as follows:

The accuracy of the model is defined as the total proportion of correct predictions out of all predictions and can be defined as follows[8]:

$$Accuracy = \frac{SP + SN}{SP + SN + FP + FN} \quad (3)$$

Sensitivity can be defined as the proportion of correct positive predictions from the total number of positive samples as:

$$Recall, TPR = \frac{SP}{SP + FN} \quad (4)$$

We can define specificity as the share of correct negative predictions from the total number of true negatives as follows:

$$Selectivity, TrueNegativeRate = \frac{SN}{SN + FP} \quad (5)$$

The accuracy of the model can be defined as the proportion of correct positive predictions and the total number of positive predictions according to the relationship:

$$Precision = \frac{SP}{SP + FP} \quad (6)$$

The harmonic mean of precision and sensitivity can be combined into a single metric as the F1 score as follows:

$$F1score = \frac{2SP}{2SP + FP + FN} \quad (7)$$

PR is the proportion of incorrectly classified negative predictions to the total number of negatives. Thus, it is the proportion of false positives among all true negatives and can be taught as:

$$FalsePositiveRate = \frac{SP}{SN + FP} \quad (8)$$

MissRate is the proportion of misclassified positive cases out of the total number of true positive cases. It is thus the proportion of false negatives among all truly positive cases. We can determine failure according to the relationship:

$$FalseNegativeRate = \frac{FN}{SP + FN} \quad (9)$$

From the point of view of machine learning algorithms, the decision tree algorithm appeared to be the most effective with a final accuracy of 96.612% with an F1 score of 95.167% and an AUC parameter of 0.938 for the case of DDoS attack classification. In the case of port scan classification, an accuracy of 95.981% was achieved, with an F1 score of 0.929% and an AUC of 0.929. Last but not least, a model was created for the classification of data exfiltration from the experimental network with an accuracy of 94.881% with an F1 score of 95.478% and an AUC equal to 0.929. These data are shown in Table IV. Therefore, it can be concluded that the model created using decision trees is suitable and accurate.

TABLE IV
COMPARISON OF RESULTS OF CLASSIFICATION ALGORITHMS

Algorithm	Attack Type	Accuracy (%)	F1 (%)	AUC
SVM	DDoS	91,918 ± 0,754	91,891 ± 0,617	0,910 ± 0,021
	Port Scan	90,412 ± 0,711	90,279 ± 0,680	0,903 ± 0,024
	Data Exf.	82,009 ± 0,713	82,107 ± 0,673	0,820 ± 0,014
K-Means	DDoS	90,771 ± 0,687	90,519 ± 0,871	-
	Port Scan	87,991 ± 0,584	87,665 ± 0,826	-
	Data Exf.	85,224 ± 0,647	85,921 ± 0,776	-
NB	DDoS	59,532 ± 1,785	54,127 ± 1,701	0,699 ± 0,011
	Port Scan	58,561 ± 1,862	52, 597 ± 1,522	0,698 ± 0,011
	Data Exf.	54,499 ± 1,92	53, 871 ± 1,445	0,680 ± 0,011
DT	DDoS	96,612 ± 0,412	95,167 ± 1,210	0,938 ± 0,014
	Port Scan	95,981 ± 0,574	94,998 ± 1,119	0,929 ± 0,014
	Data Exf.	94,881 ± 0,699	95,478 ± 1,331	0,929 ± 0,011

VI. ARCHITECTURE FOR CHANGING THE CONFIGURATION OF NETWORK DEVICES

Since the result of network traffic classification is information about ongoing anomalies in network traffic, it is possible to determine the source IP address from which the anomaly originates, the target IP address that represents the target of the attack, and the protocol or port on which the attack is carried out. This information is shown in Table V.

Using the Netmiko library[15], it is then possible to send a constructed command to the network device for setting the ACL, which has the task of stopping the communication between the source and the target of the attack as follows:

TABLE V
SOURCE TABLE FOR CHANGING NETWORK DEVICE CONFIGURATION

id	source_ip_addr	dest_ip_addr	action	port
0	10.10.0.7	192.168.0.10	acl	any

Netmiko is a library written in Python and based on the Paramiko library. This library provides an implementation of the SSHv2 protocol in Python. Compared to the Paramiko library, the Netmiko library adds a multi-level API for communicating and interacting with network devices[15].

```
acl_commands = [
    'ip access-list extended acl',
    'permit tcp any host 192.168.0.10',
    'deny ip any any',
```

It is then possible to see the change in the configuration file in the resulting configuration of the network device. The Netmiko library makes it possible to implement QoS changes on Cisco network devices. For this case, it would be appropriate to implement the Cisco AutoQoS functionality for VoIP as follows[15][16]:

```
commands = [
    'interface gi1/0/10',
    'mls qos trust cos',
    'auto qos voip cisco-phone',
]
```

The result is QoS adaptation in AutoQoS mode for VoIP. The effect on the resulting response during network operation can be seen in Fig. 6. The blue color shows the response of the device without QoS application. In the case of the applied AutoQoS mode, responses are shown in red.

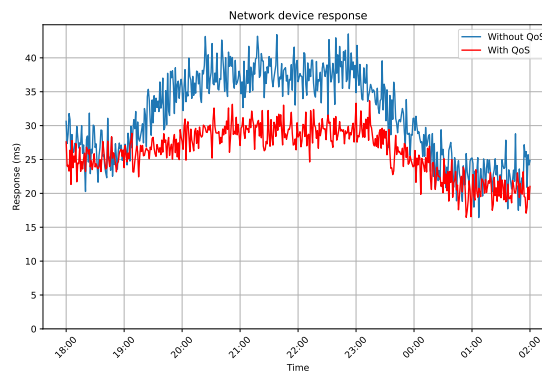


Fig. 6. Effect of QoS setting on the resulting response of VoIP equipment in the network

For a more detailed analysis of the monitored network, it is also possible to analyze other parameters such as ToS or DiffServ, which help to determine the priority of packets and can contribute to a more effective implementation of QoS policies. Another option is to prioritize different forms of internet and network services using statistical or machine learning methods.

VII. CONCLUSION

The work initially deals with the issue of collection of network flows and creates a complex comparative analysis of protocols for collection of network flows. From the available research, the NetFlow v9 protocol was chosen as the tool for collecting network flows. Subsequently, a modified architecture was created for the collection of network flows with the extension of the storage of network flows in the form of a non-relational database ElasticSearch in order to create a comprehensive overview of the measured data as well as to use the great potential in the visualization of the measured data. Using the proposed architecture, a dataset was created on real network traffic. Based on a comparative analysis of available machine learning methods and algorithms, four algorithms were chosen, with which the classification of machine learning models was realized on the measured dataset. The goal of this part of the work was to create an accurate classification model for the detection of anomalies in network operation. The most effective algorithm from the point of view of classification was the decision tree algorithm with a detection accuracy of $96.612\% \pm 0.412\%$. Algorithms that did not achieve comparable results in classification accuracy were optimized. The last part of the presented work deals with the design and implementation of the architecture, which, based on the instructions obtained from the model classification, will create an action in the form of a change in the configuration of the network device in order to eliminate anomalous network traffic and thus increase the security and efficiency of the monitored network.

Future research in this area would benefit from deploying classification models to detect other forms of anomalies in network traffic. It would also be beneficial to deploy the

architecture in real-time network operation and thus create a functional tool for the prevention of attacks and anomalies. A suitable extension would also be the classification of parameters such as QoS with the aim of increasing the efficiency in the network operation of the monitored network.

REFERENCES

- [1] Michael W Lucas. Network flow analysis. No Starch Press, 2010.
- [2] Benoit Claise, Ganesh Sadasivan, Vamsi Valluri, and Martin Djernaes. Cisco systems netflow services export version 9. 2004.
- [3] Liu Bin, Lin Chuang, Qiao Jian, He Jianping, and Peter Ungsunan. A netflow based flow analysis and monitoring system in enterprise networks. *Computer Networks*, 52(5):1074–1092, 2008.
- [4] Cisco, Inc.[online]. NetFlow for Cybersecurity. (English) [cit. 14.11.2021]. <https://www.ciscopress.com/articles/article.asp?p=2812391&seqNum=3>, 2017.
- [5] Rick Hofstede, Pavel Čeleda, Brian Trammell, Idilio Drago, Ramin Sadre, Anna Sperotto, and Aiko Pras. Flow monitoring explained: From packet capture to data analysis with netflow and ipfix. *IEEE Communications Surveys & Tutorials*, 16(4):2037–2064, 2014.
- [6] Sumitha Bhandarkar, Saurabh Jain, and AL Narasimha Reddy. Ltcp: improving the performance of tcp in highspeed networks. *ACM SIGCOMM Computer Communication Review*, 36(1):41–50, 2006.
- [7] Wei-Yin Loh et al. Classification and regression tree methods. *Encyclopedia of statistics in quality and reliability*, 1:315–323, 2008.
- [8] Dhruva Kumar Bhattacharyya and Jugal Kumar Kalita. Network anomaly detection: A machine learning perspective. Chapman and Hall/CRC, 2019.
- [9] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. A practical guide to support vector classification, 2003.
- [10] Marti A. Hearst, Susan T Dumais, Edgar Osuna, John Platt, and Bernhard Scholkopf. Support vector machines. *IEEE Intelligent Systems and their applications*, 13(4):18–28, 1998.
- [11] Benjamin S Duran and Patrick L Odell. Cluster analysis: a survey, volume 100. Springer Science & Business Media, 2013.
- [12] Shi Na, Liu Xumin, and Guan Yong. Research on k-means clustering algorithm: An improved k-means clustering algorithm. In 2010 Third International Symposium on intelligent information technology and security informatics, pages 63–67. Ieee, 2010.
- [13] Christopher Kruegel, Darren Mutz, William Robertson, and Fredrik Valeur. Bayesian event classification for intrusion detection. In 19th Annual Computer Security Applications Conference, 2003. Proceedings., pages 14–23. IEEE, 2003.
- [14] Nahla Ben Amor, Salem Benferhat, and Zied Elouedi. Naive bayes vs decision trees in intrusion detection systems. In Proceedings of the 2004 ACM symposium on Applied computing, pages 420–424, 2004.
- [15] Brendan Choi. Python network automation labs: Ssh paramiko and netmiko. In *Introduction to Python Network Automation: The First Journey*, pages 583– 628. Springer, 2021
- [16] Junghun Park, Hsiao-Rong Tyan, and C-C Jay Kuo. Internet traffic classification for scalable qos provision. In 2006 IEEE International conference on multimedia and expo, pages 1221–1224. IEEE, 2006.

Sémantický model pre popis dátovo-analytických procesov

Juliana Ivančáková

Katedra kybernetiky a umelej inteligencie
Fakulta elektrotechniky a informatiky, TUKE
Košice, Slovenská Republika
juliana.ivancakova@tuke.sk

Peter Bednár

Katedra kybernetiky a umelej inteligencie
Fakulta elektrotechniky a informatiky, TUKE
Košice, Slovenská Republika
peter.bednar@tuke.sk

Abstrakt— Ontológia a ich významná úloha, ktorá je súčasťou sémantického webu, sa ukázali byť dôležitým prvkom v mnohých aplikáciách. Na začiatku je popísaný proces dolovania znalostí a metodológia CRISP-DM, ktorá popisuje jeden z najpoužívanejších modelov pre analytický proces. Následne je predstavená hlavná motivácia práce. Ďalšia kapitola predkladá teoretický prehľad so zameraním sa na sémantický web, jeho technológie a architektúru. Prehľad uvádza problematiku ontológií, ich architektúry a poskytuje prehľad ontológií so zameraním na modelovanie úloh procesu dolovania v dátach a vedeckých experimentov v oblasti bioinformatiky. Hlavná časť tohto článku sa venuje sémantickému modelu, ktorého cieľom je zjednotenie ontológií pre popis dátovo - analytických procesov za účelom navrhnutia modelu, ktorý má slúžiť na zjednodušenie komunikácie medzi dátovým analytikom a doménovým expertom, a ktorý by umožnil automatizáciu úloh objavovania znalostí. Záver tejto publikácie opisuje experiment vykonaný na jednom dátovom súbore s použitím rozličných úloh aplikovaných v dátovej analýze.

Kľúčové slová— Ontológia; Sémantický web; Dátová analytika; OntoDM, Sémantický model

I. ÚVOD

V súčasnej dobe sú ontológie jednou z najpopulárnejších výskumných tém v oblasti reprezentácie znalostí a sú základným a dôležitým prvkom mnohých aplikácií, vrátane sémantiky v oblasti umelej inteligencie a informatiky, manažmentu znalostí, spracovania prirodzeného jazyka, elektronického obchodovania, vyhľadávania informácií, návrhu a integrácie databáz, bioinformatiky, vzdelávania, integrovania inteligentných informácií a sémantického webu. Ontológie majú nielen význam pre aplikácie, kde majú vedomosti dôležité miesto, ale môžu tiež spôsobiť významnú zmenu v aktuálnom obsahu webových stránok. Táto zmena vedie k vzniku sémantického webu, ktorý je definovaný ako konceptuálne štruktúrovaný web v explicitne strojovo-čitateľnom formáte. Na tomto webe sa vytvárajú nové aplikácie založené na ontológiách a znalostných architektúrach..

II. DOLOVANIE V DÁTACH

Dolovanie v dátach je proces zisťovania anomálií, modelov/vzorov a korelácií vo veľkých množinách údajov

na predpovedanie výsledkov. Pomocou širokej škály techník môžeme tieto informácie použiť na dosiahnutie výskumných alebo obchodných cieľov, ako je napr. zvýšenie výnosov, zníženie nákladov, zlepšenie vzťahov so zákazníkmi, zníženie rizík a ďalšie [1]. Proces dolovania v dátach na objavovanie skrytých spojení a predpovedanie budúcich trendov má dlhú históriu. Niekedy sa označuje ako "objavovanie vedomostí v databázach", termín "dolovanie dát" nebol vytvorený až do 90. rokov. Jeho základ sa skladá z troch prepojených vedeckých disciplín: štatistiky, umelej inteligencie a strojového učenia.

Existuje niekoľko metodológií pre popis procesu dolovania v dátach, ako je napríklad CRISP-DM, OSEM a SEMMA. Pre pochopenie procesu dolovania v dátach je vysvetlená metodológia, ktorá sa používa najčastejšie – CRISP-DM.

A. CRISP – DM

Metodológia CRISP-DM (CRoss Industry Standard Proces for Data Mining), predstavuje momentálne najpoužívanejší prístup ako realizovať proces dolovania v dátach. Jeho hlavnou úlohou je poskytovať prehľad o životnom cykle analytického projektu [2]. Pozostáva zo šiestich krokov:

1) Pochopenie problému

V prvom kroku je dôležité pochopenie obchodných cieľov z manažerskeho pohľadu a ich následná transformácia na príslušné úlohy z oblasti dolovania v dátach. Je potrebné si stanoviť kritériá úspechu z obchodného hľadiska, ale aj metriky merania úspechu analytického procesu. Dôležité je preskúmať dostupné zdroje a očakávané náklady.

2) Pochopenie dát

Tento krok zahŕňa primárny zber dát a získavanie popisných charakteristík, ktorých cieľom je pochopiť jednotlivé atribúty, identifikovať prípadné vzťahy medzi nimi a overiť chýbajúce hodnoty v príslušnej dátovej vzorke. Na tento účel často slúžia jednoduché štatistické analýzy a vizualizačné techniky.

3) Príprava dát

V treťom kroku metodológie sa vykonávajú základné operácie nad dátami. Výber, čistenie a transformácia dát sú riadené voľbou analytických metód, ktoré budú na dáta aplikované v nasledujúcom kroku. Často je súčasťou tejto fázy rozdelenie dát na tréningovú a testovaciu množinu.

4) Modelovanie

Štvrtý krok predstavuje aplikáciu vybraných (vhodných) analytických metód na predspracovanú vzorku dát. Výber závisí od typu úlohy, ktorú riešime. Napr. v prípade klasifikácie môžeme aplikovať algoritmy na generovanie rozhodovacích stromov (ID3, C4.5, RF, atď.), v prípade predikcie prichádzajú do úvahy rôzne typy regresie. Vygenerované modely sú vyhodnocované na základe metrick stanovených v prvom kroku celého procesu.

5) Vyhodnotenie

V tomto kroku nastáva samotné vyhodnotenie výsledkov, ktoré boli dosiahnuté na základe zvolených modelov z hľadiska dosiahnutia obchodných cieľov. Posudzuje sa, či boli obchodné ciele splnené a do akej miery. Prebieha tiež dôkladná kontrola celého procesu, či bolo vykonané všetko potrebné.

6) Nasadenie

Posledným krokom, nasadením, je vypracovanie plánu. Navrhuje sa postup pri nasadení výsledkov, teda ako budú nasadené, kto bude vykonávať tento proces a ako často. Proces je monitorovaný a udržiavaný. V závere je vypracovaný sumár dosiahnutých výsledkov a prebehne ich následná prezentácia, teda sa posúdi celý projekt [2].

Podstata metodológií pre popis procesu dolovania v dátach je rovnaká a pre komplexnosť riešenia sa vyžaduje dôsledná komunikácia dátového analytika s doménovým expertom vo fázach, ktoré odpovedajú pochopeniu cieľa a dát a taktiež pri vyhodnocovaní a interpretácii výsledkov. Pre dátového analytika je tiež veľmi dôležité, aby mal znalosti, ktoré mu umožnia pripraviť dáta, vytvoriť modely ML a právne vyhodnotiť ich výsledky. A preto vytvorenia nástroja, ktorí umožní zefektívniť komunikáciu medzi dátovým analytikom a doménovým expertom, je dôležité, tak ako aj sémantické pokrytie fáz v procese objavovania znalostí, ktoré odpovedajú príprave dát, modelovaniu a vyhodnoteniu.

III. MOTIVÁCIA

Hlavným cieľom je zjednotenie ontológií pre popis dátovo – analytických procesov. Výskum je zameraný na návrh ontológií pre vytvorenie formálneho popisu vedeckého výskumu a experimentov, ktoré sa všeobecne uplatňujú v rôznych oblastiach. Navrhovaný sémantický model pozostáva z ontológií pre opis vedeckých výskumov, údajov zozbieraných počas výskumov, rôznych metód na analýzu údajov a pre opis výsledkov analýzy údajov, ktoré mapujú výsledky na koncepty špecifickej domény. Navrhnutý znalostný model má slúžiť na zjednodušenie/zlepšenie komunikácie medzi dátovým analytikom a doménovým expertom, umožniť im efektívne zdieľanie a znovupoužitelnosť znalostí a zároveň zabezpečiť reprodukovateľnosť dátovo analytickým úlohám. Taktiež slúži na podporu automatizácie dátovo – analytických procesov, tak že umožňuje generovanie skriptov predspracovania, modelovania a vyhodnotenia výsledkov dátovej analýzy.

IV. ÚVOD DO SÉMANTICKÉHO WEBU

Sémantický web je vízia, ktorá splňa výzvu prinášať viac webových služieb orientovaných na človeka. Ide o web, kde sú

informácie štruktúrované a uchovávané podľa štandardizovaných pravidiel, čo uľahčuje hľadanie a spracovanie informácií (teda môžeme povedať, že web slúži ako úložisko informácií rôzneho druhu). Existuje veľa údajov, ktoré používame každý deň, a ktoré nie sú súčasťou webu. Sémantický web sa sústreďuje na dve veci. Na spoločné formáty, ktoré slúžia na integráciu a kombináciu údajov získaných z rôznych zdrojov dát, a na jazyk, ktorý zaznamenáva, ako sa dáta týkajú objektov skutočného sveta. To umožňuje, aby človek alebo stroj pristupoval jednotne k prepojeným databázam. Vízia sémantického webu sa realizuje pomocou základného súboru jeho technológií ako sú RDF, RDFS, OWL a SPARQL, ktoré budú ďalej opísané v nasledujúcich podkapitolách.

A. Webová architektúra sémantického webu

Základná architektúra sémantického webu je zložená z technológií rozdelených do viacerých vrstiev. Tento model obsahuje vrstvy, ktoré na seba navzájom nadväzujú a sú na sebe závislé. Evolúcia sémantického webu postupuje krokovo. Každý jeden krok vytvára novú vrstvu na základoch predchádzajúcej, teda efektívnym spôsobom. No netreba zabudnúť že z inžinierskeho pohľadu je tu potrebná štandardizácia. Po dosiahnutí štandardu tu sú ďalšie výskumné skupiny a spoločnosti ktoré ho prijímajú a nečakajú, aby uvideli ktorá alternatívna výskumnej línie bude úspešná. Vlastnosti sémantického webu sú podmienené tým, že spoločnosti ako aj používatelia sú nútení vytvárať zdroje, pridávať obsah webu a tiež používať to všetko ešte počas vývoja sémantického webu. Je nemožné čakať, kým sa naplno uskutoční jeho vízia.

Pri budovaní sémantického webu (keď budujeme novú vrstvu nad predchádzajúcou) nesmú byť porušené tieto dva princípy:

1) Klesajúca kompatibilita

Agenti musia byť schopní "naplno" interpretovať a tiež používať informácie, ktoré sú zapísané vo vrstve nižšie, napr. agent, ktorý si uvedomuje sémantiku OWL má mať plný prístup k informáciám, ktoré sú zapísané v RDF a RDFS.

2) Čiastočné porozumenie smerom nahor

Agenti by mali aspoň čiastočne pracovať z informáciami nachádzajúcich sa vo vyššej vrstve. Napríklad, agent ktorý disponuje iba s RDF a RDFS sémantikou, by mal čiastočne interpretovať znalosti, ktoré sú zapísané v OWL vrstve bez ohľadu na tie prvky, ktoré idú nad vrstvu RDF a RDFS [3].

Na najnižšej vrstve sú štandardy pre kódovanie textu ako napr. kódovanie UNICODE ktoré prevádza písané znaky do jedinečnej číselnej reprezentácie. URI – jedinečne identifikuje akýkoľvek objekt. Jazyk XML je jednou z alternatív, ktorá zabezpečuje prehľadnú štruktúru. RDF je jazyk pre popis zdrojov na webe a je základnou technológiou pre popis štruktúrovaných dát. Jeho podstatou je opis sveta pomocou objektov, vlastností a ich hodnôt. Ontológia je znalostný model, ktorý definuje pojmy a vzťahy medzi nimi, za účelom podpory porozumenia medzi ľuďmi, komunikácie medzi počítačmi a aby bolo možné zjednodušiť návrh znalostne orientovaných systémov. Logická vrstva sa používa na zlepšenie ontologického jazyka a umožňuje odôvodnenie

a zápis aplikačno – špecifických deklaračných znalostí. Vrstva dôkazu a kontroly zahŕňa aktuálne deduktívne procesy ako aj reprezentáciu dôkazov vo webových jazykoch a kontrolu platnosti dôkazov. Vrstva dôvery je reprezentovaná prostredníctvom používania elektronického podpisu a ďalších druhov znalostí, založených na odporúčení plynúceho z dôvery agentov alebo od ratingových a certifikačných agentúr a spotrebiteľských združení. Je to vrstva ktorá sa dosahuje najťažšie.

B. RDF

Štandard RDF (Resource Description Framework), je všeobecný rámec pre opis metadát na webových stránkach alebo "informácií o informáciách" na webových stránkach. Poskytuje interoperabilitu medzi aplikáciami, ktoré si vymieňajú informácie zrozumiteľné strojom na webe. Medzi základné metadáta, ktoré môže RDF popisovať o webových stránkach patria napr. informácie o tom kedy boli stránky aktualizované, aké sú kľúčové slová vyjadrujúce obsah, aké sú práva duševného vlastníctva webovej stránky atď. Jazyk XML je univerzálny metajazyk ktorý sa používa pre definíciu značiek. Bol vyvinutý pod vedením konzorcia World Wide Web Consortium a navrhnutý tak, aby umožňoval vývojárom vytvárať vyhľadávacie nástroje, ktoré sa spoliehajú na metadáta a umožňujú užívateľom internetu ľahšie zdieľať informácie o webových stránkach. RDF je založený na vytváraní ucelených častí - blokov, ktoré pozostávajú z trojíc: objekt - atribút - hodnota. Každá trojica predstavuje výrok, ktorý platí o objekte. To, že RDF je založený na syntaxe XML prináša isté benefity. Avšak je dôležité si uvedomiť, že existuje aj iná syntaktická reprezentácia RDF, ktorej základy netvorí XML. Východiskové pojmy, na ktorých RDF funguje sú: zdroje, vlastnosti a výroky [3].

1) Zdroje

Zdroj [3] je vec, na opis ktorej chceme sformulovať výrok. Môžu to byť základné atribúty knihy ako názov, autor, vydavateľ, prípadne hotel, izba, hosť ak by išlo o hotelový systém. Každý jeden zdroj má jedinečný identifikátor URI (Universal Resource Identifier). Jednou z podmnožín URI je aj forma URL (Unified Resource Locator) teda webová adresa, alebo to môže byť aj nejaký podobný druh identifikátora. Identifikátor môže a nemusí nutne poskytovať prístup k zdroju. URI nie je definovaný len v prípade že ide o webové stránky, ale taktiež napríklad pre telefónne číslo, ISBN, geografickú polohu a pod. Treba si tiež položiť otázku či je možné takto pomocou identifikátora URI jednoznačne identifikovať človeka.

2) Vlastnosti

Vlastnosť [3] je jedinečným typom zdroja. Popisuje súvis medzi zdrojmi, napríklad: kniha je napísaná autorom, bola napísaná v roku, vydaná vydavateľstvom a pod. Vlastnosti, ktoré RDF používa sú identifikované pomocou URI (Uniform Resource Identifier). URI je reťazec znakov, ktorý sa najčastejšie používa na určenie mena zdroja (webový zdroj, rodné číslo, ISBN knihy, a pod.). URL (Uniform Resource Locator) určuje doménovú adresu, napr. servera, umiestnenie zdroja na serveri a protokol na sprístupnenie zdroja. Pri identifikácii vecí a vzťahov medzi nimi, pomocou URI, je

možné dosiahnuť jednoznačnú názvovú schému a taktiež minimalizovať problém s homonymami.

3) Výroky

Výroky [3] vieme využiť na určenie vlastností zdrojov. Je to skupina troch vecí: objekt - vlastnosť - hodnota. Objekt vie byť prezentovaný zdrojom a hodnota môže byť buď zdroj alebo literál, teda atomická hodnota.

V kombinácii tieto technológie definujú tri komponenty, ktoré majú HTML a ľubovoľný výmenný jazyk spoločné: syntax (RDF/XML), dátový model (RDF) a sémantiku (RDFS). Syntax nám hovorí, ako zapísať dáta. Dátový model nám uvádza štruktúru alebo organizáciu údajov. Sémantika nám hovorí, ako tieto údaje interpretovať. Pre sémantický web je však potrebné niečo bohatšie. Je potrebný dátový model, ktorý možno použiť viacerými aplikáciami, a to nielen na opis dokumentov pre ľudí, ale aj na opis konkrétnych informácií o aplikácii. Tento dátový model musí byť nezávislý od domény. Okrem flexibilného dátového modelu je potrebný aj mechanizmus na priradenie sémantiky k informáciám reprezentovaným pomocou tohto dátového modelu [4].

C. OWL

OWL je navrhnutý ako spoločný jazyk pre reprezentáciu ontológií. Je to rozšírenie RDF schémy vyvinuté ako štandardný jazyk pre reprezentáciu ontológií, ktorý by mal umožniť používateľovi písať explicitné, formálne koncepty doménových modelov, a preto bola pri dizajne OWL udelená najvyššia priorita pre rozšíriteľnosť, modifikovateľnosť a interoperabilitu. Súčasne sa OWL snaží dosiahnuť dobrý kompromis medzi škálovateľnosťou a expresívnou silou.

OWL je bohatý slovníkový popisný jazyk pre definíciu vlastností a tried ako aj relácií medzi triedami danej domény, disponujúci bohatým typovaním vlastností a taktiež charakteristikami vlastností a vymenovaných tried [5]. Hlavné požiadavky na jazyky ontológií sú nasledovné:

- Dobre definovaná syntax – táto podmienka je veľmi dôležitá pre strojové spracovanie informácií.
- Formálna sémantika – jej úlohou je presne popísať význam znalostí.
- Konvenčnosť vyjadrení.
- Efektívna podpora odvodzovania – už samotná sémantika je predpokladom odvodzovania. To môže byť manuálne alebo automatické, a práve automatickým odvodzovaním je možné kontrolovať oveľa väčšie množstvo tried.
- Dostačujúca vyjadrovacia sila.

D. SPARQL

Skratka SPARQL znamená Simple Protocol RDF Query Language, SPARQL je teda definovaný ako štandardný dopytovací jazyk a protokol prístupu k dátam pre použitie s dátovým modelom RDF (Resource Description Framework). Funguje pre ľubovoľný zdroj údajov, ktorý je možné mapovať na RDF. Na vykonanie dopytu SPARQL je potrebný softvér. Najbežnejší, ktorý to robí, sa nazýva triple store, ktorý je v podstate databázou RDF. Akonáhle sú dáta načítané do triple store, dá sa to spýtať odoslaním SPARQL dopytov

pomocou protokolu SPARQL. Každý triple store poskytuje to, čo sa nazýva koncový bod, kde môžu byť odoslané dopyty SPARQL. Dôležitým bodom je to, že klienti posielajú dopyty na koncový bod pomocou protokolu HTTP. Pretože SPARQL používa štandardné webové technológie, je možné nájsť na webe početné koncové body SPARQL. Tieto koncové body umožňujú prístup k veľkému množstvu údajov [4].

V. ONTOLOGIE

Pojem ontológie má dvojaký význam: dá sa na neho pozeráť ako na pojem z filozofie alebo ako na pojem z počítačovej vedy. Filozofická perspektíva sa zameriava na kategorickú analýzu a odpovedá na otázky ako: "Aké sú entity na svete?" a "Aké sú kategórie subjektov?". V tomto kontexte je ontológia tradične definovaná ako "filozofická štúdia toho, čo existuje: štúdium druhov subjektov v skutočnosti a vzťahov medzi nimi". Naopak, pojem ontológie v oblasti informatiky je mladší. Existuje mnoho definícií ontológií v informatike a jedna z prvých definícií zaviedol Neches: "Ontológia definuje základné pojmy a vzťahy obsahujúce slovnú zásobu doménovej oblasti, ako aj pravidlá pre kombináciu pojmov a vzťahov na definovanie rozšírenia slovnej zásoby" [6].

V tejto kapitole je prezentovaný prehľad ontologických architektúr a stručne definované niektoré ontológie.

A. Architektúra ontológií

Architektúra ontológií [7] rozdeľuje typy ontológií na tri základné úrovne – Ontológie na vyššej úrovni, Ontológie na strednej úrovni a Doménové ontológie – podľa všeobecnosti pojmov reprezentovaných v ontológií. Najšpecifickejšie sú ontológie na spodnej úrovni, ktoré sú zamerané iba na jednu oblasť použitia (doménu). Naopak, ontológie na vrchnej úrovni obsahujú všeobecné pojmy používané vo viacerých oblastiach. Ontológie na strednej úrovni zvyčajne mapujú a prepájajú ontológie na nižšej a vyššej úrovni.

B. Ontológie pre popis úloh dolovania v dátach

Zatiaľ čo proces objavovania znalostí v databázach a proces dolovania v dátach získali v posledných rokoch veľkú popularitu a úspech, existuje výrazný nedostatok všeobecne akceptovaného rámca, ktorý by pokrýval a zjednotil doménu procesu dolovania v dátach. Súčasný nedostatok takéhoto rámca je vnímaný ako prekážka ďalšieho rozvoja tejto oblasti. Jedným krokom k vytvoreniu všeobecného rámca pre dolovanie dát je budovanie ontológie pre proces dolovania v dátach.

Tak ako už bolo spomínané, zlepšenie komunikácie medzi dátovým analytikom a doménovým expertom je veľmi aktuálna téma, spolu s popisom jednotlivých krokov procesu objavovania znalostí pre lepšiu interpretáciu v rámci sémantiky.

1) Ontológia EXPO a LABORS

EXPO ontológia [8] predstavuje návrh na formalizáciu vedeckých experimentov a jej rozšírením je ontológia LABORS [9], ktorej cieľom je zautomatizovanie vykonávania vedeckých experimentov.

2) Ontológie pre popis biomedicínskych experimentov

Ontológia BFO [10] sa zameriava na úlohu poskytovania ontológie na vyššej úrovni, ktorá sa môže použiť na podporu ontológií vyvinutých pre vedecký výskum, napríklad v biomedicíne. OBO pokrýva širokú škálu biomedicínskych domén vrátane ontológií anatómie, vývoja a ochorení pre množstvo kľúčových organizmov; ontológie pre popis biologických sekvencií, funkcií a procesov v organizmoch, ontológie biochémie, typov buniek a správania [11]. RO je ontológia vzťahov, ktoré sa používajú medzi entitami v biomedicínskych ontológiách [12]. Cieľom OBI ontológie je poskytnúť štandard pre reprezentáciu biologických a biomedicínskych vyšetrení. OBI je plne v súlade s existujúcimi formalizáciami v biomedicínskych oblastiach. OBI ontológia používa rigidnú logiku a sémantiku, pretože používa hornú úroveň ontológie (BFO) a vzťahy RO na definovanie tried vyššej úrovne a súbor vzťahov [13].

3) DSO

Ontológia DSO (Data Science ontology) [14], je báza znalostí o dátovej vede, ktorej cieľom je katalogizovať koncepty dátovej vedy, sémanticky anotovať niektoré softvérové balíky pre dátovú vedu, teda je zameraná na počítačové programovanie. Pojmy pre túto ontológiu sú extrahované zo štatistiky, strojového učenia a softvérového inžinierstva zameraného na dátovú vedu. Okrem katalogizácie a usporiadania pojmov dátovej vedy ontológia poskytuje sémantické anotácie bežne používaných softvérových knižníc pre dátovú vedu, ako sú pandas, scikit-learn a statsmodels. Anotácie mapujú typy a funkcie knižníc na univerzálne pojmy ontológie.

4) DMOP a DMWF

Ďalšie predstavené ontológie rozširuje definíciu niektorých pojmov zavedených v OntoDM. DMOP ontológia [15] sa zaoberá najmä podrobným popisom algoritmov pre úlohy dolovania v dátach aj spolu s ich princípom, ako je napr. popis použitej numerickej optimalizačnej metódy, chybovej funkcie alebo regularizácie učenia. DMOP pokrýva 3 fázy metodológie CRISP-DM. DMWF ontológia [16] definuje dátové operátory, ktoré sú použiteľné v procese objavovania znalostí v dátach vo fázach ako sú predspracovanie údajov, modelovanie a vyhodnocovanie. Operátory sú popísané ako ich vstupy, výstupy, predpoklady a účinky. Tieto vstupné a výstupné podmienky sú definované ako logické výrazy v jazyku SWRL, pokiaľ ide o možné využitie automatického odvodzovania pri skladaní pracovných tokov pri analýze údajov.

5) EXPOSÉ

Jedným z najvšeobecnejších návrhov, ktoré rozširujú OntoDM, je ontológia Exposé [17], ktorá rozširuje fázu metodológie CRISP-DM procesu analýzy údajov na úroveň experimentov s cieľom zabezpečiť napr. opakovanú použiteľnosť postupov pri analýze údajov a reprodukovateľnosť výsledkov. Okrem konceptualizácie poskytuje aj jazyk na opis experimentov založený na XML, ktorý umožňuje publikovať a zdieľať opis experimentov na webe v strojovo čitateľnom formáte.

6) *OntoDT*

Panov a kolektív popisujú OntoDT, ontológiu na reprezentáciu znalostí o dátových typoch. Táto ontológia definuje základné entity, ako sú dátové typy a ich vlastnosti, špecifikácie, charakterizačné operácie a taxonómia dátových typov [18].

7) *INPUT Ontology*

V [19] Tianxing a kolektív predstavuje meta-miningovú ontológiu, ktorá sa používa na vytvorenie doménovo orientovanej ontológie. Hlavným cieľom vytvorenia ontológie INPUT je lepšie pochopenie údajov a obchodných cieľov a jej použitie ako vstupného rozhrania pre používateľské dotazy.

8) *OntoDM*

Pre popis úloh dolovania dát je použitá ontológia OntoDM, ktorá zameraná na poskytovanie štruktúrovanej slovnej zásoby subjektov, ktoré sú dostatočné na popis scenárov a pracovných tokov pri získavaní dát. Prvá verzia OntoDM využívala filozofiu ontológie vedeckých experimentov (EXPO) a ontológiu biomedicínskych vyšetrení (OBI) na identifikáciu a organizáciu subjektov v hierarchii triedy is-a. Štruktúra ontológie bola zosúladená s najvyššou úrovňou štruktúry OBI ontológie. Bol rozšírený súbor vzťahov použitých v počiatočnej verzii so vzťahmi definovanými v OBI ontológii, aby bolo možné vyjadriť vzťahy medzi informačnými entitami, subjektmi (realizovateľnými v procese) a procesmi. Taktiež boli rozšírené triedy OBI o špecifické triedy na dolovanie dát na popis komplexných entít [20].

9) *Viso*

V rámci predspracovania a modelovania by sa nemalo zabúdať na vizualizáciu údajov a grafické znázornenie algoritmov/modelov pre úlohy dolovania v dátach. Na opis takýchto pojmov existuje ontológia VISO, ktorá formálne modeluje pojmy a fakty špecifické pre vizualizácie. Výhodami tejto ontológie, dosiahnutými aj vďaka osvedčeným sémantickým štandardom, ako sú RDF a OWL, sú technická interoperabilita, podpora spoločného porozumenia medzi interdisciplinárnymi stranami v procese vizualizácie a možnosť odvodzovať nové znalosti z existujúcich faktov [21].

VI. NÁSTROJE NA AUTOMATIZÁCIU

Automatické generovanie kódu je proces, pri ktorom sa používajú algoritmy na vytvorenie kódu pre určitú úlohu alebo program. Existujú rôzne techniky a prístupy na generovanie kódu, ktoré sa líšia v závislosti od typu problému a použitých nástrojov. Jednou z najjednoduchších foriem generovania kódu je použitie generátora kódu, ktorý umožňuje používateľom vytvoriť kód pomocou jednoduchej vizuálnej reprezentácie, napríklad blokových diagramov. Takéto nástroje sú často používané pre vytváranie kódu pre jednoduché aplikácie, ktoré nevyžadujú veľa programovacích znalostí. Na druhej strane, pri sofistikovanejších aplikáciách a programoch sa na generovanie kódu používajú algoritmy strojového učenia a umelá inteligencia. Tieto algoritmy sú schopné naučiť sa z veľkého množstva existujúceho kódu a vytvárať nový kód podľa podobnosti s existujúcim kódom.

V nasledujúcich podkapitolách sú uvedené príklady niektorých nástrojov na automatické generovanie zdrojového kódu pre dátovo-analytické úlohy.

A. *H2O AI*

H2O.AI [22] je popredná platforma pre umelú inteligenciu a strojové učenie, ktorá poskytuje open source a komerčné softvérové riešenia pre dátových vedcov a podniky. Platforma umožňuje používateľom vytvárať a nasadzovať modely strojového učenia a ponúka súbor nástrojov na analýzu údajov, vizualizáciu údajov a interpretáciu modelov.

B. *Amazon SageMaker Studio*

Amazon SageMaker Studio [23], je plne integrované vývojové prostredie (IDE) pre strojové učenie (ML), ktoré zjednodušuje celý proces vývoja ML tým, že poskytuje webové rozhranie, ktoré umožňuje jednoducho spravovať celý životný cyklus ML na jednom mieste od predspracovania dát až po vytváranie, tréning a nasadzovanie modelov ML, čím sa produktivita dátovej analýzy zvyšuje až 10 násobne. Je možné rýchlo nahrávať údaje, vytvárať nové notebooky, trénovať a ladiť modely, prechádzať medzi jednotlivými krokmi a upravovať experimenty, bezproblémovo spolupracovať v rámci organizácie a nasadzovať modely do produkcie bez toho, aby musel byť opustené prostredie SageMaker Studio.

C. *Google AutoML*

Google AutoML (Auto Machine Learning) [24] je balík produktov strojového učenia vyvinutý spoločnosťou Google, ktorý umožňuje vývojárom s obmedzenými odbornými znalosťami v oblasti strojového učenia vytvárať vlastné modely bez nutnosti písania zložitého zdrojového kódu. AutoML poskytuje používateľsky grafické rozhranie na vytváranie vlastných modelov strojového učenia bez toho, aby sa vyžadovali hlboké znalosti algoritmov strojového učenia a programovacích jazykov.

VII. SÉMANTICKÝ MODEL PRE POPIS DÁTOVO-ANALYTICKÝCH PROCESOV

Hlavnou časťou článku je vytvorenie sémantického modelu na popis dátovo-analytických procesov. V rámci výskumného procesu v predchádzajúcich prácach [25] [26] boli definované už niektoré rámce, na základe ktorých bol navrhovaný sémantický model rozdelený do niekoľkých nasledujúcich modulov:

A. *Doménové koncepty*

Doménové koncepty sú popísané pomocou metamodelu SKOS [27], ktorý špecifikuje názov, opis a definície pojmov ktoré môžu byť uvedené vo viacerých jazykových prekladoch. Koncepty možno hierarchicky usporiadať vo forme tezauru/taxonómie pomocou vzťahov `skos#broader/skos#narrower`. Je možné definovať aj polyhierarchické schémy, kedy jeden koncept môže byť súčasťou viacerých hierarchií. Okrem hierarchického usporiadania možno pojmy spájať aj asociatívne pomocou vzťahu `skos#related`. Zadeňovaný spoločný slovník doménových

konceptov možno použiť aj ako klasifikačnú schému na usporiadanie rôznych typov dokumentov, ktoré vstupujú do analýzy údajov ako doménová dokumentácia, ktorú vytvárajú dátoví analytici na zdokumentovanie samotného procesu dátovej analýzy, údajov a dosiahnutých výsledkov. Samotné typy dokumentov možno špecifikovať ako klasifikačnú taxonómiu v SKOS. Okrem naratívneho opisu je v niektorých doménach vhodné vysvetliť existujúce pojmy aj pomocou rôznych diagramov, schém a iných typov grafických zápisov (napr. pomocou diagramov BPMN na modelovanie obchodných procesov alebo procesných diagramov na vizualizáciu výrobných procesov v oblasti Priemyslu 4.0 a pod.) V tomto prípade sú jednotlivé grafické prvky (napr. blok činností v diagrame BPMN) opísané ako samostatné koncepty SKOS, ktoré sú prepojené s daným prvkom, aby bola možná obojsmerná navigácia medzi grafickým zápisom a súborom sémantických konceptov. Navrhovaný formalizmus však nedefinuje, ako sú tieto prepojenia reprezentované buď v sémantickej reprezentácii, alebo vložené priamo do formátu grafického zápisu.

B. Dátové elementy a ukazovatele výkonnosti

Dátové prvky sú špecifikované na dvoch úrovniach: logickej a fyzickej. Logické dátové prvky sa využívajú na opis každého atribútu vstupných údajov, ktorý je expertom na danú oblasť vybraný ako relevantný pre riešenie danej úlohy analýzy údajov, alebo na opis všetkých výstupných údajov vytvorených počas procesu analýzy vrátane predikovaných hodnôt analytických modelov alebo hodnôt merateľných technických ukazovateľov výkonnosti. Logická reprezentácia definuje metadáta pre každý atribút údajov, ktoré zahŕňajú jeho názov a definíciu v prirodzenom jazyku, logický typ údajov, napríklad či ide o nominálne, ordinálne, číselné údaje, skalárne, vektorové veličiny, priestorovo usporiadané údaje, časové rady atď. Logické atribúty údajov môžu byť prepojené vzájomnými závislosťami, ktoré sú reprezentované triedou *Dependency*. Trieda *Dependency* definuje vzťah medzi jedným závislým dátovým prvkom a jedným alebo viacerými nezávislými prvkami. Závislosť môže byť opísaná v texte alebo špecifikovaná formálne pomocou známeho fyzikálneho alebo ekonomického modelu atď. Závislosti možno ďalej špecifikovať podľa základného typu, napríklad vyjadriť, či je hodnota závislého prvku odvodená od nezávislého prvku prostredníctvom transformácie, alebo je agregáciou viacerých nezávislých prvkov atď. Fyzické umiestnenie údajov počas analýzy predstavujú fyzické dátové prvky a platí, že jeden logický atribút môže mať viacero fyzických realizácií. *PhysicalDataElements* sú priradené fyzické dátové formáty a typy. Fyzický typ vychádza z normy ISO na opis formátov údajov v softvérových systémoch a môže byť atomický (reálne/celočíselné číslo, logická hodnota, reťazec) alebo zložený (záznam s dátovými poľami, zoznam alebo neusporiadaná množina). Konkrétne umiestnenie údajov sa definuje pomocou IRI, ktorý predstavuje jedinečný identifikátor v softvérovom prostredí na analýzu údajov alebo URL na umiestnenie na webe. Viacero fyzických dátových prvkov možno spojiť do jedného súboru dátových prvkov, na ktorý sa je možné odkázať prostredníctvom IRI. Výsledky

dátovej analýzy sú kvantitatívne opísané merateľnými ukazovateľmi výkonnosti, ktoré sú definované ako špeciálny podtyp *DataElement*. Podobne ako pri opise vstupných údajov sa ukazovatele výkonnosti delia na dve podtriedy: *DomainIndicator* a *TechnicalIndicator*. Doménové ukazovatele bežne zavádza doménový expert na hodnotenie výsledkov z obchodného hľadiska. Zahŕňajú ukazovatele, ktoré špecifikujú napríklad finančné náklady/úspory, spotrebu energie alebo zdrojov, vplyv na životné prostredie atď. Technické ukazovatele zahŕňajú štatistiky vyjadrujúce výkonnosť modelu dolovania údajov, ako je presnosť, špecifickosť, citlivosť atď. odhadnuté na testovacom alebo validačnom súbore údajov alebo pomocou krížovej validácie. Okrem ukazovateľov výkonnosti technické ukazovatele zahŕňajú aj rôzne ukazovatele vyjadrujúce zložitnosť alebo interpretovateľnosť modelov úloh dolovania v dátach (napr. počet číselných parametrov, počet klasifikačných alebo asociačných pravidiel, počet zhlukov atď.). Technické ukazovatele dátový analytik formálne mapuje na doménové ukazovatele špecifikovaním závislostí reprezentovaných triedou *Dependency*, čo umožňuje späťne určiť, ako technická kvalita modelov dolovania v dátach kvantitatívne ovplyvňuje požadovanú kvalitu obchodných cieľov.

C. Algoritmy a dátovo-analytické modely

Koncepty popisujúce metódy a algoritmy dátovej analýzy. Používanie algoritmov pri dátovej analýze je určené definíciou úlohy dolovania v dátach, ktorú špecifikuje dátový analytik. Úloha je definovaná súborom obmedzení (v modely je tento súbor reprezentovaný triedou *Constraint*), ktoré pomocou logických výrazov ohraničujú požadované vlastnosti výsledného modelu úlohy dolovania v dátach. Tieto obmedzenia zahŕňajú vlastnosti vstupných údajov (ako sú typy údajov a výskyt chýbajúcich, alebo extrémnych hodnôt), typ požadovaného modelu, ako je klasifikácia, regresia, zhlukovanie, asociačné pravidlá, detekcia anomálií atď. Tieto obmedzenia sa ďalej delia na tvrdé obmedzenia, ktoré musia byť pri riešení úlohy úplne splnené, a mäkké obmedzenia, ktoré by sa mali pri riešení zohľadniť, ale nemusia byť nevyhnutne bezpodmienečne splnené. Hlavným konceptom, ktorý pokrýva všetky operácie nad dátami je trieda *Operator*. Algoritmy strojového učenia sú reprezentované konceptom *Algorithm*, ktorý je špeciálnym typom operátora so vstupnými a výstupnými údajmi v podobe modelu dolovania údajov. Samotný *Model* je operátor, ktorý možno aplikovať na vstupné údaje na vyhodnotenie. Výstupom operátora *Model* sú dátové prvky s predpovedanými hodnotami a s voliteľnými dodatočnými metadátami, ako sú skóre spoľahlivosti, identifikátory klasifikačných pravidiel alebo váhy vstupných hodnôt. Okrem výstupného modelu môže mať algoritmus súbor hyperparametrov, ktoré je potrebné nastaviť pred jeho vykonaním. Každý parameter má definovaný typ údajov a preddefinovanú hodnotu. Špeciálnym typom algoritmu je trieda *MetaLearningSchema* [28], ktorá interne optimalizuje nastavenia parametrov pre danú tréningovú množinu dát a základný algoritmus učenia, alebo vyberá najlepší algoritmus z množiny algoritmov pre dané tréningové dáta.

D. Procesný model pre dátovo-analytické pracovné toky

Navrhovaný procesný model možno použiť na automatické generovanie pracovných tokov pre úlohy dátovej analýzy a tiež na formálny opis existujúcich skriptov dátovej analýzy, čím sa zabezpečí ich reprodukovateľnosť a opätovná použiteľnosť. Je navrhnutý podobne ako procesný model pre choreografiu a orchestráciu webových služieb, pričom stav je reprezentovaný inštanciami priradenými k zdieľaným premenným. Proces sa skladá z uzlov, ktoré predstavujú jednotlivé operácie (trieda *Operator*) na predspracovanie, modelovanie a vyhodnocovanie údajov alebo riadiace bloky, ako sú vetvy, cykly, paralelné vykonávanie a synchronizácia (trieda *ControlNode*). Uzly sú v pracovnom postupe prepojené hranami reprezentovanými triedou *GuardedTransition*, ktorá predstavuje podmienené prechody medzi uzlami. Operátory sú opísané funkčne, pričom vstupy, výstupy, predpoklady a dopady sú definované ako logické výrazy. Priebeh procesu sa určuje spätným reťazením efektov a predpokladov, pričom sa požadované cieľové efekty preberajú zo špecifikácie úlohy dolovania v dátach. Dodatočné obmedzenia môžu byť zahrnuté v rámci algoritmov použitých v pracovnom toku, napríklad možnosť pracovať len s určitými typmi atribútov alebo spracovať chýbajúce hodnoty. Dôležitou súčasťou tohto sémantického modelu je formalizmus použitý na opis logických výrazov v špecifikácii vstupov, výstupov, predpokladov a postpodmienok operátora. Formalizmus je rozdelený na varianty s postupne sa zvyšujúcou expresivitou, čo umožňuje zvoliť kompromis medzi zložitou a expresivitou a zjednodušiť implementáciu automatického plánovania. Úplná špecifikácia bola založená na formalizme WSML (Web Service Modelling Language) [29], ktorý kombinuje konštrukcie z deskriptívnej logiky a logického programovania. Výrazy WSML pozostávajú z logických premenných, funkčných symbolov, logických operátorov (a, alebo, klasická negácia a negácia pri zlyhaní) a kvantifikátorov (existenčný a univerzálny). Náš súčasný návrh podstatne znížil expresivitu výrazov WSML s cieľom ešte viac zjednodušiť formalizmus a implementáciu automatickej metódy na skladanie procesov.

E. Koncepty pre vizualizáciu dát.

Tieto koncepty opisujú vizualizáciu v procesoch úloh dolovania v dátach. Kľúčové koncepty, sú *Algorithm*, ktorý je definovaný ako typ operátora a závisí od vstupných premenných, a *VisualizationMethod*, tento koncept je definovaný ako operátor, ktorý prijíma údaje alebo model ako vstup a výstupom je vizualizácia vo forme grafu.

VIII. EXPERIMENTY A VYHODNOTENIE

Navrhovaný sémantický model je určený na formálny opis procesov dátovej analýzy, ktorý zabezpečuje ich reprodukovateľnosť a interoperabilitu, a na automatizáciu analytických procesov. Na vyhodnotenie navrhovaného prístupu bol sémantický model aplikovaný na jednu prípadovú štúdiu z oblasti medicíny. Pri vyhodnocovaní boli najprv ručne anotované všetky skripty použité na predbežné spracovanie, modelovanie a vyhodnocovanie modelov strojového učenia s

konceptmi z navrhovaného sémantického modelu; a potom bol porovnaný kód ručne vytvorený dátovými analytikmi s kódom automaticky vygenerovaným pomocou znalostného grafu.

Na porovnanie kódu boli definované najmä metriky založené na pokrytí kódu [30], a to počet riadkov kódu (vrátane riadiacich príkazov pre vetvenie a cyklus), počet presných zhodných operátorov a počet operátorov s čiastočne zhodnými parametrami. Sémantické obmedzenia boli definované tak, aby sa výsledný model dolovania údajov presne zhodoval s výsledkami expertného kódu bez dodatočnej automatickej optimalizácie pomocou schémy metaučenia. Okrem vyhodnotenia pokrytia kódu sa testovala aj presnosť naučených modelov, pričom medzi vygenerovaným a pôvodným kódom neboli zistené výrazné rozdiely v kvalite naučených modelov [31].

Existuje niekoľko typov metrických pokrytia kódu, niektoré z nich sú vysvetlené nasledovne:

1) Pokrytie riadkov (*Line coverage*):

Meranie percentuálneho podielu riadkov kódu, ktoré boli vykonané počas testovania.

2) Pokrytie vetvení (*Branch coverage*):

Percentuálny podiel vetiev v kóde, ktoré boli vykonané počas testovania. Vetvy sú obvykle spojené s podmienenými príkazmi (if, else, atď.).

3) Pokrytie funkcií (*Function coverage*):

Percentuálny podiel definovaných funkcií v kóde, ktoré boli volané počas testovania.

4) Pokrytie príkazov (*Statement coverage*):

Percentuálny podiel príkazov a algoritmov v kóde, ktoré boli vykonané počas testovania.

5) Pokrytie podmienok (*Condition coverage*):

Percentuálny podiel podmienok (ako napríklad podmienený príkaz if) v kóde, ktoré boli vykonané počas testovania.

6) Pokrytie vstupov a výstupov funkcií (*Entry/Exit coverage*):

Percentuálny podiel vstupov a výstupov funkcií v kóde, ktoré boli vykonané počas testovania.

Evaluácia sémantického modelu bola vykonaná na rôznych úlohách, zameraných na dátovú analýzu. Jedna z prípadových štúdií, uvedená v nasledujúcej podkapitole, popisuje medicínsku doménu, kde bolo vykonané prediktívne dolovanie – binárna klasifikácia.

A. Vyhodnotenie sémantického modelu

Evaluácia bola vykonaná na binárnej klasifikácii aplikovanej na diagnostiku rakoviny prsníka. Príznaky boli vypočítané z digitalizovaného obrazu aspirátu tenkou ihlou (FNA) z prsníka. Dáta opisujú charakteristiky bunkových jadier prítomných na snímke, ako je napríklad polomer, textúra, obvod, plocha atď. Všetky znaky boli reálne hodnotené číselné atribúty bez chýbajúcich hodnôt. Niektoré atribúty boli odvodené zo zdrojového atribútu, ako agregáčnej funkcie (napr. celková priemerná symetria vypočítaná zo symetrie každej anotovanej bunky atď.). Súbor údajov [32] obsahuje spolu 32 atribútov. Zdrojový kód bol vytvorený dátovými analytikmi [33] v jazyku R a zahŕňa

predspracovanie, modelovanie a vyhodnotenie šiestich algoritmov objavovania znalostí použitých na klasifikáciu – rozhodovacie stromy, náhodné lesy, k-najbližších susedov, naivný Bayesov klasifikátor (s normálnym rozdelením pravdepodobností atribútov), SVM a algoritmus neurónových sietí. Modely boli vyhodnotené pomocou štandardnej metriky - celkovej presnosti modelu pre pozitívnu triedu.

Pri vyhodnocovaní boli definované sémantické obmedzenia pre konečný model dolovania dát tak, aby presne zodpovedal výsledkom expertného kódu bez dodatočnej automatickej optimalizácie pomocou schémy metaučenia. Porovnávaný je zdrojový kód na základe metrick pokrytia kódu, medzi kódom expertov – dátových analytikov a automaticky generovaným kódom. Výsledky sú zhrnuté v tabuľke 1.

TABUĽKA 1 VÝSLEDKY EXPERIMENTU

Metrika	Kód experta	Generovaný kód	Pokrytie %
Počet riadkov	328	-	-
Počet riadkov s operátormi	126	-	-
Počet algoritmov	6	6	100
Počet vizualizácií	6	4	67
Počet vyhodnocovacích metrick	1	1	1
Počet argumentov	26	21	81

Skripty pre tento príklad boli vyvinuté ako aplikácia RShiny, kde veľká časť kódu súvisí s programovaním používateľského rozhrania, ktoré nie je relevantné pre úlohu dátovej analýzy. Zo zvyšného kódu 126 riadkov obsahuje operácie nad dátami a modelmi. Automaticky vygenerovaný kód pokrýva všetkých 6 hodnotených algoritmov (rozhodovací stromy, náhodné lesy, naivný Bayesov klasifikátor, k-najbližších susedov, neurónové siete a SVM). Niektoré parametre (t. j. počet rozhodovacích stromov v náhodnom lese a počet susedov pre kNN) boli v pôvodnom kóde optimalizované pomocou jednoduchého cyklu. Táto časť optimalizácie bola v sémantickom modeli nahradená ekvivalentným operátorom vyhľadávania v sieti. Modely sa vyhodnocovali pomocou rovnakého súboru technických ukazovateľov KPI pre binárnu klasifikáciu (presnosť a kontingenčná tabuľka). Pokrytie vizualizačných operátorov zodpovedalo 67 %, pričom dve vizualizačné metódy neboli pokryté súčasným modelom. Nepokryté vizualizačné metódy vykazujú závislosť medzi vybraným technickým KPI (napr. presnosť) a jedným z parametrov algoritmu (napr. počet stromov v algoritme náhodného lesa). Pokrytie premenných zahŕňa aj všetky aliasy premenných vo východiskovom kóde (t. j. keď je rovnaká hodnota údajov priradená dvom premenným s rôznym názvom). Celkovo má sémantický model oveľa bohatší a konzistentnejší súbor

jedinečných premenných. Všetky vetvenia v pôvodnom kóde boli pokryté, pretože súviseli s výberom modelu. Okrem toho pôvodný kód obsahuje definíciu dvoch pomocných funkcií. Obe sa používali ako operátory predspracovania údajov a obe boli vo vygenerovanom kóde nahradené ekvivalentnými funkciami zo štandardných balíkov R.

POĎAKOVANIE

Táto publikácia vznikla vďaka podpore projektov APVV-16-0213 a VEGA 1/0685/21.

REFERENCIE

- [1] Data Mining: What it is and why it matters, online, [cit. 2019-01-17]. Dostupné na internete: <https://www.sas.com/en_us/insights/analytics/data-mining.html#dmhistory>
- [2] J. Smatana, Metodika CRISP-DM ako proces získavania znalostí z databáz [online]. [cit. 2017-04-25]. Dostupné na internete: <<http://smartvia.sk/?p=469>>.
- [3] K. Machová, Od adaptívneho k sémantickému webu, 132 s., (*Technická univerzita v Košiciach*, 2013), ISBN: 978-80-553-1489-1
- [4] G. Antoniou, F. Van Harmelen, A semantic Web primer. 2nd ed. Cambridge, Mass: MIT Press; 2008. 264 s. (Cooperative information systems).
- [5] G. Antoniou, F. Van Harmelen, Web Ontology Language: OWL, online, [cit. 2019-01-13]. Dostupné na internete: <<https://www.math.vu.nl/~frankh/postscript/OntoHandbook03OWL.pdf>>
- [6] R. Neches, R.E. Fikes, T. Finin, T. Gruber, R. Patil, T. Senator, and W.R. Swartout, (1991). Enabling technology for knowledge sharing. *AI Magazine* 12, 36.
- [7] P. Panov, A modular ontology of data mining, Doctoral dissertation, *Jožef Stefan Inter-national Postgraduate School*, Ljubljana, 2012.
- [8] L.N. Soldatova, and R.D. King, An ontology of scientific experiments. *Journal of The Royal Society Interface* 3, 795–803, 2006.
- [9] R.D. King, et al.: The Automation of Science. *Science* 324(5923), 85–89 (2009)
- [10] M. Ashburner, et al. Gene ontology: Tool for the unification of biology. *Nature Genetics* 25, 25–29, 2000.
- [11] B. Smith, et al. The OBO Foundry: Coordinated evolution of ontologies, 2007.
- [12] B. Smith, et al. Relations in biomedical ontologies. *Genome Biology* 6, r46., 2005.
- [13] D. Schober, et al. Towards naming conventions for use in controlled vocabulary and ontology engineering. In: *Proceedings of BioOntologies SIG*, ISMB 2007, pp. 29.
- [14] Data Science Ontology. [online] [cit. 2023-05-29] Dostupné na internete: <<https://www.datascienceontology.org/help>>
- [15] M. Hilario, P. Nguyen, H. Do, A. Woznica and A. Kalousis. "Ontology-Based Meta-Mining of Knowledge Discovery Workflows". In: *Meta-Learning in Computational Intelligence*. 2011.
- [16] J. Kietz, F. Serban, A. Bernstein and S. Fischer. "Towards Cooperative Planning of Data Mining Workflows". In: *Proceedings of the Third Generation Data Mining Workshop at the 2009 European Conference on Machine Learning*. 2009.
- [17] J. Vanschoren and L. Soldatova. "Exposé: An ontology for data mining experiments". In: *International workshop on third generation data mining: Towards service-oriented knowledge discovery*. 2010. pp. 31–46.
- [18] P. Panov, L. N. Soldatova, and S. Džeroski. "Generic ontology of datatypes", *Information Sciences*. 2016. pp. 900–920. doi: <https://doi.org/10.1016/j.ins.2015.08.006>.

- [19] M. Tianxing, N. Zhukova, A. Vodyaho, and T. Aung. "A Meta-Mining Ontology Framework for Data Processing". *International Journal of Embedded and Real-Time Communication Systems*. 2021. pp. 37–56. doi: 10.4018/IJERTCS.2021040103.
- [20] P. Panov, S. Džeroski, and L. Soldatova, *OntoDM: An Ontology of Data Mining*, IEEE International Conference on Data Mining Workshops, 2008, s. 752–760.
- [21] J. Polowinski, m. Voight. „VISO: A Shared, Formal Knowledge Base as a Foundation for Semi-automatic InfoVis Systems, In CHI '13 Extended Abstracts on Human Factors in Computing Systems, 1791-1796. CHI EA '13. New York, NY, USA: ACM, 2013 [online] [cit. 2023-04-29] Dostupné na internete: <http://www-st.inf.tu-dresden.de/semvis/papers/authors_version_polowinski_viso_chi2013_wip.pdf>
- [22] H2O Driverless AI. Democratizing AI with Automated Machine Learning. [online] [cit. 2023-05-29] Dostupné na internete: <<https://h2o.ai/platform/ai-cloud/make/h2o-driverless-ai/>>
- [23] Amazon web service. [online] [cit. 2023-05-29] Dostupné na internete: <https://aws.amazon.com/?nc2=h_lg>
- [24] AutoML. [online] [cit. 2023-05-29] Dostupné na internete: <<https://cloud.google.com/automl>>
- [25] P. Bednár, J. Ivančáková and M. Sarnovský. "Semantic automatization of the data-analytical processes". In: *International Symposium on Applied Computational Intelligence and Informatics*. 2022.
- [26] M. Sarnovsky, P. Bednar, and M. Smatana. "Cross-Sectorial Semantic Model for Support of Data Analytics in Process Industries". *Processes* 7, no. 5: 281. 2019. doi: <https://doi.org/10.3390/pr7050281>.
- [27] A. Miles and S. Bechhofer. *SKOS Simple Knowledge Organization System Reference*. W3C Recommendation. World Wide Web Consortium. 2009. doi: <https://www.w3.org/TR/skos-reference/>.
- [28] M. Sarnovsky and J. Marcinko. "Adaptive Bagging Methods for Classification of Data Streams with Concept Drift". *Acta Polytechnica Hungarica*. 2021. pp. 47–63. doi: 10.12700/APH.18.3.2021.3.3.
- [29] J. Bruijn, H. Lausen, A. Polleres, and D. Fensel. "The web service modelling language WSML: An overview". 2006. pp. 604. doi: 10.1007/11762256_43.
- [30] S. Pittet. "What is code coverage?" [online] [cit. 2023-01-29] doi: <<https://www.atlassian.com/continuous-delivery/software-testing/code-coverage>>.
- [31] Flash Mob Computing. What Is The Purpose Of Code Coverage .[online] [cit. 2023-05-29] Dostupné na internete: <<https://www.flashmobcomputing.org/what-is-the-purpose-of-code-coverage/>>
- [32] Breast Cancer Wisconsin (Diagnostic) Data Set. [online] [cit. 2023-02-25]. doi: <<https://www.kaggle.com/uciml/breast-cancer-wisconsin-data>>.
- [33] J. Ivančáková, F. Babič and P. Butka. "Comparison of different machine learning methods on Wisconsin dataset". 2018 IEEE 16th World Symposium on Applied Machine Intelligence and Informatics. Kosice and Herlany, Slovakia. 2018. pp. 173-178. doi: 10.1109/SAMI.2018.8324834.

Impact of magnetic and electric field on structure of magnetic fluids

Maksym Karpets

Department of Physics
Faculty of Electrical Engineering and Informatics
Technical University of Kosice
Kosice, Slovakia
maksym.karpets@tuke.sk

Milan Timko, Peter Kopčanský, Michal Rajňák

Department of Magnetism
Institute of Experimental Physics SAS
Kosice, Slovakia
timko@saske.sk, kopcan@saske.sk, rajnak@saske.sk

Abstract—Magnetic fluids are suspensions of magnetic nanoparticles in a liquid carrier. Their structural behavior in magnetic fields is quite intuitive and scientifically well explored. In the field of fundamental research on magnetic fluids, there are well-known electric properties and phenomena, the understanding of which and subsequent practical applications in electrical engineering require deeper experimental study. The submitted work reviews the state of art of influence of magnetic and electric field on structure of magnetic fluids to highlight the significance of research on transformer oil based magnetic fluids due to their potential application in high voltage engineering. Within the dissertation thesis, the progressive magnetic fluids based on insulating liquids were investigated. The objective of the dissertation was the experimental investigation of bulk structural changes in magnetic fluids as well as the magnetic fluid-solid interfaces under the influence of various electric and magnetic fields. To attain the objective, the small angle neutron scattering and neutron reflectometry were employed as primary experimental methods. The observed effects could be used to increase the dielectric breakdown voltage of working fluids and to enhance the heat transfer from the current carrying conductor to the surrounding cooling liquid.

Abstrakt— Magnetické kvapaliny sú suspenzie magnetických nanočastíc v nosnej kvapaline. Ich správanie v magnetickom poli z hľadiska štruktúry je celkom intuitívne a na vedeckej úrovni veľmi dobre preskúmané. V oblasti základného výskumu magnetických kvapalín sú známe aj ich elektrické vlastnosti a javy, ktorých pochopenie a následná aplikácia v praxi (v elektrotechnike) si vyžadujú prehlbujúce experimentálne štúdium. V predkladanej práci je stručne zhrnutý literárny prehľad súčasného stavu vplyvu magnetického a elektrického poľa na štruktúru magnetických kvapalín, poukazujúci na dôležitý význam výskumu magnetických kvapalín na báze transformátorových olejov vzhľadom na ich potenciálne využitie v elektroenergetike. V rámci tejto dizertačnej práce boli skúmané progresívne magnetické kvapaliny na báze elektroizolačných kvapalín. Cieľom dizertačnej práce je experimentálne skúmať štruktúrne zmeny magnetickej kvapaliny v celom objeme, ako aj na rozhraní s tuhým telesom pod vplyvom rôznych magnetických a elektrických polí. Pre splnenie tohto cieľa boli použité najmä metódy malouhlového rozptylu neutrónov a neutrónovej reflektometrie. Pozorované efekty by sa mohli použiť na zvýšenie dielektrického prierného napätia pracovných kvapalín a na

zvýšenie prenosu tepla z vodiča nesúceho prúd do okolitej chladiacej kvapaliny.

Keywords— *electric field; magnetic fluid; transformer oil; magnetic nanoparticles; neutron reflectometry; SANS.*

I. INTRODUCTION

The study of liquid systems occupies a significant place in formation of the worldview of the liquid state of matter. One example of liquid systems are magnetic fluids. Magnetic fluids (MFs), or so-called ferrofluids (FF), are colloidal systems that form from a liquid base, which includes magnetic polydisperse particles with characteristic size of several nanometers, and the classification of liquid systems are colloidal suspensions. Every magnetic particle in the system covered with a layer of surfactants, which do not give particles contact each other and combine into aggregates. Depending on desired characteristics of the ferrofluid, the carrier fluid in magnetic fluids can serve as both polar and non-polar fluid. The result is a stable two-phase liquid system, which is similar in characteristics and properties to a homogeneous liquid.

Magnetic fluids are widely used in various fields they have great potential for use in engineering and medicine [1]. One of the key equipment in the electrification system is an electric transformer, which transforms voltage at least between two voltage levels and transfers electric energy. During the transformer operation the heat energy is generated in the transformer windings and core [2]. It is necessary to transfer the heat from the transformer. Recently, nanodielectric fluids have been proposed to be used as transformer liquid insulation. The addition of particles seeks to improve the dielectric and thermal properties of the base liquids, giving rise to a new generation of insulating fluids. Despite the large number of works devoted to the study of the structure of magnetic fluids in the volume of the sample, still there are properties and phenomena, the understanding of which and subsequent practical applications in electrical engineering require deeper experimental study.

The interaction of MNPs with external magnetic fields can lead to different structural re-organizations of ferrofluids including the formation of aggregates, chains and more

complex patterns, which are the subjects of intensive studies for many years by various experimental methods and theoretical approaches. Effect of an external magnetic field and the induced heterogeneous nanoparticle (NP) structure results in the magneto-viscous effect and magneto-dielectric anisotropy [3]. Also, it was found that the thermal conductivities parallel to the magnetic field direction are, in most cases, much higher than the perpendicular ones [4]. Thus, the magnetic field intensity, direction and distribution plays a key role in the study of MF. Analogously, structural transitions in various colloidal suspensions can be induced by electric fields leading to particle-particle electrostatic interactions, playing an important role in numerous self-assembly phenomena [5].

The study of magnetic fluid systems is of both fundamental and applied interest. Studies of the behavior of their self-organization at the interface can significantly expand the understanding of the phenomenon of adsorption in liquid systems. The practical application of magnetic fluid systems leads to the need for the synthesis of ferrofluids with pre-known stability parameters. Therefore, obtaining the relationship of structural parameters in the volume and at the interface under external fields will predict the behavior of magnetic fluid systems during their using in real working conditions.

II. THEORETICAL BACKGROUND AND RELATED LITERATURE

The term magnetic fluid evokes the idea of a fluid with ferromagnetic behavior. Although the ferromagnetism of a homogeneous liquid is possible on a theoretical level, a substance with a Curie point exceeding the melting point of the substance has not yet been discovered. For this reason, it is only possible to prepare a magnetic fluid synthetically as a stable colloidal suspension of monodomain or multidomain magnetic particles in a liquid carrier [6].

In early publications, MFs were produced by grinding magnetite with heptane or long chain hydrocarbon and a grinding agent, e.g. oleic acid. Then MFs were produced by precipitation of an aqueous $\text{Fe}^{3+}/\text{Fe}^{2+}$ solution with a base, coating these particles with an adsorbed layer of oleic acid and then dispersing them in a non-aqueous fluid. Later many other perspective methods of manufacturing of magnetic fluids were discovered. Although most of the works in the field of nano-dielectric fluids are focused on mineral-oil based liquids, some investigations have also been reported on the use of natural ester or synthetic ester based nanofluids [7].

Finite size and surface effects can modify particles properties when compared with bulk materials. With decreasing the particle size coercivity increases until critical size d_{SD} is reached. Below the value d_{SD} , a single domain state is preferred and over d_{SD} a multidomain state is preferred. Below d_{SD} coercivity begins to decrease with lowering the particle dimensions. Magnetic moment ordering inside single domain particle depends mostly on magnetic anisotropy and the particle shape. For particles with high crystallography

anisotropy magnetic moments are aligned in the easy axis direction. The nanoparticle does not contain domain walls.

In general, magnetization M of a magnetic particle depends on magnetic field strength H . In nanoparticle body there appears demagnetizing field. Therefore, induced field inside of magnetic nanoparticle (MNP) H_{ind} equals to applied magnetic field H reduced by a demagnetizing field.

For small nanoparticles (in order of nm) a similar behavior to paramagnetism can be observed even at temperatures lower than Curie temperature. In such state particles behave like huge spins. Magnetic moments of particles are several times larger than atomic moments. They reach values 102 to 105 times greater. The state was called superparamagnetism due to similarity with paramagnetism and the existence of giant superspin [8].

In most magnetic fluids applications, the magnetic material is represented by one of various ferrites. The most commonly used ferrites are magnetite (Fe_3O_4) and maghemite ($\gamma\text{-Fe}_2\text{O}_3$). One of the most common methods for preparing nanocrystalline magnetic iron oxide particles for magnetic fluid is the wet-chemical co-precipitation method. Magnetic fluid systems are characterized by relatively large specific surface areas, and therefore significant surface energy. Since MFs are colloidal suspensions, they are not, as they can undergo processes with a decrease in surface energy through the formation aggregates, adsorption of various molecules on the surface of particles and, consequently, their gradual precipitation. Such processes violate the stability of the system. There are two types of stability of colloidal systems: stability against sedimentation (associated with subsidence under the action equilibrium of gravity) and stability against aggregation [9].

There are many parameters of magnetic fluids sensitive to electric and magnetic fields such as density, viscosity, dielectric permeability, magnetic permeability, conductivity, inductivity, heat conductivity, heat capacity, surface tension, ultrasound absorption, optical absorption, refraction index, reflection index. With regards to the potential application of transformer oil (TO)-based magnetic nanofluids (MNFs) in power transformers, most research studies have focused on three key properties: thermal, dielectric, and viscous. On the other hand, this property is a function of the base liquid, dispersed particles, size distribution, volume fraction, surfactant, and the external magnetic and electric field.

The magnetic nanoparticle interactions and subsequent structural reorganization, aggregates, chains and pattern formation have been intensively studied by various experimental methods and theoretical approaches [11-15]. The magnetic field induced microscopic structures have a substantial impact on macroscopic behavior of ferrofluids. Especially, the effect of the increased viscosity due to the applied magnetic field influences the specific adaptability of ferrofluids applications. Within this context, magnetoviscous effect has been introduced by Odenbach [10], which expresses the measure of the increased dynamic viscosity in relation to the basic viscosity in the absence of a magnetic field.

Static liquid mountains are emerging on the free surface of a magnetic fluid when a critical value of the vertical magnetic

induction is surpassed. When a magnetic field is applied to ferrofluid, the viscosity of this ferrofluid increases in a phenomenon known as the magnetoviscous effect [11].

Two distinct mechanisms exist for the rotation of the magnetic moments in magnetic fluids. One is the rotation of the magnetic particle inside the liquid carrier, known by the names of Debye rotation or Brownian rotation. The other mechanism is the rotation of the magnetic moment with respect to the particle, known as Néel rotation. The relaxation time for this rotation is strongly dependent on the particle's volume and on the temperature.

Study [12] showed that the dielectric response of MF is dependent on the strength and frequency of the electric field. It was also proved [13] that the low frequency relaxation process is slowed down when the inter-electrode gap is increased in units of micrometers and when dc bias field is applied to the sample. The electric field induced changes in the ferrofluid structure at nanoscale were recently confirmed by *situ* small angle neutron scattering (SANS) experiments [5]. The insulating and dielectric properties of TO were found to exhibit peculiar enhancement upon the addition of MNPs [14]. The obtained theoretical results of P.A. Selyshchev et al. [15] show that the homogeneous distribution of dielectric particles in the dielectric carrier could become inhomogeneous under the applied electric field.

Neutron reflectometry (NR) is one of the most promising methods of studying the structure and physical properties of thin layers of matter. The dependence of the FF ordering at SiO₂ interfaces on magnetic field was first studied systematically by Vorobiev et al. [16] using *situ* NR. Then Theis-Bröhl et al. [17] performing NR studies observed a close-packed wetting layer of magnetite NPs on a silicon dioxide surface of a Si substrate. The effect of MNP assembly formation at a planar interface between a transformer oil-based ferrofluid and single-crystal silicon under non-homogeneous magnetic fields was also studied by Nagorny et al. [18].

Structural changes in transformer oil-based magnetic fluids under magnetic fields were investigated using surface acoustic waves [19].

III. SELECTED METHODS OF INVESTIGATIONS OF STRUCTURE OF MAGNETIC FLUIDS

Nowadays, the SANS method is effectively used in structural research of bulk ferrofluid specimens. The method makes it possible to analyse the ferrofluid structure on the microscopic scale and, accordingly, relates it to macroscopic properties. It should be noted that, since recently, the method of neutron reflectometry has also been actively used for studying the adsorption of ferrofluid nanoparticles on the solid surface. In specular reflectometry experiments the reflectivity R (ratio between reflected (I_{refl}) and incident (I_0) beam intensities) is measured as a function of q , the component of the momentum transfer vector along the normal to the interface plane.

The reflection of neutrons from the surface, however, is very different from most methods of neutron scattering. In the case of reflection in the region and below the critical angle (where neutrons are completely reflected from a smooth

surface), the processes of multiple scattering cannot be neglected. Since a neutron exhibits wave properties, we can use the theory of classical optics to show the behavior of neutrons under these conditions. Fresnel formulas are also valid for a multilayer system. The falling beam will be repeatedly reflected and refracted at each boundary of the 2 media. The calculation of the coefficients of reflection and transmission in each layer and the calculation of the reflectivity in air are therefore possible [20].

The small-angle scattering techniques have been used for structural characterization of materials at nanometer length scales in the range of 1–100 nm using X-rays and neutrons [21]. SANS is undertaken with thermal neutrons by interaction with nuclei and small angle X-ray scattering is sensitive to electron density only.

The general theory of the scattering of thermal neutrons shows that the differential scattering cross section $d\Sigma/d\Omega$, can be considered with good approximation, as a function of momentum transfer q as a sum over all nuclei.

In the SANS experiment we measure $d\Sigma/d\Omega$, the differential scattering as [22]:

$$q = 4\pi/\lambda \sin(\theta/2) \quad (1)$$

where λ is the neutron's wavelength and θ is the scattering angle.

From the relation (1) it can be seen that to obtain the dependence $d\Sigma(q)/d\Omega$ there can be measured or the scattering angle θ , or the wavelength λ , or both parameters.

The main parameter of material that defines interaction with neutrons is scattering length density (SLD):

$$\rho = \rho_m N_A / M \sum n_i b_i \quad (2)$$

where M is atomic mass, ρ_m - material density, N_A - Avogadro's number, n_i - quantity of nucleus i .

IV. EXPERIMENTS AND RESULTS

The experimental part is divided into three sections. The first section focuses on preparation and basic characterization of MF. The second section follows the SANS experiments with MF under electric and magnetic field. The third section deals with the electric field effect on MF founded by neutron reflectometry.

A. Preparation and Characterization of Investigated Magnetic Fluids

The investigated MFs are based on a commercially available inhibited insulating transformer oils: MOL TO 40A, SHELL Diala S4 ZX-I and Mogul Trafo CZ-A (Paramo). The iron oxide and Mn–Zn ferrite nanoparticles were synthesized by chemical co-precipitation method from aqueous solution of ferrous and ferric ions in the presence of NH₄OH at 80–82 °C. After the co-precipitation, the iron oxide NPs were sterically stabilized by chemisorption of a single oleic acid layer (C₁₈H₃₄O₂, Merck) on the particle surfaces. The coated MNP were purified in Acetone and magnetic decantation was applied to remove the excess of the oleic acid. So, the excess of unbound oleic acid was removed in the sample according to the standard procedure [23]. Synthesis of ferrofluid was performed according to the previous reports [24].

The magnetic properties of the powder MNPs and TOFF samples were measured by means of a vibrating sample magnetometer. The magnetization curves were measured at 298 K in the field ranging up to 6 T. The ferrofluid under study shows typical superparamagnetic behavior with zero coercivity and remanence in the magnetization curve. The fitting of the magnetization curve with superposition of Langevin functions gives the mean magnetic diameter (magnetic core) of log-normal distributed MNPs. Magnetization curves for MOL TO 40A based TOFF and distribution of the powder magnetite MNPs are presented in Figure 1 [25].

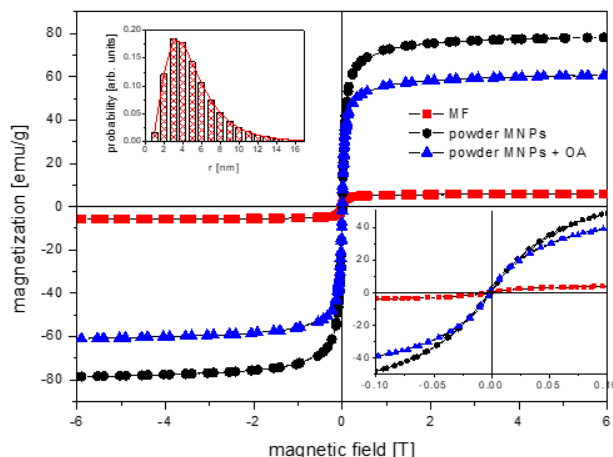


Fig. 1 Magnetization curves of the powder magnetite MNPs, MNPs with surfactant (oleic acid) and the magnetic fluid. Inset 1 (left upper corner): distribution of MNPs magnetic radius in TOFF calculated from the magnetization. Inset 2 (right bottom corner): superparamagnetic behavior (zero hysteresis) of the samples [25].

To confirm the considered lossy dielectric response of the MF samples, we present the dielectric loss of both the MF samples and the transformer oil in the form of temperature-dependent dissipation factor [26]. Eltel ADTR-2K plus was used. Figure 2 shows that the dissipation factor of transformer oil Mogul Trafo CZ-A exhibits low and quasi-constant values at various temperatures. However, with the addition of the oleic acid coated nanoparticles, the dissipation factor exponentially increases with temperature.

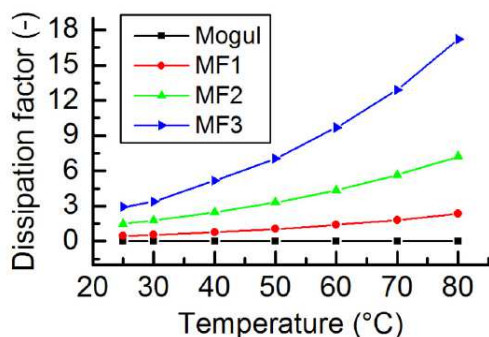


Fig. 2 Temperature-dependent dissipation factor of the transformer oil Mogul and magnetic nanofluid samples [26]

AC magnetic susceptibility of samples was measured by an AC susceptometer (IMEGO, DynoMag, SE). Not only magnetite but Mn–Zn ferrite nanoparticles based magnetic fluids were characterized. In the following study, we applied transformer oil prepared by the gas-to-liquid technology (Shell, Diala S4 ZX-I). Figure 3 shows the real and the imaginary magnetic susceptibility in dependence on the magnetic field frequency at room temperature [27]. It is seen that the susceptibilities of ferrofluid are quasi-constant in the whole frequency range, meaning that the nanoparticles are well relaxed and easily follow the field changes.

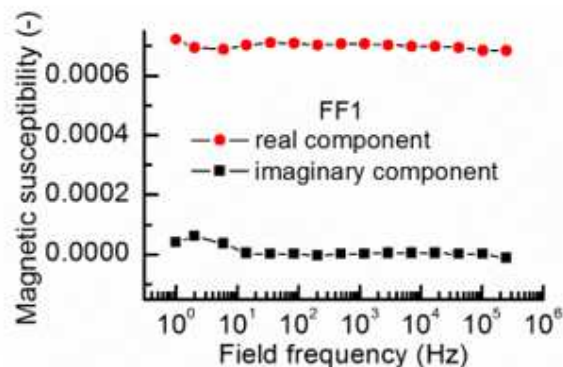


Fig. 3 AC magnetic susceptibility of the ferrofluids vs. magnetic field frequency. The amplitude of the excitation field is 0.5 mT [27].

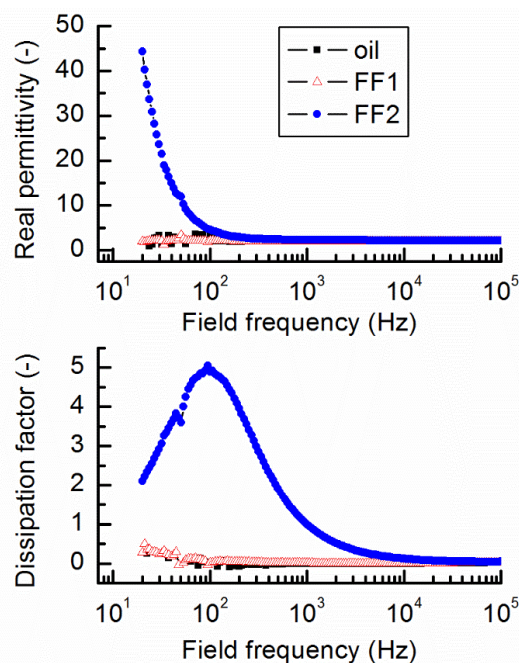


Fig. 4 The spectra of real permittivity and dissipation factor of the oil and the ferrofluid samples measured at room temperature [27].

Analogously to the AC magnetic susceptibility spectra, the real part of dielectric permittivity and dissipation factor of the oil and the ferrofluid samples in dependence on the electric field frequency are presented in Figure 4. The measurements were performed by means of an LCR meter (Agilent E4980A). As the figure shows, the oil and the diluted FF1 sample exhibit low values and quasi constant behavior of the two parameters. On the other hand, the permittivity of more concentrated FF2 sample increases remarkably with decreasing frequency (below 100 Hz). This increase, accompanied with the maximum in the dissipation factor, is attributed to the nanoparticle interfacial dielectric relaxation.

B. SANS Study of Transformer Oil-Based Ferrofluids under Electric Field

SANS experiments with TOFF (based on SHELL and MOL oil) were carried out on the small-angle neutron diffractometer - time-of-flight YuMO instrument [22] at the pulsed IBR-2 reactor at JINR. For the SANS experiments, we used a standard quartz cuvette (Hellma, 1 mm thick) equipped with two stainless tubular electrodes inside, fixed 1 cm apart in the Teflon stoppers (Figure 5) [28]. Our simulation confirmed a homogenous field in the middle of the cuvette, while the near electrode regions exhibit gradient field with the strongest intensity in the stoppers.

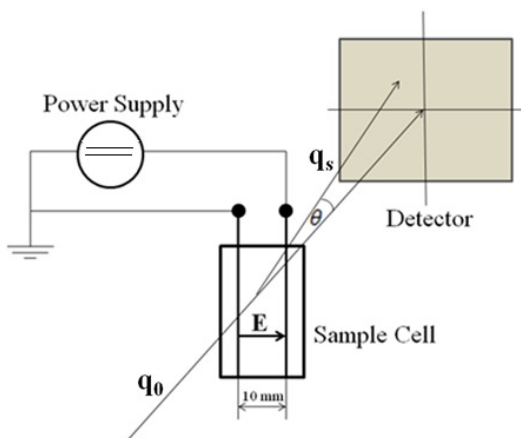


Fig. 5 Schematic representation of the cuvette for sample with the electrical circuit and neutron beam [28]

The TOFF structure at nanoscale under the dc electric field was followed by SANS representing the scattered neutron intensity I as a function of the momentum transfer modulus q . Experimental measurements were performed for a cuvette fully filled with ferrofluid. During the experiment, the sample was fixed and only the intensity of the dc applied electric field was varied. Each curve was measured from 2 to 6 times. All received experimental SANS data could be well treated with theoretical model of core-shell particles. Core-shell type nanoparticles are a type of biphasic materials which have an inner core structure and an outer shell made of different components (surfactant in the case of MNP). The mean radius of such particles was about 4 nm. The obtained experimental SANS dependencies and fit for the ferrofluids based on

insulating transformer oils SHELL and MOL with different nanoparticles concentration for voltage 0 kV are plotted in Figure 6. One can see significant increasing in scattering intensity with increasing of iron oxide nanoparticles concentration. Similar behavior is confirmed also in presence of electric field.

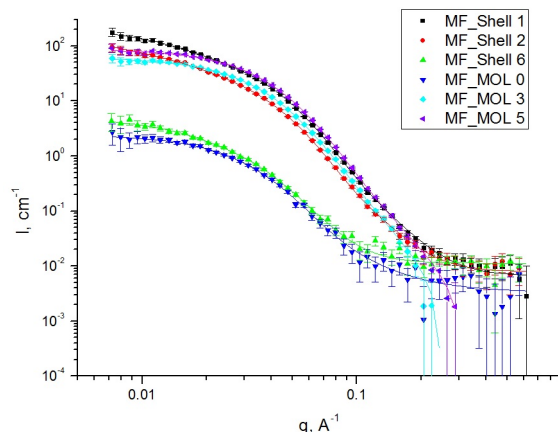


Fig. 6 SANS data (points) and fit (curves) for the ferrofluids based on insulating transformer oils SHELL and MOL with electrodes without the dc electric field (0 kV/cm) [28]

Actually, there are almost no changes in SANS data with applying and increasing dc electric field strengths for all samples. One can see the curves for MF_MOL 3 in Figure 40. There is a little increasing in intensity for small q for higher concentration, but it is in the error range. Therefore, expected aggregates are probably of greater size which is not possible to see in the q -range covered in the experiment.

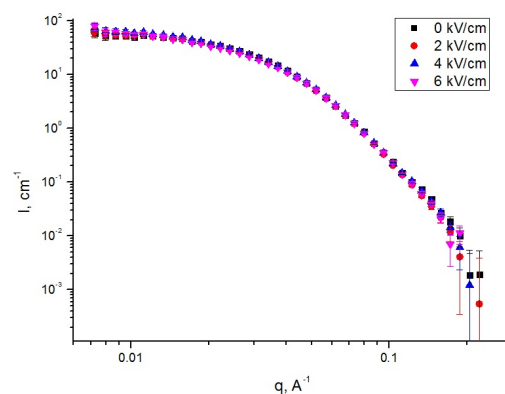


Fig. 7 SANS data for the MF_MOL 3 sample exposed to the dc electric field strengths from 0 kV/cm (initial) to 6 kV/cm [28]

Another in situ SANS experiments with TOFF based on Mogul Trafo CZ-A oil were carried out on the small-angle neutron diffractometer D33 at the ILL (Grenoble, France). SANS curves of a ferrofluid was measured in 7 dc electric fields. This protocol was repeated in the presence of an external magnetic field (1 and 425 mT, parallel and perpendicular to the electric field). There were no remarkable difference between

the scattering intensity of up and down neutron spins, so pointing out the absence of any magnetic structure in the created aggregates. The identical SANSPOL curves are observed for both cases, in the absence and presence of the electric field. A noticeable difference in the SANSPOL curves has been detected in the case, when the sample is simultaneously exposed to both, electric and magnetic field. In magnetic field only, the SANSPOL curves indicate the well-known chain like anisotropy in the magnetic field induced aggregates (**Chyba! Nenašel sa žiaden zdroj odkazov.**). The most pronounced difference in the SANSPOL curves is seen in the case when both, magnetic (425 mT) and electric (7 kV/cm) fields act on the ferrofluid sample. As the magnetic field was applied firstly on the sample and subsequently the electric voltage was switched on, one can assume that the electric field enhanced the magnetic anisotropy in the formed aggregates.

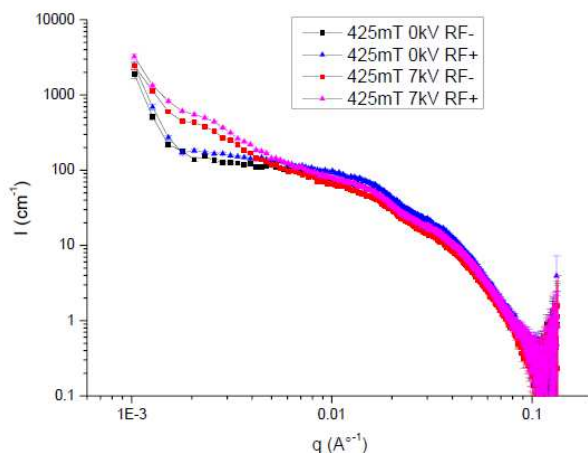


Fig. 8 Simultaneous effect of electric and magnetic field on the SANSPOL curves

C. Neutron Reflectometry Study of Magnetic Fluids under Electric Field

Specular reflectivity of non-polarized neutrons was measured at the neutron reflectometer GRAINS [29] with a horizontal sample plane configuration installed at the IBR-2 pulsed reactor of JINR. The measurements were carried out in the ‘beam from the bottom’ configuration in the time-of-flight mode that is standard positioning to study the solid-liquid interfaces when a plane neutron beam passes through a single-crystal silicon substrate to meet and scatter from the interface between a liquid sample and an electrode film. The sample in the experimental cell was exposed to direct current (DC) electric field of various intensities (from 100, 300, 500 and 700 kV/m), which are close to the operating values commonly used in high voltage power engineering [25]. The configuration of the sample cell and crystal is schematically depicted in Figure 9.

The obtained NR curves were processed in a standard way in terms of the Parratt formalism using the Motofit package for the IGOR Pro software. In this model the interface is split into n layers sandwiched between the super- and sub-phases. Since the incident neutron beam is refracted by each of the

layers and the interface between layers is unlikely to be perfectly smooth the roughness/diffuseness of each interface modifies the Fresnel coefficient. The algorithm used in the program - genetic optimization - is very efficient at finding global minima in the χ^2 function, even if the starting guesses are far from the final solution. The characteristics of the formed layers at the interface were obtained in a homogeneous approximation including 3 parameters for each layer: the mean nuclear SLD, thickness and interlayer roughness. The obtained experimental and best model NR curves for different electric field intensity applied are presented in Figure 10.

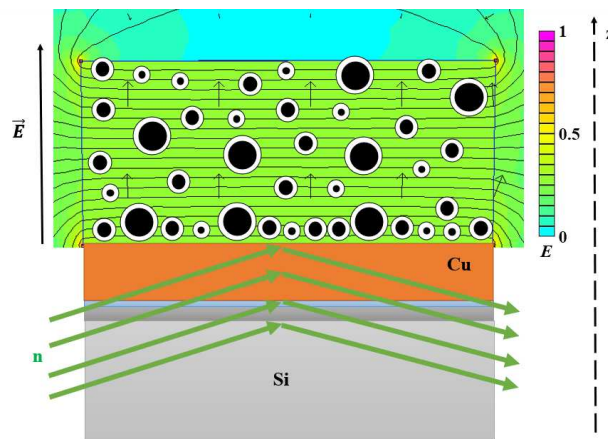


Fig. 9 Schematic of the neutron reflectivity measurement configuration that depicts the ferrofluid in the sample cell; color scale of the simulated electric flux density and related electric equipotential field lines distributed between the planar electrodes used during the experiments [25].

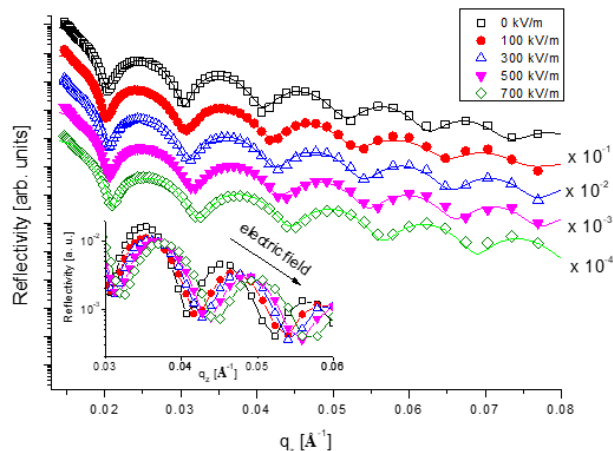


Fig. 10 Specular reflectivity curves obtained in the NR experiment when applying electric field from 0 to 700 kV/m. Solid lines represent best fits [25].

The oscillations in the reflectivity curves (the so-called ‘Kiessig fringes’) are a result of the interference effects regarding the electrode copper layer formally located between two (substrate and ferrofluid) semi-infinite media. The deviations from this model are from the intermediate titanium layer between copper electrode and silicon substrate, as well

as possible oxide layers on silicon and copper and diffusive layers between the solid interface components. The reflectivity curves for the interface under electric field were fitted with all fixed values of the solid components parameters and all varying parameters of FF components. The received fit parameters from the previous dataset (with smaller electric field) are used as a starting point for the next dataset (with higher electric field). The electric field induced evolution of the ‘liquid part’ of the interface in terms of the changes in the SLD profile is followed in Fig. 2. As expected, the increase in the electric field slightly affects the structure of layer #1 determined by the surfactant shell; one can see small reduction in thickness, roughness and SLD of this layer. More significant changes are observed for layer #2. It becomes thicker and more concentrated thus evidencing the enhanced adsorption of MNPs. The layer thickness grows with the electric field increase together with the SLD. Starting from 300 kV/m, the fits of the reflectivity curves are better if an additional layer #3 is introduced. It describes a transition to the bulk ferrofluid. At the same time, some redistribution of MNP in the two layers takes place at maximum field intensity, reflecting a decrease of the MNPs content in the first layer. The reason of such development is the polarization of the particles and their interaction as dipoles. Moreover, from the relaxation measurements we can conclude that a residual electric field and potentially present magnetic and Van der Waals forces influence nanoparticles assembling even some time after switching the field off.

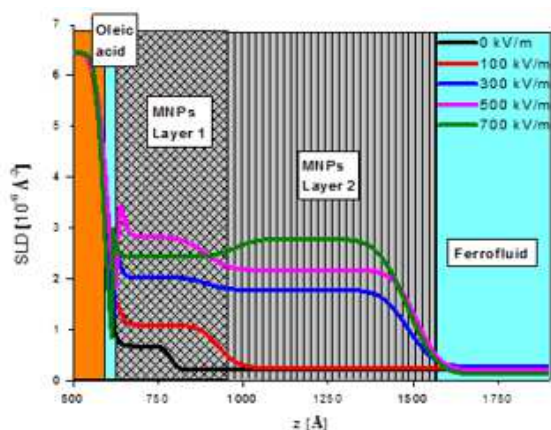


Fig. 11. SLD profiles at the electrode-ferrofluid interface at different intensities of the external electric field [25].

V. CONCLUSIONS

We report assembling of superparamagnetic nanoparticles of a dilute classical ferrofluid - MNPs coated with oleic acid in transformer oil. Forming of magnetic nanoparticles layers was found on a planar surface of copper electrode film on a single-crystal silicon substrate under perpendicular homogeneous electric field by neutron reflectometry. It has been demonstrated nanoparticles concentration influence on SANS measurements. From the SANS measurement of magnetic

fields in combined electric and magnetic field one can assume that the electric field enhanced the magnetic anisotropy in the formed aggregates. The obtained results contribute to a better understanding of ferrofluids. The observed self-assembled layering could be used as an additional barrier at the inner surface of transformer to increase dielectric breakdown voltage of working fluids.

ACKNOWLEDGMENT

The work is a part of the projects supported by the Programme of SAS grants for PhD students (No. APP0140); the Science Grant Agency of the Slovak Republic (No. 2/0011/20 and 2/0016/17); the Slovak Research and Development Agency (No. APVV-18-0160).

REFERENCES

- [1] M. V. Avdeev *et al.*, *Langmuir* **26**, 8503 (2010).
- [2] M. Rafiq, Y. Lv, and C. Li, *Journal of Nanomaterials* **2016**, 1 (2016).
- [3] M. Rajnak *et al.*, *Energies* **12**, 4532 (2019).
- [4] A. Katiyar *et al.*, *Journal of Magnetism and Magnetic Materials* **419**, 588 (2016).
- [5] M. Rajnak *et al.*, *Appl. Phys. Lett.* **107**, 073108 (2015).
- [6] R. E. Rosensweig, *Ferrohydrodynamics*, Dover edition (Dover Publications, Inc, Mineola, New York, 2014).
- [7] Y. Lv *et al.*, *IEEE Electr. Insul. Mag.* **30**, 23 (2014).
- [8] B. Pacakova *et al.*, in *Magnetic Spinels - Synthesis, Properties and Applications*, edited by M. S. Seehra (InTech, 2017).
- [9] L. Vékás, M. V. Avdeev, and D. Bica, in *NanoScience in Biomedicine*, edited by D. Shi (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009), pp. 650–728.
- [10] S. Odenbach, L. M. Pop, and A. Yu. Zubarev, *GAMM* **30**, 195 (2007).
- [11] S. Odenbach, editor, *Colloidal Magnetic Fluids: Basics, Development and Application of Ferrofluids* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009).
- [12] M. Rajnak *et al.*, *Journal of Applied Physics* **114**, 034313 (2013).
- [13] M. Rajnak *et al.*, *Phys. Rev. E* **90**, 032310 (2014).
- [14] J. Kurimský *et al.*, *Journal of Magnetism and Magnetic Materials* **496**, 165923 (2020).
- [15] P. A. Selyshchev *et al.*, *Journal of Molecular Liquids* **278**, 491 (2019).
- [16] A. Vorobiev *et al.*, *Phys. Rev. Lett.* **93**, 267203 (2004).
- [17] K. Theis-Bröhl *et al.*, *Soft Matter* **11**, 4695 (2015).
- [18] A. Nagomyi *et al.*, *Applied Surface Science* **473**, 912 (2019).
- [19] P. Bury *et al.*, *Acta Phys. Pol. A* **137**, 964 (2020).
- [20] L. G. Parratt, *Phys. Rev.* **95**, 359 (1954).
- [21] A. M. Mebert *et al.*, in *Advanced Spectroscopic Methods to Study Biomolecular Structure and Dynamics* (Elsevier, 2023), pp. 271–307.
- [22] A. I. Kuklin, A. Kh. Islamov, and V. I. Gordeliy, *Neutron News* **16**, 16 (2005).
- [23] D. Bica, L. Vékás, and M. Raša, *Journal of Magnetism and Magnetic Materials* **252**, 10 (2002).
- [24] L. Vékás, D. Bica, and M. V. Avdeev, *China Particuology* **5**, 43 (2007).
- [25] M. Karpets *et al.*, *Journal of Molecular Liquids* **362**, 119773 (2022).
- [26] M. Rajňák *et al.*, *Journal of Molecular Liquids* **309**, 113243 (2020).
- [27] M. Rajňák *et al.*, *J. Phys. D: Appl. Phys.* **55**, 345002 (2022).
- [28] M. Karpets *et al.*, *Ukr. J. Phys.* **65**, 729 (2020).
- [29] M. V. Avdeev *et al.*, *Crystallogr. Rep.* **62**, 1002 (2017).

Vysvetliteľnosť umelej inteligencie pri spracovaní videosnímkov z USG záznamov

Michal Kolárik
KKUI

Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
e-mail: michal.kolarik@tuke.sk

Ján Paralič
KKUI

Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
e-mail: jan.paralic@tuke.sk

Martin Sarnovský
KKUI

Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
e-mail: martin.sarnovsky@tuke.sk

Abstrakt—Tento príspevok má za cieľ prezentovať výsledky dizertačnej práce, ktorá sa zameriava na aplikovanie vysvetliteľných metód na modely hlbokého učenia v medicínskej oblasti. Prinášame aktuálny prehľad stavu vysvetliteľnosti v oblasti medicíny. V prehľade súčasného stavu zachytávame základné definície, požiadavky a potreby umelej inteligencie v medicínskej oblasti ako aj základné rozdelenie metód vysvetliteľnosti. V ďalšej časti poskytneme a predstavíme riešenie pre detekciu a klasifikáciu videosnímkov využitím metód hlbokého učenia pre spracovanie videosnímkov. Na výsledný model boli aplikované metódy vysvetliteľnosti a výstupy boli konzultované s lekármi.

Index Terms—XAI, USG, hlboké učenie, klasifikácia videa, medicínske snímky

Abstract—This paper aims to present the results of a dissertation that focuses on the application of explainable methods to deep learning models in the medical domain. We provide an up-to-date overview of the state-of-the-art of explainability in the medical domain. In the state-of-the-art review, we capture the basic definitions, requirements and needs of artificial intelligence in the medical domain as well as the basic distribution of explainability methods. In the next section, we provide and present a solution for video image detection and classification by using deep learning methods for video image processing. Explainability methods were applied to the resulting model and the outputs were consulted with medical practitioners.

Index Terms—XAI, USG, deep learning, video classification, medical image

I. ÚVOD

Metódy hlbokého učenia sú súčasným trendom v oblasti spracovania údajov. V medicínskej oblasti je však okrem samotného výkonu modelu dôležité vedieť popísať a vysvetliť jeho správanie v rôznych situáciách. Okrem toho tento trend sleduje aj legislatíva, ktorá vyžaduje, aby rozhodnutia modelov umelej inteligencie boli bezpečné, eticky spravodlivé a transparentné. Zákon GDPR nám ponúka právo na vysvetlenie modelu pri práci s našimi údajmi. V oblasti medicíny je zaobchádzanie s osobnými údajmi ešte kritickejšie. Vysvetliteľnosť umelej inteligencie (XAI) je typ umelej inteligencie, ktorá dokáže poskytovať zrozumiteľné vysvetlenia o svojich rozhodovacích procesoch. XAI je navrhnutá tak, aby boli systémy umelej inteligencie pre ľudí transparentnejšie a interpretovateľnejšie, aby sme mohli pochopiť, ako dospeli k určitému rozhodnutiu.

V tomto príspevku uvádzame naše dve nedávne publikácie zamerané na vysvetliteľné metódy umelej inteligencie

pri spracovaní medicínskych dát. V našom prípade ide o videodáta z oblasti ultrasonografie (USG). V prvom, prehľadovom článku [1] sme sa venovali identifikácii problémov, požiadaviek či potrieb umelej inteligencie v oblasti medicíny. Článok poskytuje analýzu súčasného stavu ako aj prehľad metód použitých pri spracovaní videozáznamov. V druhom článku [2] predsavujeme postup, ktorým je možné spracovať USG videozáznamy a klasifikovať na nich pohyb plúc pomocou modelov hlbokého učenia.

II. ANALÝZA SÚČASNÉHO STAVU V OBLASTI VYSVETLITEĽNEJ UMELEJ INTELIGENCIE

AI (angl. Artificial Intelligence) systémy alebo systémy umelej inteligencie na báze modelov typu čierna skrinka (angl. black box) sa dnes používajú v mnohých oblastiach. Systémy AI používané v spotrebnej elektronike akou sú smartfóny alebo v online službách nemajú kľúčové požiadavky na transparentnosť modelov, ale sústredia sa hlavne na výkonnosť modelov. Ak takýto model zlyhá a napríklad nerozpozná osobu pri prihlasovaní do systému alebo prekladač spraví gramatickú chybu, nemá to zásadné dôsledky. Požiadavky na transparentnosť a dôveru týchto modelov sú nízke. Naopak, v aplikáciách, ktoré sú kritické na bezpečnosť ľudí, hrajú dôležitú úlohu a dokonca môžu tvoriť rozhodujúci faktor v prípade nasadenia takéhoto systému. Najmä, ak dôsledky rozhodnutia AI môžu byť život ohrozujúce napr. autonómne riadenie alebo oblasť medicíny. Preto je transparentnosť dôležitejšia najmä v týchto oblastiach a okrem transparentnosti je potrebné aj vytvorenie dôvery v AI systém, ktorý má byť nasadený. Okrem toho vznikajú aj legislatívne požiadavky na vysvetliteľnosť a transparentnosť [3] pri spracovaní osobných údajov.

Vysvetliteľnosť posilňuje dôveru a overiteľnosť [3]. Možnosť overiť AI systém je dôležitá v situáciách kde AI tvorí podpornú úlohu napr. v diagnostike chorôb v medicíne a tiež v situáciách, kde AI robí rozhodnutia napr. autonómne riadenie. Vysvetliteľnosť môže poskytnúť extra informácie o systéme, čo môže pomôcť lekárovi spraviť si lepší obraz o pacientovi a dôvodoch prečo sa systém rozhodol určitú danú diagnózu. Dôležitá je aj možnosť overiť, že sa systém rozhoduje správne a na základe relevantných príčin.

Vysvetliteľnosť je dôležitá aj z hľadiska sociálneho [3], pretože ľudia sú zvyknutí veci opisovať a vysvetľovať, ako

došli k svojim záverom, napr. doktor vysvetľuje pacientovi jeho diagnózu. Vysvetliteľné modely by preto mali popisovať postup ako prišli k danému záveru takým spôsobom aby im ľudia rozumeli. To umožňuje zvyšovať dôveru ľudí v tieto systémy.

Ďalšie dôvody pre potrebu vysvetliteľnosti podľa U. Bhatta [4]:

- *ladenie modelov (debugging)* - výskumníci majú často problém s ladením modelov. Chcú zistiť prečo dáva model zlé výsledky pri určitých vstupoch alebo v ktorých oblastiach má model slabú výkonnosť. Taktiež hľadajú zefektívnenie funkcií, možnosti zrušenia funkcií a možnosti pre zvýšenie výkonnosti modelu. Schopnosť vysvetliť môže pomôcť aj pri identifikácii príčin problematických miest.
- *sledovanie modelov (monitoring)* - nastávajú aj obavy z posunu distribúcie sledovaných funkcií a tým pádom k nesprávnym predikciám skôr natrénovanými modelmi. Preto je potrebné sledovať model aj po jeho nasadení a byť včas upozornený ak dôjde k významnému posunu v porovnaní s tréningovou fázou. Tiež to pomôže zistiť dôsledky a prínosy pre úpravu modelu.
- *transparentnosť modelov* - organizácie ktoré nasadzujú modely pre vytváranie rozhodnutí priamo ovplyvňujúce koncových užívateľov, hľadajú vysvetlenia pre predikcie modelov. Cieľom týchto vysvetlení je zvýšiť transparentnosť modelu a dosiahnuť súlad medzi aktuálne platnými alebo budúcimi predpismi. Vo všeobecnosti sa odborníci domnievajú, že vysvetlenia môžu tiež pomôcť pri komunikácii s externými užívateľmi. Rovnako spoločnosti potrebujú vysvetliť správanie sa modelov svojim zákazníkom.
- *auditovateľnosť modelov* - vo finančných organizáciách musia všetky nasadené modely z dôvodu regulačných požiadaviek prejsť interným auditom. Dátoví analytici vytvorené modely nechávajú kontrolovať internými tímami pre kontrolu rizík a súladu s právnymi predpismi. Jedným z cieľov auditovania modelu je vykonať rôzne druhy testov stanovené predpismi.

A. Vysvetliteľnosť v doméne zdravotnej starostlivosti

V oblasti zdravotnej starostlivosti môže byť umelá inteligencia veľmi prospešná. Už dnes existujú praktické nasadenia umelej inteligencie, ktorá pomáha lekárom pri identifikácii problémov so zlyhaním srdca [5]. Avšak plné využitie možností ktoré AI ponúka, je obmedzené neschopnosťou algoritmov vysvetliť ľudským odborníkom svoje výsledky a rozhodnutia. Toto je veľký problém hlavne v medicínskej oblasti, kde lekári musia pochopiť prečo sa AI nejakým rozhodla a ako dospela k danému rozhodnutiu. V niektorých prípadoch môže na tomto rozhodnutí závisieť ľudský život. Transparentné algoritmy by navyše mohli primerane zvýšiť dôveru lekárskeho odborníka v budúce AI systémy [6]. Preto výskum zameraný na vytváranie vysvetliteľných AI systémov pre aplikácie v medicíne si vyžaduje udržanie vysokej miery výkonnosti pri učení pre rôzne techniky strojového učenia

a interakcie človeka s počítačom. Existuje tu určité napätie medzi presnosťou a vysvetliteľnosťou metód strojového učenia. Najvýkonnejšie modely (najmä hlboké učenie) sú často aj najmenej transparentné a metódy ktoré poskytujú jasné a zrozumiteľné vysvetlenia (napr. rozhodovacie stromy) sú menej presné [7].

V oblasti zdravotníctva je motivácia pre vysvetliteľnosť AI modelov celkom jasná. V mnohých prípadoch si konečný používateľ aj kritická povaha predpovedí vyžaduje určitú transparentnosť, či už pre zapojenie používateľov alebo pre bezpečnosť pacientov. Niekedy však samostatné vysvetlenie predikcií algoritmu nie je dostatočné, je preto dôležité zamyslieť sa nad tým ako sú výsledky interpretované koncovým používateľom, ako sú začlenené do pracovných procesov alebo ako sú využívané. Zdravotní pracovníci sú často preťažení návalom pacientov, návalom dát o týchto pacientoch a súvisiacich úloh, ktoré sa od nich vyžadujú (zadávanie údajov do systému, elektronické záznamy, poskytovanie starostlivosti). Preto ak AI systémy a ich vysvetliteľnosť nebude prezentovaná správnym spôsobom, nebude pomáhať zdravotníkom ale naopak bude to pre nich ďalšia úloha navyše. Preto by mali byť tieto systémy tvorené z hľadiska domény a pohľadu používateľa, ktorý s nimi bude pracovať [6].

Častokrát sa v oblasti zdravotnej starostlivosti spája AI s predstavou, že by umelá inteligencia mala nahradiť rozhodnutia zdravotníkov. Avšak nie je nutné tvoriť systémy takýmto spôsobom. Naopak, AI môže byť veľmi nápomocná pri dôležitých rozhodnutiach ktoré musia lekári vykonávať a to najmä v prípade, ak sú dôvody AI rozhodnutí či predikcií správne vysvetlené.

B. Požiadavky na AI systémy v oblasti medicíny

Oblasť medicíny kladie špecifické požiadavky na všetky počítačové systémy, pretože vyžaduje aby tieto systémy boli bezpečné, spoľahlivé, zabezpečené, certifikované či auditované. Navyše, systémy musia byť schopné spolupracovať a byť odolné voči chybám. Chyba systému môže spôsobiť výpadok elektriny, či podanie zlého lieku čo môže mať za následok smrť. Preto je nutné aby bola definovaná zodpovednosť za správne fungovanie všetkých systémov. Túto zodpovednosť nesú správcovia systémov alebo certifikačné authority.

V tejto oblasti sa čoraz častejšie objavujú obavy aj etické či regulačné aspekty umelej inteligencie v zdravotníctve. Tieto obavy zahŕňajú napríklad zaujatosť modelov, nedostatočnú transparentnosť, obavy o ochranu súkromia súvisiacu s citlivými dátami využívanými na tréningovanie modelov či otázky zodpovednosti. Hoci tieto obavy sú často témou diskusií, praktických odporúčaní či príkladov je veľmi málo.

Legislatívne požiadavky na AI systémy v zdravotníctve sa môžu v jednotlivých častiach sveta líšiť [8]. V USA podlieha schvaľovanie nových systémov a zariadení s umelou inteligenciou pod organizáciu FDA (Food, Drug, and Cosmetic Act). V EU narozdiel od USA nie sú zdravotnícke pomôcky schvaľované centralizovanou agentúrou. Zdravotnícke pomôcky sú rozdelené do tried podľa rizika, pričom pri triede 1 s najnižším rizikom je za zariadenie zodpovedný priamo jeho výrobca.

Zdravotníckymi pomôckami vysokorizikových tried (IIa, IIb a III) sa zaoberajú súkromné tzv. notifikované orgány - t. j. organizácie, ktoré boli akreditované na vykonávanie posudzovania zhody a vydávanie značky Conformité Européenne (CE).

Americký úrad pre kontrolu potravín a liečiv (FDA) len nedávno zverejnil [9] prvý akčný plán agentúry pre softvér ako zdravotnícku pomôcku (SaMD) založený na umelej inteligencii/strojovom učení (AI/ML). V tomto akčnom pláne je opísaný viacnásobný prístup, ktorého cieľom je pokročiť v dohľade agentúry nad zdravotníckym softvérom založeným na AI/ML.

Môžeme očakávať, že EU bude v dohľadnom čase nasledovať USA v rámci zlepšenia dohľadu nad kontrolou AI/ML systémov v zdravotníctve.

III. VYSVETLITELNÉ METÓDY UMELEJ ITELIGENCIE

Z dôvodu narastajúceho počtu metód v oblasti vysvetliteľnosti je náročné pochopiť výhody, nevýhody či konkurenčné výhody v rôznych doménach. Okrem toho existuje aj veľké množstvo variácií či kategórií vysvetliteľných metód ako sú globálne metódy (tzn. vysvetľujú správanie modelu na celej dátovej množine), lokálne metódy (vysvetľujú predpoveď alebo rozhodnutie na konkrétnom príklade), ante-hoc (model sa vytvára vo fáze tréningu AI modelu), post-hoc (model sa vytvára až na natrenovanom modeli), náhradný - surrogate (AI model je nahradený interpretovateľným modelom) alebo je použitý priamo interpretovateľný model (strom alebo rozhodovacie pravidla).

Molnar vo svojej knihe [10] všeobecne kategorizuje metódy vysvetlenia na tri druhy: metódy s vnútornou interpretáciou, modelové agnostické metódy a vysvetlenia založené na príkladoch.

Existujú aj iné taxonómie [11], ktoré sa napríklad pozerajú na XAI metódy na základe typu dát, ako sú tabuľkové, obrazové a textové dáta. Na obrázku 1 je zobrazená častokrát využívaná kategória XAI metód [12].

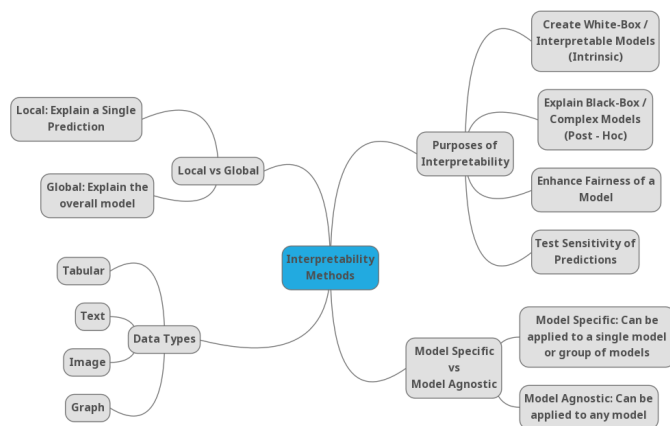


Figure 1. Taxonómia XAI metód podľa [12]

V tomto príspevku sa zameriame najmä na metódy, ktoré sme použili v rámci svojich experimentov.

A. Interpretovateľné modely

Jedným z najjednoduchších spôsobov, ako získať vysvetliteľnosť modelu umelej inteligencie, je použiť interne interpretovateľný model, ako je lineárna regresia, logistická regresia, rozhodovacie stromy alebo rozhodovacie pravidlá. Tým sa môžeme vyhnúť používaniu modelov "čiernej skrinky". Táto prirodzená vysvetliteľnosť modelov je často na úkor presnosti modelu.

V oblasti medicíny sa tento prístup v súčasnosti široko používa. Vysvetliteľné modely v medicíne [6] poskytujú transparentnosť a auditovateľnosť pri zachovaní dostatočnej výkonnosti modelu, najmä pri spracovaní tabuľkových údajov. Medzi najbežnejšie modely patria *Lineárna regresia*, *logistická regresia*, *rozhodovacie stromy* a *Rozhodovacie pravidlá*. Lineárna regresia predpovedá cieľový atribút ako vážený súčet vstupných atribútov. Linearita naučených vzťahov sa jednoducho interpretuje. Jednotlivé parametre môžeme z modelu priamo extrahovať. Logistická regresia je rozšírením lineárnej regresie a používa sa na úlohy binárnej klasifikácie. Naučený model používa logistickú funkciu na výpočet pravdepodobnosti klasifikácie. Vzhľadom na použitie logistickej funkcie je výsledný model o niečo zložitejší na vysvetlenie ako lineárna regresia. Rozhodovacie stromy možno použiť v situáciách, keď máme nelineárne údaje alebo klasifikáciu viacerých tried. Interpretácia modelu je jednoduchá. Strom začína od koreňového uzla, pričom uzly predstavujú podmienky a hrany predstavujú splnenie alebo nespĺnenie podmienky. Listový uzol zodpovedá predpovedanej výstupnej hodnote. Hĺbka a šírka stromu určujú zložitosť stromu, a teda aj jeho interpretáciu. Rozhodovacie pravidlá používajú jednoduchý princíp IF-THEN pozostávajúci z podmienky a predpovede. Podmienka môže byť jednoduchá alebo zložená, pričom sa používa spojka AND. Rozhodovacie pravidlá sú pravdepodobne najjednoduchším modelom na interpretáciu.

V poslednom období pribúdajú aj vysoko výkonné interpretovateľné modely *Soznamy padajúcich pravidiel*[13] usporiadaný zoznam rozhodovacích pravidiel, *GA2M*[14] nazývaný aj Explainable Boosting Machines a *SLIM*[15] Supersparse Linear Integer Model, ktorý používa techniku skórovacieho systému. Modely ako Boosted trees *Random forest* a *Xgboost* sú často presnejšie, ale aj menej interpretovateľné ako predtým uvedené modely.

B. Modelovo-agnostické metódy

Modelovo agnostické metódy oddeľujú vysvetlenia od modelu strojového učenia. To prináša výhodu oproti modelovo špecifickým metódam v ich flexibilitate [16] a univerzálnosti. Agnostické metódy možno použiť pre širokú škálu modelov strojového učenia, ako sú ansámblové metódy alebo hlboké neuronové siete. Dokonca aj výstup metódy XAI, či už ide o vizuálne alebo textové používateľské rozhranie, sa stáva nezávislým od použitého modelu strojového učenia. Jedna agnostická metóda môže vysvetliť každý z viacerých natrenovaných modelov strojového učenia a pomôcť rozhodnúť o najvhodnejšom modeli nasadenia. Tieto metódy možno ďalej rozdeliť na globálne a lokálne metódy. Globálne metódy

opisujú vplyv prvkov na model v priemere a lokálne metódy vysvetľujú model na základe predpovedí jednotlivých príkladov. Medzi tieto metódy patria hlavne dve populárne metódy LIME a SHAP.

C. Model-špecifické vysvetlenia

V tejto skupine existuje niekoľko metód XAI pracujúcich so špecifickými modelmi hlbokého učenia, napr. CNN (konvolučné neuronové siete), LSTM alebo GAN na spracovanie obrazu [17] alebo modely na spracovanie videa [18]. Existujú aj metódy XAI pre špecifické typy údajov, ako je text, hlas alebo časové rady.

1) *Class Activation Mapping (CAM)*: CAM [19] predstavuje jednu zo základných metód z vizuálnej oblasti. Na jej princípe sú založené aj ďalšie metódy. CAM pridáva medzi poslednú konvolučnú vrstvu a poslednú plne prepojenú vrstvu neuronovej siete CNN vrstvu globálneho priemerneho združovania. Plne prepojená vrstva, riadená aktivačnou funkciou softmax, nám následne na výstupe poskytuje požadované pravdepodobnosti. Dôležitosť váh vzhľadom na kategóriu môžeme získať spätnou projekciou váh na tepelné mapy z poslednej konvolučnej vrstvy. To nám umožňuje vizualizovať funkcie CNN z vrstvy, ktorá je zodpovedná za klasifikáciu.

Matematická formulácia metódy CAM: Nech $f(x, y)$ je aktivačná mapa váhy u v poslednej konvolučnej vrstve v priestorovom umiestnení (x, y) . Výstup globálnej priemernej združovacej vrstvy je prezentovaný ako:

$$F_u = \sum_{x,y} f_u(x, y)$$

Pre triedu c , vstup do softmax funkcie bude:

$$S_c = \sum_u w_u^c F(u)$$

Výstup zo softmax vrstvy:

$$P_c = \frac{e^{S_c}}{\sum_c e^{S_c}}$$

Teda finálna rovnica pre výpočet aktivačnej mapy pre triedu c bude:

$$M_c(x, y) = \sum_u w_u^c f_u(x, y)$$

CAM sú dobrou a jednoduchou technikou na interpretáciu vlastností modelov CNN. Nevýhodou tejto metódy je šum, ktorý spôsobuje stratu priestorových informácií. CAM si vyžadujú model CNN, ktorý obsahuje vrstvu GAP, a tepelné mapy CAM možno generovať len pre poslednú konvolučnú vrstvu. Preto boli vyvinuté iné algoritmy, ako napríklad Grad-CAM.

2) *Gradient-weighted Class Activation Mapping (Grad-CAM)*: Grad-CAM [20] je zovšeobecnením CAM, ktoré možno aplikovať na akýkoľvek typ CNN. Grad-CAM je použiteľný na rôzne typy architektúr CNN: CNN, VGG, DenseNet. Grad-CAM nevyžaduje vrstvu GAP a môže sa použiť na tepelné mapy pre akúkoľvek vrstvu. Rozdiel medzi CAM a Grad-CAM je vo výpočte váh pre každú

tepelnú mapu. Grad-CAM vezme mapu príznakov konvolučnej vrstvy a na základe gradientu skóre vypočíta, ktorý atribút je dôležitý pri vybranej cieľovej triede. Potom sa váhy neurónov získajú globálnym spriemerovaním gradientov. Týmto spôsobom získame váhy príznakov pre cieľovú triedu. Vynásobením máp príznakov ich váhami získame tepelnú mapu zvyrazňujúcu oblasti, ktoré pozitívne alebo negatívne ovplyvňujú záujmovú triedu. Nakoniec použijeme funkciu ReLU, ktorá nastaví záporné hodnoty na 0, pretože nás zaujímajú len pozitívne príspevky vybranej triedy. Týmto spôsobom získame mapy príznakov, ktoré zvyrazňujú dôležité oblasti vstupného obrazu pre vybranú cieľovú triedu.

IV. SPRACOVANIE VIDEOSEKVENCIÍ Z ULTRASONOGRAFIE PĽÚC

Pri analýze ultrazvukových (USG) snímok si lekár všima určité znaky (artefakty), ktoré pomáhajú posúdiť zdravotný stav pacienta. V prípade pľúc ide najmä o artefakty, ako sú *pohyb pľúc*, *A línie* a *B línie*. Pohyb pľúc je prirodzený artefakt, ktorý sa vyskytuje u ľudí. Parietálna a viscerálna pleura u zdravého jedinca tesne priliehajú k sebe, pričom tekutina medzi nimi je minimálna. Táto skutočnosť umožňuje, aby sa počas dýchania posúvali jedna po druhej. Počas diagnostického procesu môže lekár chcieť vyšetriť pacienta na identifikáciu prítomnosti pohybu pľúc alebo jeho absencie. Absencia pohybu pľúc je príznakom jedného alebo viacerých pľúcnych ochorení. Tento pohyb pleury dozadu a dopredu možno pozorovať na snímkach v *B-mode* alebo *M-mode*.

M-mode predstavuje snímky, ktoré sú 2D časovým rezom videozáznamu v oblasti pľúc. Na týchto snímkach je možné pozorovať dva vzory: *čiarový kód* alebo *plážový vzor*. Tieto vzory sú základným znakom pri klasifikácii *M-mode* snímok, či sa na danom videu vyskytuje kľzavý pohyb pľúc [21]. Lichtenstein a kol. v práci [22] klasifikovali kľzavý pohyb pľúc u pacientov prichádzajúcich na jednotku intenzívnej starostlivosti (JIS) s akútnymi respiračnými problémami, pričom použili uvedené vzory na *M-mode* snímkach. Následne v kombinácii s ďalšími sledovanými príznakmi na ultrazvuku dokázali diagnostikovať pľúcne ochorenia s vysokou presnosťou. Na obrázku 2 môžeme porovnať tieto dva vzory a príklad vybraného rezu obrazu *B-mode* a obrazu *M-mode* v čase.

A. Použitá metóda

V predkladanej práci analyzujeme *B-mode* videá z USG vyšetrenia pľúc pre klasifikáciu prítomnosti alebo neprítomnosti *kľzavého pohybu pľúc*. USG vyšetrenie pľúc je relatívne nová, ale účinná technika vyšetrenia pľúc, ktorej hlavnou výhodou je absencia zbytočného transportu pacienta, ako aj absencia ožiarenia pacienta, ku ktorému dochádza pri použití konvenčného röntgenového vyšetrenia hrudníka. Analýza takýchto snímok z ultrazvukového vyšetrenia si vyžaduje určitú úroveň odborných znalostí rádiológa, ktorý takýto systém obsluhuje [23].

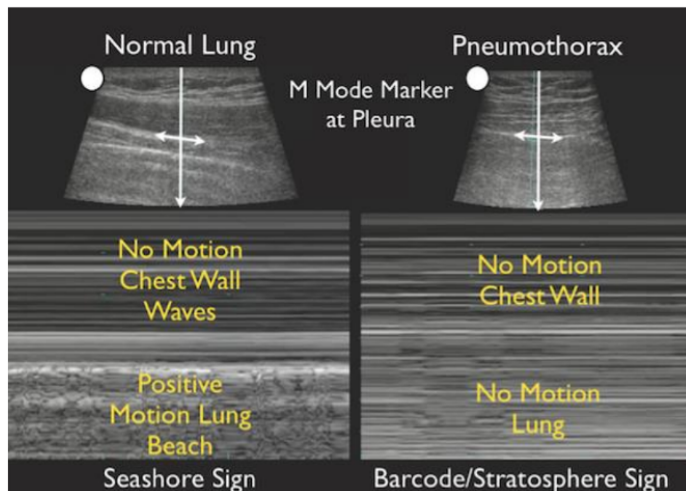


Figure 2. Príklad B-mode a M-mode snímkov s plážovým a čiarovým vzorom.

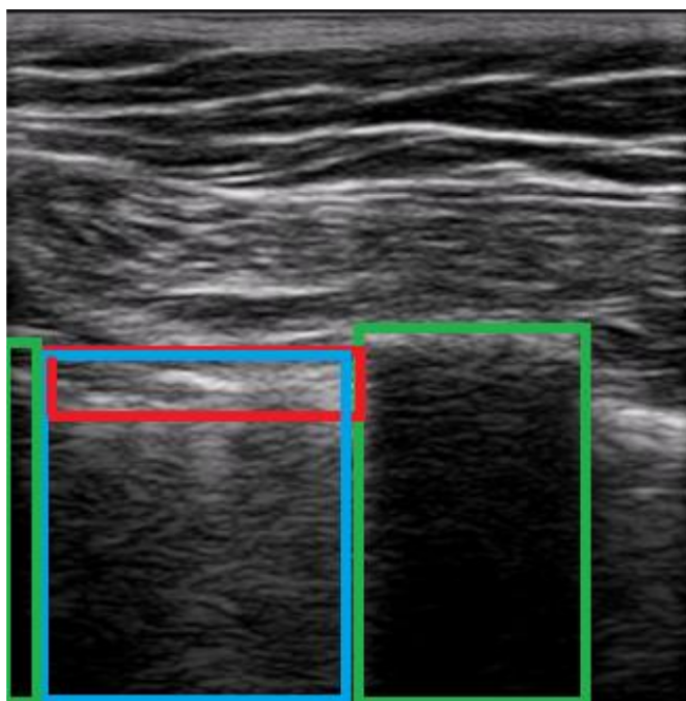


Figure 3. Snímka z USG videá s označením oblasti ohraničujúcej rebrá (zelená), pľúca (modrá) a pleuru (červená).

Na obrázku 3 vidíme, ako vyzerá videozáznam z ultrazvukového vyšetrenia pľúc. Zelene vyznačená oblasť predstavuje rebrá, modro oblasť pľúc a červeno pohrudnicu.

V tejto práci sme sa zamerali na klasifikáciu videí, ktoré obsahujú *kľzavý pohyb pľúc* a v ktorých *kľzavý pohyb pľúc* chýba. Na základe predchádzajúcich výsledkov prác využívajúcich *M-mode* obrazy sme pochopili, že pre rozhodnutie modelu je potrebný obraz v čase, ktorý vytvára vzor *pobrežia* alebo *čiarového kódu* viditeľný v celej oblasti pľúc medzi rebrami. Okrem toho, obraz 64 snímkov v priereze pľúc

je už dostatočnou vzorkou. Rozhodli sme sa využiť informácie o pohybe pľúc priamo z videozáznamu. Naopak, na klasifikáciu použijeme len malú oblasť videa, v ktorej sa nachádzajú pľúca, čo považujeme za dostatočné na klasifikáciu pohybu alebo absencie pohybu pľúc. Modely, ktoré budeme používať, sú 3D konvolučné neurónové siete.

1) *Popis dát:* V tejto práci sme použili dataset s videonímkami publikovaný v [23]. Uvedené videonímky boli získané z ultrazvukového vyšetrenia pľúc lekármi z Kliniky hrudníkovej chirurgie Jesseniovej lekárskej fakulty UK v Martine. Získané videá sú z vyšetrenia pneumotoraxu, pri ktorom pacient leží na chrbte a lekár lineárnou sondou skenuje prednú a bočnú časť hrudníka a zameriava sa na tri hlavné znaky: *pohyb pľúc*, *B línie* a *lung point*. Všetky tieto videá boli zaznamenané v režime *B-mode*, ktorý zachytáva impedanciu dvojrozmerného prierezu tkaniva. Následne lekári rozdelili tento súbor údajov do dvoch tried, a to do triedy s videami s prítomnosťou kľzavého pohybu pľúc (28 videí) a bez kľzavého pohybu pľúc (20 videí).

Predspracovanie dát: Pri spracovaní súboru videodát sme sa najprv zamerali na zjednotenie charakteristík videí. Videá sú uložené v rozličných formátoch (AVI a MP4) a z rozdielnou vzorkovacou frekvenciou (30FPS a 60FPS). Najskôr sme zjednotili videá na formát MP4 pri cieľovom rozlíšení 640x480 a frekvencii 30FPS.

V ďalšom kroku sme videá rozdelili na 2-sekundové (60 snímkov) intervaly, ktoré by mali byť dostatočné na klasifikáciu prítomnosti pohybu pľúc na základe už publikovaných prác. Týmto spôsobom sme rozšírili súbor údajov na 115 krátkych videonímkov s prítomnosťou pohybu pľúc a 48 krátkych videonímkov s neprítomnosťou pohybu pľúc.

V ďalšom kroku sme z videí extrahovali oblasti pľúc a vytvorili tak menšie snímky dostatočné pre klasifikáciu snímku. Na obrázku 3 môžeme vidieť červenú oblasť v ktorej prebiehalo označovanie pohrudnicových oblastí. Táto oblasť sa nachádza medzi rebrami, v celej šírke pohrudnice (pľúčne tkanivo medzi dvoma rebrami).

2) *Zhodnotenie výsledkov z experimentov:* Nakoniec sme sa rozhodli porovnať výsledky všetkých našich experimentov. Pri porovnaní experimentov v tabuľke I vidíme, že najlepšie výsledky dosiahol model *Resnet3D-18* na súbore údajov 30x30 pixelov z celej oblasti pľúc, model *Resnet3D-34* na údajoch 60x60 prvých dvoch riadkov a model *Resnet3D-18* na súbore údajov 90x90 pixelov z celej oblasti pľúc. Zaujímavé je, že podľa lekárov je pre rozhodovanie o pohybe pľúc dôležitá oblasť okolo pohrudnice a pohyb by mal byť viditeľný práve pri pohrudnici.

Výsledky našich experimentov sme potom porovnali so štúdiou Jaščura a kol. [23], ktorá použila *M-mode* prístup k časozberným videonímkom a ich klasifikáciu pomocou sietí Resnet. V tabuľke 5 sú uvedené výsledky porovnania našich najlepších modelov s modelom *M-mode* siete Resnet na klasifikáciu obrazov. Najlepšie fungujúci model dosiahol lepšiu presnosť a citlivosť, ale horšiu špecifickosť. V prípade citlivosti sme dosiahli o viac ako 7,5 % lepšiu presnosť. Naš model sa naučil lepšie rozpoznávať pozitívne príklady, t. j.

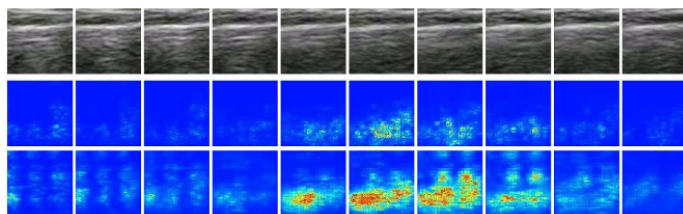


Figure 4. Example of processed frames (first row), and result of XAI methods Vanilla Saliency Map (second row) and SmoothGrad (third row) to each processed frame.

tie, pri ktorých bol na videu zaznamenaný kľzavý pohyb pľúc. Metriky sú porovnateľné s tými, ktoré dosiahli autori, pričom sme mali k dispozícii menší objem údajov pre fázu tréningu ako autori pre 64-snímkovú architektúru.

Model	Dataset	Accuracy	Sensitivity	Specificity	F1
Resnet3D-50	2_30x30	0.8420	0.8788	0.7624	0.8838
Resnet3D-18	all_30x30	0.8913	0.9360	0.7853	0.9238
Resnet3D-34	2_60x60	0.8596	0.8403	0.8985	0.8889
Resnet3D-50	all_60x60	0.7909	0.8722	0.6118	0.8516
Resnet3D-50	2_90x90	0.5541	0.5098	0.6522	0.6118
Resnet3D-18	all_90x90	0.8815	0.8768	0.8940	0.9150

Table I

TABUĽKA S NAJLEPŠÍMI DOSIAHNUTÝMI VÝSLEDKAMI.

B. Vysvetliteľnosť modelu

Po natréningu a vyhodnotení modelu sme chceli preskúmať možnosť vysvetlenia modelu. Rozhodli sme sa použiť metódu Vanilla saliency map [24], ktorá generuje tepelnú mapu, ktorá ukazuje, ktorá časť vstupného obrazu (v našom prípade ktorý pixel vo videu) ovplyvňuje výstupnú klasifikáciu. Ako druhú metódu sme zvolili SmoothGrad [25], ktorá je založená na metóde mapy salencie. Funguje na podobnom princípe, ale na vstupe svojich výpočtov tepelnej mapy pridáva šum, aby znížila šum na výstupe tejto vysvetľovacej mapy. Poskytuje tak výraznejšie oblasti, na základe ktorých sa model rozhoduje. Na obrázku 4 môžeme vidieť použité metódy vysvetľovania pre konkrétne video v priebehu času. Dané snímky sú od seba vzdialené šesť snímok, aby sa ukázala daná vysvetľovanosť pre celé video. Pôvodný model videa (vstup) je v prvom riadku. V druhom riadku sú zobrazené snímky z videa s aplikovanou Vanilla Saliency Map a v treťom riadku je uvedený výsledok SmoothGrad. V oboch prípadoch zvýraznené oblasti vizualizujú pixely, ktoré najvýznamnejšie prispievajú ku klasifikácii.

V. ZÁVER

Predložená práca prináša pohľad na dva naväzujúce publikované články v rámci doktorandského štúdia. V prvom článku, ktorého časť tvorí analýzu tejto práce, sme sa snažili priblížiť oblasť vysvetliteľnosti hlbokého učenia a potreby či požiadavky na vysvetliteľné metódy v medicínskej oblasti. Oblasť medicíny prináša veľké množstvo požiadaviek na modely umelej inteligencie, z ktorých veľká časť má priamo súvis so správnym vysvetlením rozhodovania modelov. Existencia

transparentných a vysvetliteľných modelov je v kritickom rozhodovaní v medicíne dôležitá.

V rámci druhého, prakticky orientovaného článku sme predstavili postup akým možno spracovať videozáznamy, ak poznáme charakter dát a vieme dáta vhodne predspracovať, čo nám umožní znížiť výpočtové nároky pri učení modelu, zvýšiť rýchlosť učenia, a hlavne sa zamerať na správnu oblasť pri tréningu modelu. Poskytnuté riešenie bolo diskutované s lekármi, avšak bolo obtiažné vysvetliť prečo aj veľmi malá časť videa zo správnej oblasti nesie informácie pre klasifikáciu celého videa a čo je výstupom vizualizácie z metód vysvetliteľnosti.

POĎAKOVANIE

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-20-0232 and by the Slovak VEGA research grant No.1/0685/21.

REFERENCES

- [1] M. Kolarik, M. Sarnovsky, J. Paralic, and F. Babic, "Explainability of deep learning models in medical video analysis: a survey," *PeerJ Computer Science*, vol. 9, p. e1253, 2023.
- [2] M. Kolárik, M. Sarnovský, and J. Paralič, "Detecting the absence of lung sliding in ultrasound videos using 3d convolutional neural networks," *Acta Polytechnica Hungarica*, vol. 20, no. 6, 2023.
- [3] W. Samek and K.-R. Müller, *Towards Explainable Artificial Intelligence*. Cham: Springer International Publishing, 2019, pp. 5–22. [Online]. Available: https://doi.org/10.1007/978-3-030-28954-6_1
- [4] U. Bhatt, A. Xiang, S. Sharma, A. Weller, A. Taly, Y. Jia, J. Ghosh, R. Puri, J. M. Moura, and P. Eckersley, "Explainable machine learning in deployment," *FAT* 2020 - Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 648–657, 2020.
- [5] E. Choi, A. Schuetz, W. F. Stewart, and J. Sun, "Using recurrent neural network models for early detection of heart failure onset," *Journal of the American Medical Informatics Association*, vol. 24, no. 2, pp. 361–370, 08 2016. [Online]. Available: <https://doi.org/10.1093/jamia/ocw112>
- [6] M. A. Ahmad, A. Teredesai, and C. Eckert, "Interpretable machine learning in healthcare," *Proceedings - 2018 IEEE International Conference on Healthcare Informatics, ICHI 2018*, p. 447, 2018.
- [7] G. Bologna and Y. Hayashi, "Characterization of symbolic rules embedded in deep DIMLP networks: A challenge to transparency of deep learning," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 7, no. 4, pp. 265–286, 2017.
- [8] U. J. Muehlethaler, P. Daniore, and K. N. Vokinger, "Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis," *The Lancet Digital Health*, vol. 3, no. 3, pp. e195–e203, 2021. [Online]. Available: [http://dx.doi.org/10.1016/S2589-7500\(20\)30292-2](http://dx.doi.org/10.1016/S2589-7500(20)30292-2)
- [9] US Food and Drug Administration (FDA), "Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan," *International Medical Device Regulators Forum*, no. January, pp. 1–9, 2021. [Online]. Available: <http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf>
- [10] C. Molnar, "A guide for making black box models explainable," URL: <https://christophm.github.io/interpretable-ml-book>, 2018.
- [11] F. Bodria, F. Giannotti, R. Guidotti, F. Naretto, D. Pedreschi, and S. Rinzivillo, "Benchmarking and survey of explanation methods for black box models. arxiv e-prints," *arXiv preprint arXiv:2102.13076*, 2021.
- [12] P. Linardatos, V. Papastefanopoulos, and S. Kotsiantis, "Explainable ai: A review of machine learning interpretability methods," 2021.
- [13] F. Wang and C. Rudin, "Falling rule lists," *Journal of Machine Learning Research*, vol. 38, pp. 1013–1022, 2015.
- [14] R. Caruana, Y. Lou, J. Gehrke, P. Koch, M. Sturm, and N. Elhadad, *Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-Day Readmission*. New York, NY, USA: Association for Computing Machinery, 2015, p. 1721–1730. [Online]. Available: <https://doi.org/10.1145/2783258.2788613>

- [15] B. Ustun and C. Rudin, "Supersparse linear integer models for optimized medical scoring systems," *Machine Learning*, vol. 102, no. 3, pp. 349–391, 2016.
- [16] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should i trust you?' Explaining the predictions of any classifier," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.
- [17] H. Alshazly, C. Linse, E. Barth, and T. Martinetz, "Explainable COVID-19 detection using chest CT scans and deep learning," *Sensors (Switzerland)*, vol. 21, no. 2, pp. 1–22, 2021. [Online]. Available: <http://arxiv.org/abs/2011.05317>
- [18] D. R. Chittajallu, B. Dong, P. Tunison, R. Collins, K. Wells, J. Fleshman, G. Sankaranarayanan, S. Schwaitzberg, L. Cavuoto, and A. Enquobahrie, "XAI-CBIR: Explainable ai system for content based retrieval of video frames from minimally invasive surgery videos," *Proceedings - International Symposium on Biomedical Imaging*, vol. 2019-April, no. Isbi, pp. 66–69, 2019.
- [19] B. Zhou, A. Khosla, L. A., A. Oliva, and A. Torralba, "Learning Deep Features for Discriminative Localization." *CVPR*, 2016.
- [20] R. R. S. Ramakrishna, "Grad-CAM : Visual Explanations from Deep Networks," *International Journal of Computer Vision (IJCV)*, pp. 1–23, 2019.
- [21] A. Singh, S. Sengupta, and V. Lakshminarayanan, "Explainable deep learning models in medical image analysis," *arXiv*, pp. 1–18, 2020.
- [22] D. A. Lichtenstein, "Lung ultrasound in the critically ill," *Annals of intensive care*, vol. 4, no. 1, pp. 1–12, 2014.
- [23] M. Jaščur, M. Bundzel, M. Malík, A. Dzian, N. Ferenčík, and F. Babič, "Detecting the Absence of Lung Sliding in Lung Ultrasounds Using Deep Learning," *Applied Sciences*, vol. 11, no. 15, 2021. [Online]. Available: <https://www.mdpi.com/2076-3417/11/15/6976>
- [24] K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep inside convolutional networks: Visualising image classification models and saliency maps," *2nd International Conference on Learning Representations, ICLR 2014 - Workshop Track Proceedings*, pp. 1–8, 2014.
- [25] D. Smilkov, N. Thorat, B. Kim, F. Viégas, and M. Wattenberg, "Smooth-Grad: Removing noise by adding noise," *arXiv*, 2017.

Výskum supravodivosti v silne neusporiadaných systémoch

Marek Kuzmiak

Katedra fyziky

Fakulta elektrotechniky a informatiky, TUKE

Košice, Slovensko

marek.kuzmiak@tuke.sk

Pavol Szabó

Centrum fyziky nízkych teplôt

Ústav experimentálnej fyziky SAV

Košice, Slovensko

pszabo@saske.sk

Abstrakt— Tento článok sa zaoberá výskumom vplyvu rastúcej neusporiadanosti na supravodivé vlastnosti tenkých filmov Mo_2N . Cieľom práce bolo študovať mechanizmus prechodu supravodič-izolant (SIT) v prítomnosti rastúcej neusporiadanosti, príp. rastúceho magnetického poľa v supravodičoch, ktoré sa nachádzali v tesnej blízkosti kritickej neusporiadanosti. Aplikáciu transportných a lokálnych STM meraní sme zistili, že SIT prechod v tomto systéme sa realizuje fermiónovou cestou tak v prípade rastúcej neusporiadanosti, ako aj v prípade rastúceho magnetického poľa. Pri skúmaní lokálnych supravodivých vlastností filmov Mo_2N z tesnej blízkosti kritickej neusporiadanosti sme zistili, že v ultratenkých filmoch na supravodivé vlastnosti môžu mať veľký vplyv aj pair-breaking efekty, pochádzajúce z rozhrania substrát/tenký film. V tejto práci prezentujem výsledky prvého kompletnej spektroskopického štúdia SIT prechodu na vzorkách z oboch strán tohto prechodu.

KLúčové slová— supravodivosť, prechod supravodič-izolant (SIT), rastrovací tunelová mikroskopia (STM), fermiónový scenár, bozónový scenár, pair-breaking, molybden nitríd

Abstract— This article deals with the study of the influence of increasing disorder on the superconducting properties of Mo_2N thin films. The aim of the work was to study the mechanism of the superconductor-insulator transition in the presence of increasing disorder and increasing magnetic fields in superconductors which were in close proximity to the critical disorder. By applying transport and local STM measurements, we found that the SIT transition in this system is realized through the fermionic mechanism both in the case of increasing disorder and/or increasing magnetic field. When studying the local superconducting properties of Mo_2N films from close proximity to critical disorder, we found that in ultrathin films the superconducting properties can be strongly influenced by pair breaking effects coming from the thin film/substrate interface. In this work I present the results of the first complete spectroscopic study of the SIT transition on samples from both sides of this transition.

Keywords— superconductivity, superconductor-insulator transition (SIT), scanning tunneling microscopy (STM), fermion scenario, boson scenario, pair-breaking, molybdenum nitride

I. ÚVOD

Už antický filozof Demokritos z Abdéry (460 – 370 pred n. l.) veril v existenciu malých, ďalej už nedeliteľných častíc

z ktorých sú zložené všetky veci. Dalo by sa povedať, že tieto už ďalej nedeliteľné častice by sme mohli aproximovať, ba aj chápať, ako nami ponímané atómy. Traduje sa, že práve Demokritos si ako prvý položil otázku, ktorá je aktuálna až do dnešných dní. Pýtal sa totiž, ako veľmi sa budú meniť vlastnosti materiálov (napr. jablka) v prípade, ak ich budeme zmenšovať až do ich atomárnych rozmerov. Tiež ho zaujímalo, či si tieto pôvodné materiály zachovávajú svoje základné vlastnosti (chuť, vôňu, ... esenciu jablka) aj potom, ako ich zmenšíme na spomínané najmenšie možné rozmery.

Demokritovo kladenie si otázok zamerané na potenciálne menenie vlastnosti vecí po ich zmenšení sa dostáva do popredia práve v dnešnej tzv. nanomateriálovej dobe, kedy nám súčasné poznanie a technológie umožňujú reálne manipulovať s atómami. Veď aj ja sám sa môžem pochváliť tým, že na pracovisku, na ktorom som realizoval moju dizertačnú prácu, je pri využívaní rastrovacej tunelovej mikroskopie (STM) úplne bežné a zároveň nevyhnutné mať na konci meracieho hrotu práve jeden dobre definovaný atóm, ktorý následne vypovedá o ostromi ihly, ktorou uskutočňujeme rastrovanie po povrchu vzorky, ako aj všetky spektroskopické (STS) merania v STM. Táto nanomateriálová doba umožnila zrealizovať pre nás, možno pred pár rokmi nepredstaviteľné veci, ako je napríklad spraviť z mobilného telefónu neoddeliteľného spoločníka každého jedného z nás, čo by pri jeho pôvodných rozmeroch a funkciách určite nebolo možné. A to je prítom mobilný telefón len „čerešničkou na torte“ v procese zavádzania nanotechnológie do praxe. Avšak do akých najmenších rozmerov je toto zmenšovanie materiálov (elektrotechnických súčiastok) možné? Zmenia tieto materiály svoje vlastnosti natoľko, že zo svojej novej prirodzenosti prestanú byť vhodné pre svoje pôvodné určenie?

Táto práca v sebe zahŕňa výsledky našich prác [1]–[6], v ktorých sme sa priamo zamerali na to, ako sa menili vlastnosti Mo_2N supravodivého systému, ak doň bola postupne zavádzaná neusporiadanosť (disorder). Tá sa do systému zavádzala postupným stenčovaním supravodivých tenkých filmov nami skúmaného systému od hrúbky 30 nm, až do rozmerov, kedy postupné skracovanie strednej voľnej dráhy l_{σ} v systéme prítomných elektrónov zapríčinilo, že tenké filmy začali v rozmedzí hrúbok 3 nm – 2 nm postupne prechádzať do izolačného stavu, tak ako [7]. Takýmto spôsobom došlo k prechodu supravodič-izolant (SIT), pričom o najviac „neusporiadaných“ tenkých filmoch z Mo_2N

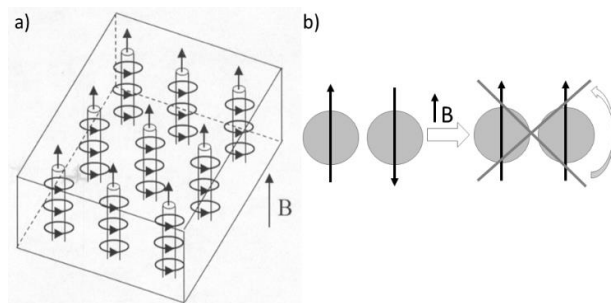
systému o hrúbke 2 nm a 1,5 nm už bolo možné povedať, že vykazovali nie supravodivé, ale väčšmi izolačné vlastnosti.

Spomínaný SIT všeobecne predstavuje kvantový fázový prechod z limity nízkych teplôt, vo všeobecnosti vyvolaný pôsobením vonkajších parametrov, akými môžu byť napríklad zavádzanie neusporiadanosti, pôsobenie vonkajšieho magnetického poľa, alebo zmena koncentrácie prítomných nosičov náboja v skúmanom systéme [8]. SIT sa môže uskutočniť podľa fermiónového, alebo bozónového scenára, pričom to, ktorý z týchto mechanizmov je v skúmanom systéme dominantný sa dá určiť podľa toho, akým spôsobom je ovplyvnený komplexný parameter usporiadania popisujúci supravodivý stav podľa vzťahu

$$\psi(\vec{r}) = \Delta e^{i\Theta(\vec{r})}. \quad (1)$$

V prípade, ak zavádzaná neusporiadanosť vplyva na supravodivý kondenzát tak, že dochádza k zmene hodnoty amplitúdy parametra usporiadania Δ , sa SIT uskutočňuje podľa fermiónového scenára [9]. Avšak v prípade, ak sa mení fáza Θ sa SIT uskutočňuje podľa bozónového scenára [10].

Je dobre známe, že aplikácia vonkajšieho magnetického poľa potláča supravodivé vlastnosti. Magnetické pole môže rozbiť Cooperove páry supravodivého kondenzátu dvoma rôznymi spôsobmi. Prvý mechanizmus rozbíjania Cooperových párov je orbitálny (orbital pair-breaking), pri ktorom dochádza k potláčaniu supravodivých vlastností prostredníctvom abrikosovových vírov, skrze ktoré preniká magnetické pole do vnútra supravodiča. Druhý mechanizmus rozbíjania Cooperových párov je paramagnetický (paramagnetic pair-breaking), ktorý sa uskutočňuje cez preklopenie spinov Cooperových párov do smeru pôsobiaceho vonkajšieho magnetického poľa, vďaka čomu je supravodivosť v systéme potláčaná skrze Zeemanovské štiepenie. Pre náš výskum je podstatné, že keď v silne neusporiadaných supravodivých tenkých filmoch v kolmom magnetickom poli pozorujeme prejav spinových Zeemanovských efektov, získavame priamy dôkaz o tom, že v systéme sa uskutočňuje potláčanie supravodivosti fermiónovou cestou. Oba mechanizmy rozbíjania Cooperových párov magnetickým poľom sú vyobrazené na Obr. 1.



Obr. 1 Potláčanie supravodivosti v supravodiči II. druhu vyvolané pôsobením vonkajšieho magnetického poľa B , zodpovedného za rozbíjanie Cooperových párov skrze tzv. orbitálny pair-breaking (a), alebo paramagnetický pair-breaking (b). Obr. (a) vyobrazuje prenikanie B do vnútra supravodiča skrze Abrikosove víry. Obr. (b) ukazuje Cooperov pár aj s jeho spinami pred, a po aplikácii B na skúmaný systém, pričom preklopenie spinov do smeru

pôsobiaceho magnetického poľa má za následok, že sa Cooperov pár rozpadne.

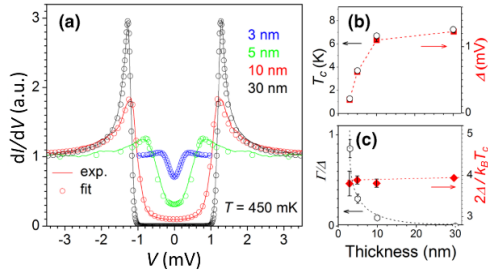
Prezentovaný článok sa zaoberá podrobným výskumom SIT, ktorý je v skúmanom Mo_2N systéme vyvolaný skrz zavádzanie „neusporiadanosti“, teda cez stenčovanie hrúbky tenkých filmov nášho systému. K získaniu experimentálnych dát boli využité transportné a STM/STS merania. Podstatným cieľom tejto práce bolo zistiť, či SIT vyvolaný rastúcou neusporiadanosťou alebo rastúcim magnetickým poľom, obzvlášť v prípade kritickej neusporiadanosti, vykazuje viac vlastností fermiónového, alebo bozónového scenára. Taktiež nás zaujímalo, či vplyv pôsobiaceho magnetického poľa na potláčanie supravodivosti v blízkosti kritickej neusporiadanosti, vykazuje viac orbitálny, alebo paramagnetický pair-breaking.

II. TEÓRIA

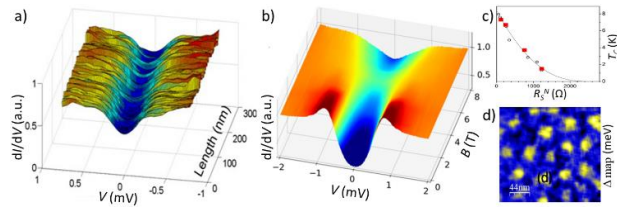
A. Typické črty fermiónového scenára SIT

V roku 2016 naša skupina z CFNT ÚEF SAV v Košiciach opublikovala článok [11], v ktorom ukázala, že v tenkých filmoch MoC sa SIT realizuje fermiónovou cestou (viď. nižšie vymenované argumenty), čo priamo naznačilo, že bozonický scenár SIT nemôže byť univerzálny, ako to bolo naznačované v [12], [13]. Naše poznatky o MoC systéme boli ešte neskôr doplnené v [14], [15]. Typické vlastnosti fermiónového SIT možno zhrnúť do nasledujúcich bodov:

1. Počas STM experimentov nebýva na tunelových spektrách dI/dV pozorovaná pseudomedzera, viď. Obr. 2 (a) [11], [14],
2. S rastúcou neusporiadanosťou sa rovnomerne znižuje kritická teplota T_C aj hodnota energetickej medzery Δ , a preto je sila supravodivej väzby $2\Delta/k_B T_C$ konštantná, viď. Obr. 2 (b, c) [11],
3. V blízkosti SIT býva pozorovaná homogénna supravodivosť, viď. Obr. 3 (a) [11], [14],
4. Býva na ňom pozorovaná mriežka Abrikosovových vírov, viď. Obr. 3 (d) [11],
5. Závislosť T_C od plošného elektrického odporu R_S celého systému je opisateľná cez Finkel'shtein model [16], viď. Obr. 3. (c) [11],
6. V hustote stavov (DOS) normálneho stavu, u nesupravodivých vzoriek a nad horným kritickým magnetickým poľom B_{C2} supravodičov, býva pozorovaná zosilnená elektrón-elektrónová interakcie - Altshuler-Aronov (AA) efekt, viď. Obr. 3 (b) [14],
7. AA efekt býva polovo závislý, pričom sa s rastúcim magnetickým poľom prehĺbuje pri hodnote Fermiho energie E_F . Táto polová závislosť je typickou vlastnosťou fermiónového systému s extrémne malou strednou voľnou dráhou l_{tr} , kde už magnetické pole vplyva na systém prostredníctvom paramagnetických efektov (paramagnetic pair-breaking) [14].



Obr. 2 (a) Typické tunelové spektrá dI/dV merané na 3 nm, 5 nm, 10 nm, 30 nm tenkom filme z MoC systému pri teplote $T = 0,45$ K v poli $B = 0$ T, na ktoré bol aplikovaný fit cez Dynesov model [17]. (b) Hrúbková závislosť tenkých filmov z MoC systému voči kritickej teplote T_C (čierne symboly) a hodnote energetickej medzery Δ (červené symboly). (c) Hrúbková závislosť parametra spektrálneho rozmazania Γ normalizovaného na Δ (čierne symboly) a sily supravodivej väzby $2\Delta/k_B T_C$ (červené symboly) [11].

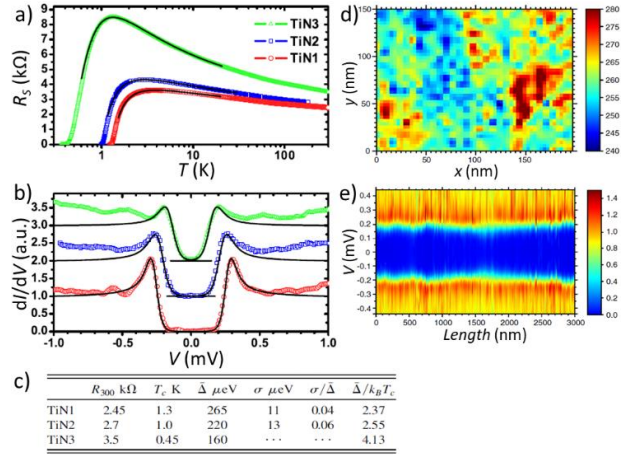


Obr. 3 Merania na MoC systéme: (a) Homogénne tunelové spektrá dI/dV merané pozdĺž 300 nm čiary topografie, na 3 nm tenkom filme, pri teplote $T = 0,45$ K v poli $B = 0$ T v blízkosti SIT [11]. (b) Poľová závislosť tunelových spektrier dI/dV meraná na 3 nm tenkom filme, pri $T = 0,5$ K s hodnotou horného kritického magnetického poľa $B_{C2} = 6,3$ T [14]. (c) Závislosť kritickej teploty od plošného elektrického odporu meraného v normálnom stave $T_C(R_{300}^N)$ určená z tenkých filmov o hrúbke od 30 nm do 3 nm dobre opísateľná pomocou Finkel'shteinového modelu [16] z [11]. (d) Mriežka 21 Abrikosovových vírov pozorovaná na 5 nm tenkom filme pri $T = 0,45$ K a $B = 1$ T [11].

B. Typické črty bozónového scenára SIT

Prítomnosť bozónového SIT bola pozorovaná na homogénne neusporiadaných tenkých filmoch TiN [18], [19] a NbN [20], ako aj na objemových vzorkách NbN [21]–[23] a InO [24]. Typické črty tohoto prechodu sú:

1. S rastúcou neusporiadanosťou sa viac potláča T_C ako Δ , a preto sila supravodivej väzby $2\Delta/k_B T_C$ narastá, vid'. Obr. 4 (a-c) [18], [19], [23], [24],
2. Blízko kritickej neusporiadanosti SIT boli pozorované priestorové nehomogenity parametra Δ , vid'. Obr. 4 (e) [18], [19], [21]–[24],
3. Počas STM experimentov býva na tunelových spektrách dI/dV na systéme v normálnom stave nad T_C pozorovaná pseudomedzera, ktorá poukazuje na prítomnosť fázovo nekoherentných Cooperových párov schopných vytvárať lokálne koherentné supravodivé oblasti/ostrovčeky o veľkosti koherenčnej dĺžky, vid'. Obr. 4 (d) [18]–[22], [24],
4. Najviac neusporiadané objemové vzorky zvyknú vykazovať potláčanie koherenčných maxim energetickej medzery v DOS, vid'. Obr. 4 (a, b) [21], [22], [24],
5. Nebýva na ňom pozorovaná mriežka Abrikosovových vírov, keďže k vytvoreniu mriežky je zapotreby prítomnosť fázovej koherencie na dlhú vzdialenosť [20].



Obr. 4 Merania na tenkých filmoch TiN o hrúbke 3,6 nm (TiN1), 5 nm (TiN2) a hrúbkovo redukovaná vzorka z pôvodne tej 5 nm (TiN3), publikované v [18]. (a) Závislosť plošného elektrického odporu od teploty $R_{300}(T)$. (b) Typické tunelové spektrá dI/dV zaznamenané na spomínaných tenkých filmoch. Hodnoty ako plošný odpor pri 300 K - R_{300} , kritická teplota T_C , šírka energetickej medzery Δ , štandardná odchýlka pre supravodivú medzeru pri viacnásobných meraniach σ a porovnanie pôsobenia supravodivej väzby $2\Delta/k_B T_C$ sú zaznamenané v Tabuľke (c). (d) Vyobrazenie Δ -mapy meranej na topografii 200×150 nm² pre TiN1 tenký film s farebne rozšíriteľnými supravodivými nehomogenitami. (e) Tunelové spektrá dI/dV merané pozdĺž 3000 nm čiary topografie na TiN2 filme.

III. EXPERIMENT

A. Výroba vzoriek

Naše vzorky vznikali metódou reaktívneho magnetronového naprašovania [25]. Z literatúry [26] vieme, že takto vzniknutý Mo_xN tenký film môže pod vplyvom pôsobenia teploty a koncentrácie N₂ pri výrobe vytvárať tetragonálnu β -Mo₂N, hexagonálnu δ -MoN a kubickú γ -Mo₂N fázu. Uskutočnením röntgenovej difrakcie (XRD) na našich vzorkách sme sa dozvedeli, že vytvárajú plošne centrovanú kubickú kryštalovú mriežku (fcc), príznačnú pre γ -Mo₂N fázu [27].

B. Transportné merania

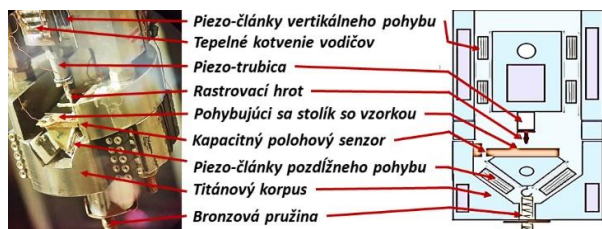
Pre uskutočnenie transportných meraní sme využívali Physical Properties Measurement System (PPMS) od spoločnosti Quantum Design, ktorý umožňuje skúmať transportné vlastnosti materiálov v teplotnom rozsahu 0,8 K – 400 K a v magnetických poliach až do 9 T. Taktiež sme využívali VTI systém od spoločnosti Oxford Instruments, ktorý pracuje v teplotnom rozsahu 1,2 K – 300 K a v magnetických poliach až do 8 T. Skonštruovaním držiaka vzorky na meranie v konfigurácii Van der Pauw, upevniteľného na ³He refrigerátor od spoločnosti Janis SSV, sme mohli merať aj v teplotnom rozsahu 0,3 K – 300 K, avšak bez možnosti využitia pôsobenia vonkajšieho magnetického poľa.

Transportné merania na našich tenkých filmoch sa uskutočnili v konfigurácii Van der Pauw [28]. Presné podmienky pre úspešné uskutočnenie meraní v tejto konfigurácii sú uvedené v článkoch [29], [30].

C. STM/STS merania

Lokálne na nanoúrovni je možné skúmať povrchy a vlastnosti materiálov pomocou rastrovacieho tunelového mikroskopu (STM). Ten náš bol skonštruovaný v našom Centre fyziky nízkych teplôt, pričom jeho konštrukčný popis,

ako aj reálny pohľad naň je vidieť na Obr. 5. Tento systém pracuje v teplotnom rozsahu 0,4 K – 300 K, v magnetickom poli až do 8 T.



Obr. 5 Reálny pohľad na našu hlavu STM aj s popisom.

Využitím STM je možné merať aj v móde rastrovacej tunelovej spektroskopie (STS), ktorá vychádza z meraní volt-ampérových (I - V) charakteristík na tunelovom prechode medzi dvoma kovovými materiálmi, ktorých študovaním je možné získať informácie o hustote elektrónových stavov (DOS) na oboch stranách tunelovej bariéry. V prípade, že na jednej strane tunelového prechodu je supravodič, z prvej derivácie I - V charakteristiky (dI/dV) môžeme priamo určiť DOS v supravodiči, kde pri energii energetickej medzery môžeme vidieť výrazné koherentné maximá v oboch polaritách a pri nižších energiách môžeme vidieť skutočnú medzeru excitačného spektra. Pomerne často býva využívaný aj CITS mód, v ktorom sa merajú spektrálne mapy, kedy sa v každom bode rastrovaním získanej topografie uskutočňuje STS meranie.

IV. DISKUSIA

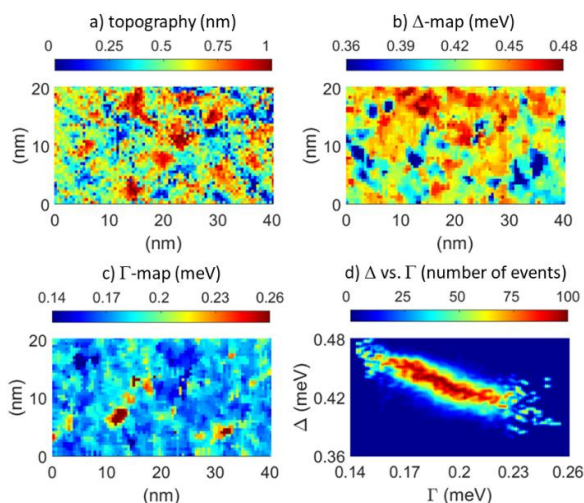
A. Vplyv substrátu na supravodivosť v ultratenkých filmoch

V prácach [1], [4] boli publikované naše výsledky namerané pomocou STM v CITS móde na 3 nm Mo_2N tenkom filme, ktoré boli veľmi podobné s výsledkami, ktoré boli zaznamenané na tenkých filmoch MoC [11], [15]. Možno povedať, že nami zaznamenaná topografia povrchu z Obr. 6 (a) vcelku dobre korelovala s lokálnou variáciou supravodivej energetickej medzery Δ z Obr. 6 (b), a antikorelovala s parametrom spektrálneho rozmazania Γ z Obr. 6 (c). V tenších oblastiach 3 nm Mo_2N tenkého filmu bola hodnota Δ menšia a Γ väčšia, ako na povrchu vyčnievajúcich oblastí. To, že lokálna zmena hodnoty energetickej medzery Δ sa menila so zmenou parametra spektrálneho rozmazania Γ možno vyčítať aj z Obr. 6 (d), ktorý znázorňuje v 3D histograme vzájomnej početnosti lineárne potláčanie Δ , ktoré sa uskutočňuje pri narastajúcom Γ . Z tejto $\Delta(\Gamma)$ závislosti je zrejmé, že s rastúcou hodnotou Γ hodnota Δ klesá. To naznačuje prítomnosť lokálnych centier pair-breakingu.

V našich ultratenkých filmoch boli najskôr zdrojmi lokálneho rozbitia Cooperových párov oblasti rozhrania substrát/tenký film, kde difúziou atómov kyslíka mohli vzniknúť nevykompenzované elektrónové orbitály, ktoré pravdepodobne boli zdrojom lokálneho magnetizmu. Navyše v [31] bolo ukázané, že zařirový substrát môže za určitých podmienok vykazovať povrchový feromagnetizmus, ktorý môže zosilňovať tieto netienené magnetické momenty. Podobný zdroj lokálneho magnetického pair-breakingu bol

detegovaný na Nb, kde substechiometrické oxidy Nb_2O_3 vytvárali na povrchu netienené magnetické momenty v d-pásme [32]. Nedávno bolo v [33] uvedené, že magnetické poruchy môžu byť prítomné aj na povrchu amorfných supravodivých filmov bez akýchkoľvek magnetických nečistôt, vychádzajúc pri tom z povrchových úprav týchto vzoriek.

Pri nazeraní na všetky uvedené argumenty sme pre ultratenký Mo_2N 2D systém usúdili, že kľúčové pre jeho vlastnosti je rozhranie medzi zařirovým substrátom a Mo_2N tenkým filmom. Práve na tomto rozhraní substrát/tenký film prítomné nevykompenzované elektrónové orbitály produkujú lokálne magnetické momenty, ktoré sú podľa všetkého zdrojom lokálneho magnetického potláčania Cooperových párov v tomto systéme (local magnetic pair-breaking). Naše výsledky poukazujú na to, že pri štúdiu supravodivých vlastností ultratenkých filmov treba brať do úvahy aj vplyv rozhrania substrát/tenký film, ktorý sa môže prejavovať veľmi podobne, ako lokalizované bozonické stavy.



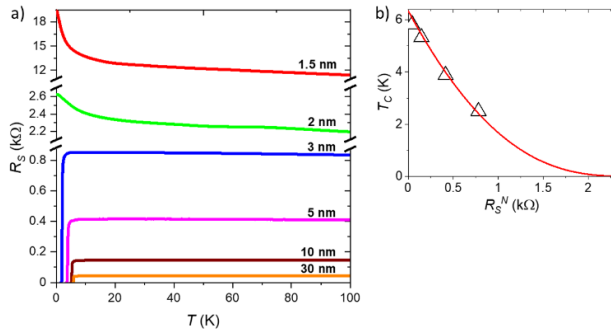
Obr. 6 Grafické výstupy z meraní na STM v CITS móde pre 3 nm Mo_2N supravodivý tenký film, ktorý sa nachádza v tesnej blízkosti SIT. Tieto merania prebiehali pri teplote $T = 0$, 5 K, aplikovanom napätí $V = 3$ mV a prúde $I = 0,4$ nA, na ploche o rozmere 40×20 nm². Na Obr. (a) je vyobrazená topografia povrchu, na ktorom sa uskutočnili CITS merania. Fitovaním diferenciálnych tunelových spektier cez Dynesov model [17] boli vykreslené mapy, zahrňujúce v sebe informáciu o prislúchajúcej hodnote energetickej medzery - Δ -mapa (b) a spektrálnom rozmazaní - Γ -mapa (c). Vyobrazenie početnosti závislosti Δ od Γ z Obr. (b, c) pre všetky zaznamenané tunelové spektrá získané z merania v CITS móde je na Obr. (d).

B. SIT študovaný pomocou transportných a STM meraní

V prácach [2], [6] sme uverejnili výsledky z transportných meraní na tenkých filmoch z Mo_2N systému, do ktorých bola stencovaním ich hrúbky zavádzaná neusporiadanosť (disorder). Vykonaním transportných meraní sme zostrojili graf závislosti plošného elektrického odporu od teploty $R_S(T)$ na tenkých filmoch z tohto systému, ktorý je vyobrazený na Obr. 7 (a). Prítomná najhrubšia 30 nm vzorka vykazovala počas merania závislosti plošného elektrického odporu v normálnom stave od teploty $R_S^N(T)$ teplotne nezávislú hodnotu, ktorá sa začala znižovať tesne pred dosiahnutím prechodu do supravodiveho stavu, pri kritickej teplote $T_C = 5,83$ K. Následným postupným znižovaním hrúbky nášho systému, bolo pozorované zvyšovanie R_S^N . U tenkých

filmov, ktoré vykazovali supravodivý stav (30 – 3 nm), spôsobilo prítomné zredukované ich hrúbky aj pokles hodnoty im príručnej T_C , a to až na 2,8 K pri 3 nm vzorke. V prípade 2 a 1,5 nm vzoriek sa evidentne v porovnaní so supravodivými tenkými filmami mení ich krivosť $R_S(T)$, a pri najnižších teplotách dochádza k výraznému nárastu hodnoty R_S , čo naznačuje, že systém postupne prechádza do izolačného stavu.

Obr. 7 (b) zobrazuje závislosť kritickej teploty od plošného elektrického odporu v normálnom stave $T_C(R_S^N)$ jednotlivých vzoriek (čierne symboly Δ). Tieto hodnoty boli následne prefitované pomocou funkcie z Finkel'shteinovo modelu [16]. Fitovacia funkcia je vo vloženom obrázku vyobrazená červenou farbou. Využitím tohoto modelu sme získali hodnotu parametra kritickej teploty supravodivého prechodu pre objemový vzorku z nášho systému $T_{C0} = 6,34$ K, priemernú dobu medzi jednotlivými zrážkami častíc v systéme $\tau_{cr} = 2,66 \times 10^{-16}$ s, a tiež sme vytčali hodnotu elektrického odporu pri kritickej neusporiadanosti v našom systéme $R_{cr} = 2,3$ k Ω . To, že Mo₂N systém je opisateľný cez Finkel'shteinov model [16] naznačuje, že sa tento systém riadi *fermiónovým scenárom SIT*.

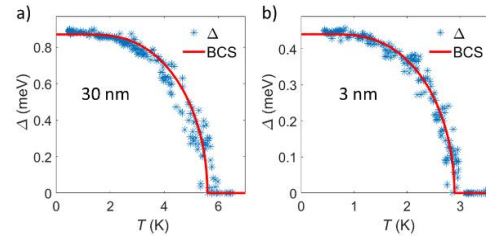


Obr. 7 Neusporiadanosťou indukovaný SIT v tenkých filmoch Mo₂N o hrúbke 30 - 1,5 nm. Na Obr. (a) sú zachytené závislosti plošného elektrického odporu od teploty $R_S(T)$ pre tenké filmy o rôznych hrúbkach, pričom príslušajúca hrúbka d zobrazovaného tenkého filmu je uvedená nad konkrétnou $R_S(T)$ závislosťou. Na Obr. (b) je zobrazená závislosť kritickej teploty od plošného elektrického odporu v normálnom stave $T_C(R_S^N)$ pre tu zobrazené supravodivé tenké filmy o hrúbkach 30 – 3 nm. Červenou farbou je vyobrazená aplikovaná fitovacia funkcia z Finkel'shteinovo modelu [16], prispôbená na nameranú $T_C(R_S^N)$ závislosť.

V ďalšom kroku sme na Mo₂N systéme uskutočnili merania pomocou STM. Hneď po nájdení čistých dobre definovaných povrchov s nízkou úrovňou korugácie boli na našich tenkých filmoch vykonané merania v režime STS, z ktorých boli získané teplotné (v konštantnom magnetickom poli), ale aj pol'ové (pri konštantnej teplote) závislosti.

U teplotných závislostí tunelových spektier 30 nm a 3 nm supravodivých tenkých filmov bolo možné s rastúcou teplotou pozorovať postupné potlačanie supravodivej medzery, ktorá bola po stranách ohraničená koherenčnými píkmí, a to až po hodnotu T_C . Nad touto hodnotou sa vyskytovali už iba tunelové spektrá normálneho stavu [1], [2], [4]. Pre následné presné určenie hodnoty energetickej medzery Δ , ako aj kritickej teploty T_C bolo použité fitovanie nameraných tunelových spektier z teplotných závislostí na Dynesovu formulu [17]. Takýmto spôsobom sa nám numericky podarilo získať závislosť $\Delta(T)$, ktorá je zobrazená aj na Obr. 8, pre prípad zaznamenaný na tenkom filme o hrúbke 30 nm (a

a 3 nm (b) v poli $B = 0$ T. Závislosť $\Delta(T)$ sme následne prefitovali podľa BCS teórie, čím sme získali príslušajúce hodnoty Δ a T_C pre skúmané tenké filmy. Zo získaných údajov sme vyjadrili aj silu supravodivej väzby $SC = 2\Delta/k_B T_C$. Pre 30 nm vzorku sme tak získali hodnoty $\Delta(0) = 0,87$ meV, $T_C = 5,6$ K a pre 3 nm tenký film $\Delta(0) = 0,44$ meV, $T_C = 2,9$ K. Pri 30 nm a 3 nm supravodivom tenkom filme bola pre oba tenké filmy zaznamenaná konštantná hodnota sily supravodivej väzby na úrovni $2\Delta/k_B T_C \sim 3,55$ -3,6. Táto konštantnosť súvisí s tým, že so zavádzaním neusporiadanosti do systému dochádza k rovnomernému poklesu hodnôt parametrov Δ a T_C , čo taktiež svedčí o tom, že v Mo₂N systéme sa *SIT uskutočňuje fermiónovým spôsobom*.

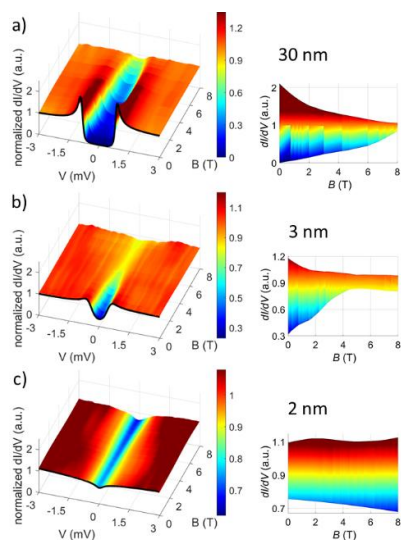


Obr. 8 Teplotné závislosti supravodivej energetickej medzery $\Delta(T)$, získané z meraní na tenkých filmoch Mo₂N o hrúbke 30 nm (a) a 3 nm (b) (modré symboly) v poli $B = 0$ T. Príslušné hodnoty Δ (modrý symbol *) boli získané fitovaním teplotne závislých tunelových spektier dI/dV podľa Dynesovo modelu [17]. Z tohoto modelu sme získali aj hodnoty parametra spektrálneho rozmazania Γ , ktorého hodnota sa počas fitovania líšila približne o 5 %. Hodnoty $\Delta(T)$ boli prefitované podľa BCS teórie (červená čiara), čím sa získali hodnoty počiatočných parametrov $\Delta(0)$ a T_C .

V článku [2] sú popísané pol'ovo závislé merania tunelových spektier na supravodivých (30 a 3 nm) a nesupravodivých (2 nm) tenkých filmoch, vid'. Obr. 9. U najhrubšieho, 30 nm supravodivého tenkého filmu, vid'. Obr. 9 (a), dochádza pod vplyvom narastajúceho pôsobenia magnetického poľa k postupnému uzatváraniu energetickej medzery Δ , ako aj nárastu vodivosti pri nulovom predpätí ZBC (meraná pri Fermiho energii E_F) spôsobom, ktorý je typický pre väčšinu supravodičov II. typu v rastúcom magnetickom poli. Vplyv pôsobenia vonkajšieho kolmého magnetického poľa na 3 nm supravodivý tenký film Mo₂N je vyobrazený na Obr. 9 (b). Na tomto grafe vidieť, že s narastajúcim magnetickým poľom dochádza k potlačaniu ZBC, avšak šírka energetickej medzery Δ sa v tomto prípade akoby zachováva, a to až po dosiahnutie hodnoty horného kritickeho magnetického poľa $B_{C2} \sim 4 - 5$ T [6], kedy sa v systéme začína objavovať nový efekt. Nad touto hodnotou B_{C2} sa totiž začne objavovať medzera, ktorej prítomnosť sa s rastúcim magnetickým poľom zvyrazňuje [34].

U nesupravodivých vzoriek z Mo₂N systému s hrúbkou 2 nm a 1,5 nm bol badateľný zosilňujúci účinok narastajúceho magnetického poľa na zníženie vodivosti tunelovania dI/dV pri ZBC. Konkrétne to bolo možné vidieť na Obr. 9 (c), kde sa nachádza pol'ová závislosť tunelových spektier 2 nm tenkého filmu Mo₂N. Analýzou sme zistili, že s rastúcim magnetickým poľom, v škále 0–8 T, klesá hodnota vodivosti pri nulovom predpätí ZBC (na E_F) približne o ~ 10 %. Taktiež všetky vyobrazené tunelové spektrá vykazovali zreteľnú logaritmickú redukciu DOS smerom k E_F , čo je typické pre AA efekt. Navyše sa s narastajúcim poľom zväčšuje akoby aj vnútorná šírka týchto spektier v smere ich novej závislosti od Zeemanovskej energie E_Z . Pozorovaná pol'ová závislosť

AA efektu je typickou vlastnosťou *fermiónového systému* s extrémne malou strednou voľnou dráhou l_r , kde už magnetické pole vplyva na systém prostredníctvom paramagnetických efektov (paramagnetic pair-breaking).



Obr. 9 V kolmom poli merané poľové závislosti tunelových spektier dI/dV pre tenké filmy o hrúbke 30 nm (a), 3 nm (b) a 2 nm (c) pri konštantnej teplote $T = 0.5$ K. Vľavo je vyobrazený 3D graf, kdežto obrázky priložené vpravo predstavujú pohľad z boku na tieto závislosti.

V. ZÁVER

Poznatky odprezentované v tomto článku prinášajú nové výsledky z oblasti základného výskumu neusporiadaných tenkých filmov z Mo_2N systému. Ako bolo aj v tomto článku odprezentované, v tomto systéme sa uskutočňuje mechanizmus prechodu supravodič-izolant (SIT) *fermiónovou cestou*. Tiež sa ukázalo, že s narastajúcou neusporiadanosťou v tomto systéme sa vplyvom pôsobenia vonkajšieho magnetického poľa uplatňuje potlačanie supravodivosti väčšími paramagnetickým, ako orbitálnym pair-breakingom, ktorý prevládal skôr u menej neusporiadaných tenkých filmov. Celkovo naše výsledky dosiahnuté pri štúdiu vytýčenej problematiky reprezentujú prvé kompletne štúdium fermiónového mechanizmu SIT pomocou lokálnych spektroskopických meraní (STS), ktoré bolo na skúmanom systéme vykonané z oboch strán prechodu, t.j. na supravodivom 3 nm a nesupravodivom 2 nm tenkom filme.

POĎAKOVANIE

Činnosť uskutočnená pri našom výskume bola vo veľkej miere podporená Ministerstvom školstva, vedy, výskumu a športu SR a Slovenskou akadémiou vied na základe zmlúv č. VEGA 1/0743/19, VEGA 2/0058/20, APVV-18-0358, APVV-20-0425, Európskej platformy Microkelvin, akcie COST CA21144 a spoločnosťou U. S. Steel Košice.

REFERENCIE

[1] M. Kuzmiak *et al.*, “Suppressed Superconductivity in Ultrathin Mo_2N Films due to Pair-Breaking at the

Interface,” *J Supercond Nov Magn*, vol. 35, no. 7, pp. 1775–1780, Jul. 2022, doi: 10.1007/S10948-022-06197-6/FIGURES/3.

- [2] M. Kuzmiak and et al., “Disorder- and magnetic field-tuned fermionic superconductor-insulator transition in Mo_2N thin films. Transport and STM studies,” *zaslané do Physical Review B*, 2023.
- [3] M. Kuzmiak, “Research of superconductivity in strongly disordered systems by scanning tunneling microscopy,” in *SCYR 2020*, Košice: TUKE, Apr. 2020, pp. 151–154.
- [4] M. Kuzmiak, “Superconducting properties of strongly disordered molybdenum nitride ultrathin films in the vicinity of superconductor-insulator transition,” in *SCYR 2021*, Košice: TUKE, Apr. 2021, pp. 39–41.
- [5] M. Kuzmiak, “Experimental study of the vortex lattice in strong disordered ultrathin 3 nm Mo_2N film,” in *SCYR 2022*, Košice: TUKE, Apr. 2022, pp. 159–161.
- [6] M. Kuzmiak, “Effect of increased disorder on the pair-breaking mechanism in Mo_2N thin films,” in *SCYR 2023*, Košice: TUKE, Apr. 2023, pp. 73–75.
- [7] D. B. Haviland, Y. Liu, and A. M. Goldman, “Onset of superconductivity in the two-dimensional limit,” *Phys Rev Lett*, vol. 62, no. 18, p. 2180, May 1989, doi: 10.1103/PhysRevLett.62.2180.
- [8] F. Wang, J. Biscaras, A. Erb, and A. Shukla, “Superconductor-insulator transition in space charge doped one unit cell $\text{Bi}_2.1\text{Sr}_{1.9}\text{CaCu}_2\text{O}_{8+x}$,” *Nature Communications 2021 12:1*, vol. 12, no. 1, pp. 1–6, May 2021, doi: 10.1038/s41467-021-23183-z.
- [9] A. M. Finkel’stein, “Metal-Insulator Transition in Disordered Two-Dimensional Electron Systems,” *Sov. Phys. JETP*, vol. 59, no. 1, pp. 212–219, 1984.
- [10] M. P. A. Fisher, “Quantum phase transitions in disordered two-dimensional superconductors,” *Phys. Rev. Lett.*, vol. 65, no. 7, pp. 923–927, Aug. 1990, doi: 10.1103/PhysRevLett.65.923.
- [11] P. Szabó *et al.*, “Fermionic scenario for the destruction of superconductivity in ultrathin MoC films evidenced by STM measurements,” *Phys Rev B*, vol. 93, no. 1, p. 014505, Jan. 2016, doi: 10.1103/PHYSREVB.93.014505/FIGURES/4/MEDIUM.
- [12] N. Trivedi, “10 Superconductor-Insulator Transitions: Present Status and Open Questions,” *Conductor-Insulator Quantum Phase Transitions*, vol. 9780199592593, Sep. 2012, doi: 10.1093/ACPROF:OSO/9780199592593.003.0010.
- [13] B. Sacépé, M. Feigel’man, and T. M. Klapwijk, “Quantum breakdown of superconductivity in low-dimensional materials,” *Nature Physics 2020 16:7*, vol. 16, no. 7, pp. 734–746, Jul. 2020, doi: 10.1038/s41567-020-0905-x.
- [14] M. Žemlička *et al.*, “Zeeman-driven superconductor-insulator transition in strongly disordered MoC films: Scanning tunneling microscopy and transport studies in a transverse magnetic field,” *Phys Rev B*, vol. 102, no. 18, p. 180508, Nov. 2020, doi: 10.1103/PHYSREVB.102.180508/FIGURES/3/MEDIUM.

- [15] V. Hašková *et al.*, “On the origin of in-gap states in homogeneously disordered ultrathin films. MoC case,” *Appl Surf Sci*, vol. 461, pp. 143–148, Dec. 2018, doi: 10.1016/J.APSUSC.2018.06.228.
- [16] A. M. Finkel’stein, “Superconducting transition temperature in amorphous films,” *JETP Lett*, vol. 45, p. 46, 1987, Accessed: Dec. 09, 2022. [Online]. Available: <https://ui.adsabs.harvard.edu/abs/1987PZETF..45..37F/abstract>
- [17] R. C. Dynes, V. Narayanamurti, and J. P. Garno, “Direct Measurement of Quasiparticle-Lifetime Broadening in a Strong-Coupled Superconductor,” *Phys. Rev. Lett.*, vol. 41, no. 21, p. 1509, Nov. 1978, doi: 10.1103/PhysRevLett.41.1509.
- [18] B. Sacépé, C. Chapelier, T. I. Baturina, V. M. Vinokur, M. R. Baklanov, and M. Sanquer, “Disorder-induced inhomogeneities of the superconducting state close to the superconductor-insulator transition,” *Phys Rev Lett*, vol. 101, no. 15, p. 157006, Oct. 2008, doi: 10.1103/PHYSREVLETT.101.157006/FIGURES/3/MEDIUM.
- [19] B. Sacépé, C. Chapelier, T. I. Baturina, V. M. Vinokur, M. R. Baklanov, and M. Sanquer, “Pseudogap in a thin film of a conventional superconductor,” *Nat Commun*, vol. 1, no. 1, pp. 1–6, Dec. 2010, doi: 10.1038/ncomms1140.
- [20] Y. Noat *et al.*, “Unconventional superconductivity in ultrathin superconducting NbN films studied by scanning tunneling spectroscopy,” *Phys Rev B Condens Matter Mater Phys*, vol. 88, no. 1, p. 014503, Jul. 2013, doi: 10.1103/PHYSREVB.88.014503/FIGURES/7/MEDIUM.
- [21] M. Chand *et al.*, “Phase diagram of the strongly disordered s-wave superconductor NbN close to the metal-insulator transition,” *Phys Rev B Condens Matter Mater Phys*, vol. 85, no. 1, p. 014508, Jan. 2012, doi: 10.1103/PHYSREVB.85.014508/FIGURES/9/MEDIUM.
- [22] M. Mondal *et al.*, “Phase fluctuations in a strongly disordered s-wave nbn superconductor close to the metal-insulator transition,” *Phys Rev Lett*, vol. 106, no. 4, p. 047001, Jan. 2011, doi: 10.1103/PHYSREVLETT.106.047001/FIGURES/3/MEDIUM.
- [23] A. Kamlapure, T. Das, S. C. Ganguli, J. B. Parmar, S. Bhattacharyya, and P. Raychaudhuri, “Emergence of nanoscale inhomogeneity in the superconducting state of a homogeneously disordered conventional superconductor,” *Sci Rep*, vol. 3, no. 1, pp. 1–8, Oct. 2013, doi: 10.1038/srep02979.
- [24] B. Sacépé *et al.*, “Localization of preformed Cooper pairs in disordered superconductors,” *Nature Physics* 2010 7:3, vol. 7, no. 3, pp. 239–244, Jan. 2011, doi: 10.1038/nphys1892.
- [25] M. Kozejova *et al.*, “Evaluation of hydrogen evolution reaction activity of molybdenum nitride thin films on their nitrogen content,” *Electrochim Acta*, vol. 315, pp. 9–16, Aug. 2019, doi: 10.1016/J.ELECTACTA.2019.05.097.
- [26] K. Inumaru, K. Baba, and S. Yamanaka, “Structural distortion and suppression of superconductivity in stoichiometric B1-MoN epitaxial thin films,” *Phys Rev B Condens Matter Mater Phys*, vol. 73, no. 5, p. 052504, Feb. 2006, doi: 10.1103/PHYSREVB.73.052504/FIGURES/2/MEDIUM.
- [27] H. Ihara, Y. Kimura, K. Senzaki, H. Kezuka, and M. Hirabayashi, “Electronic structures of B1 MoN, fcc Mo2N, and hexagonal MoN,” *Phys Rev B*, vol. 31, no. 5, p. 3177, Mar. 1985, doi: 10.1103/PhysRevB.31.3177.
- [28] L. J. van der Pauw, “A method of measuring specific resistivity and Hall effect of discs of arbitrary shape,” *Philips Res. Rep.*, vol. 13, no. 1, pp. 1–9, Feb. 1958.
- [29] S. H. N. Lim, D. R. McKenzie, and M. M. M. Bilek, “Van der Pauw method for measuring resistivity of a plane sample with distant boundaries,” *Review of Scientific Instruments*, vol. 80, no. 7, Jul. 2009, doi: 10.1063/1.3183503/351980.
- [30] S. K. Pandey and A. Manivannan, “A fully automated temperature-dependent resistance measurement setup using van der Pauw method,” *Review of Scientific Instruments*, vol. 89, no. 3, Mar. 2018, doi: 10.1063/1.4998340/368154.
- [31] A. Sundaresan and C. N. R. Rao, “Implications and consequences of ferromagnetism universally exhibited by inorganic nanoparticles,” *Solid State Commun.*, vol. 149, no. 29–30, pp. 1197–1200, Aug. 2009, doi: 10.1016/J.SSC.2009.04.028.
- [32] T. Proslie *et al.*, “Tunneling study of cavity grade Nb: Possible magnetic scattering at the surface,” *Appl. Phys. Lett.*, vol. 92, no. 21, p. 212505, May 2008, doi: 10.1063/1.2913764.
- [33] I. Tamir, M. Trahms, F. Gorniaczyk, F. Von Oppen, D. Shahar, and K. J. Franke, “Direct observation of intrinsic surface magnetic disorder in amorphous superconducting films,” *Phys Rev B*, vol. 105, no. 14, p. L140505, Apr. 2022, doi: 10.1103/PHYSREVB.105.L140505/FIGURES/4/MEDIUM.
- [34] M. Eltschka *et al.*, “Superconducting scanning tunneling microscopy tips in a magnetic field: Geometry-controlled order of the phase transition,” *Appl Phys Lett*, vol. 107, no. 12, Sep. 2015, doi: 10.1063/1.4931359/29083.

SKÚMANIE OBMEDZENÍ PRENOSU DÁT V HYBRIDNÝCH FSO/RF SYSTÉMOCH

Maroš Lapčák

Katedra elektroniky a multimediálnych telekomunikácií
Technická univerzita v Košiciach
Košice, Slovenská republika
maros.lapcak@tuke.sk

Ľuboš Ovseník

Katedra elektroniky a multimediálnych telekomunikácií
Technická univerzita v Košiciach
Košice, Slovenská republika
lubos.ovsenik@tuke.sk

Abstract—Tento článok sa zaoberá problematikou bezvláknovej optickej komunikácie a hybridnými FSO/RF systémami. V dôsledku vplyvu poveternostných podmienok na optickú komunikáciu vo voľnom prostredí je v práci uvedený spôsob predikcie parametra RSSI (indikátor prijatého výkonu), ktorý je rozhodujúcim faktorom v metóde tvrdého prepínania medzi oboma linkami. Okrem samotnej predikcie je potrebné skúmať a zaznamenávať poveternostné vplyvy pomocou meteorostanice. Ďalšou dôležitou súčasťou hybridného systému je záložná RF (rádiofrekvenčná) linka. Tá slúži ako záloha v prípade, že FSO linka prestane komunikovať, teda ak hodnota parametra RSSI klesne pod prahovú hodnotu – 42 dBm. Pre predikciu parametra RSSI boli využité metódy Decision Tree a Decision Tree s AdaBoost regresorom. Po testovaní všetkých možných vstupných množín dát bola zvolená metóda s využitím regresora, ktorá dosahovala lepšie výsledky. Výstupom je teda fungujúci systém na predikciu parametra RSSI. V prípade záložnej RF linky bolo vytvorených niekoľko návrhov. Najlepšie sa javilo pole mikropásových antén. Práve tento návrh je možné v budúcnosti využiť pre výrobu sekundárnej RF linky.

Keywords— FSO, hybridný FSO/RF systém, optická komunikácia voľným prostredím, rádiová komunikácia, RSSI, strojové učenie

I. BEZVLÁKNOVÁ OPTICKÁ KOMUNIKÁCIA (FSO)

„Optika voľného priestoru“ bola navrhnutá ako životaschopné riešenie pre vyššie prenosové rýchlosti a stala sa konkurenciou pre tradičné komunikačné systémy. FSO systém pracuje v nelicencovanom pásme so schopnosťou poskytovať služby až do vzdialenosti 15 - 16 km. Nedávno, v roku 2018, nemecké letecké a vesmírne centrum v spolupráci s optickými sieťami ADVA dosiahlo rekordných 13,16 Tb/s v prevádzke na 10,45 km [1]. Napriek tomu, že FSO systém môže dosiahnuť veľmi vysoké prenosové rýchlosti za nízke náklady, má aj svoje vlastné nevýhody. FSO teda nemôže nahradiť RF systém úplne všade. Vplyv atmosféry nie je na FSO a RF rovnaký. Útlm je spôsobený predovšetkým rozptylom a absorpciou. Mraky, sneh a dážď sú príčinami rozptylu, zatiaľ čo absorpcia signálu FSO je spôsobená hlavne oxidom uhličitým a vodnými časticami. Dôvodom významného útlmu systému FSO v dôsledku rozptylu je veľkosť častice hmly, ktorá je porovnateľná s prevádzkovou vlnovou dĺžkou optického signálu. Tento druh rozptylu sa nazýva Mieho rozptyl a je

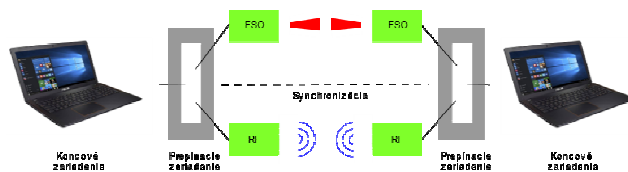
všeobecne použiteľný na horizontálne väzby v blízkosti zemského povrchu. Signál FSO nie je citlivý na dážď, ale v hmle sa výrazne zhoršuje. V prípade RF signálu sú frekvencie pod 10 GHz ovplyvnené iba kvapôčkami hmly, zatiaľ čo pri frekvenciách nad 10 GHz dochádza k výraznému rozptylu [2]. RF signál nie je veľmi citlivý na hmlu, ale jeho výkon sa v daždi zhoršuje. FSO a RF teda môžu byť dobrými doplnkovými systémami na spoločné nasadenie [1, 68].



Obr. 1 Princíp komunikácie FSO systému.

II. HYBRIDNÝ FSO/RF SYSTÉM

Hybridné FSO/RF (rádiofrekvenčné) systémy sú považované za konkurencie schopné riešenia vzhľadom k vysokorychlostným bezdrôtovým komunikáciám typu point-to-point. Dokážu v sebe kombinovať výhody týchto spojení, teda vysokej prenosovej rýchlosti systémami FSO a spoľahlivosti sekundárnych RF liniek. Tieto systémy poskytujú v súčasnosti rýchlosť prenosu od 10 Mb/s až do niekoľkých Gb/s. Hybridný FSO/RF systém, zobrazený na Obr. 2, sa zvažuje aj pre mobilné aplikácie, napríklad pre diaľkové ovládanie robotov [43, 73].



Obr. 2 Princíp fungovania hybridného FSO/RF systému.

Systémy FSO a RF majú odlišné nosné frekvencie a aj v dôsledku toho rozdielne poruchy prenosových kanálov. Krátke vlnové dĺžky v linkách FSO sú závažne ovplyvnené rozptylom v dôsledku hmly a atmosférických scintilácií. Prenosové linky RF sú na rozdiel od FSO obzvlášť citlivé na rozptyl spôsobený dažďom. Tieto faktory spôsobujú, že

kombinácia s FSO linkou nie je konvenčnou technológiou v porovnaní s komunikáciou pomocou optických vlákien. Ak sa na to pozrieme z druhej strany, komunikácia RF sa dokáže vysporiadať s poveternostnými podmienkami. Vytvorenie hybridného FSO/RF systému dokáže kombinovať výhody oboch týchto technológií, ktorými sú nízke náklady na prevádzku siete, aplikácie pracujúce bez dohľadu, vysoké prenosové rýchlosti dát a relatívne nízka latencia. Veľkou nevýhodou je, že RF linky neposkytujú takú veľkú kapacitu kanálov ako optické komunikácie. Ďalšou nevýhodou RF systému je nedostatočná šírka pásma kvôli licencovaniu. Tak isto sú citlivé na rušenie a zraniteľné pri odpočúvaní. V hybridných FSO/RF systémoch existujú dve konfigurácie týchto sietí založené na princípe prepínania [43, 73].

A. Poveternostné vplyvy na FSO a RF systémy

Jednou z hlavných výhod kombinácie FSO a RF systémov je eliminácia negatívnych javov spôsobených poveternostnými podmienkami, keďže obe linky komunikujú cez atmosférický kanál. Každá z týchto technológií je ovplyvnená inými javmi. FSO systémy sú najnáchylnejšie na:

- hmlu,
- opar,
- koncentráciu častíc,
- viditeľnosť.

Naopak, RF systémy sú voči niektorým týmto vplyvom odolné, avšak aj na tento systém majú vplyv niektoré poveternostné podmienky. Ide o:

- dážď,
- rozptyl,
- difrakciu,
- factory straty šírenia,
- odraz.

Práve v dôsledku kombinácie týchto dvoch systémov vieme zabezpečiť takmer 100% funkčnosť systému.

B. Spôsoby prepínania v hybridnom FSO/RF systéme

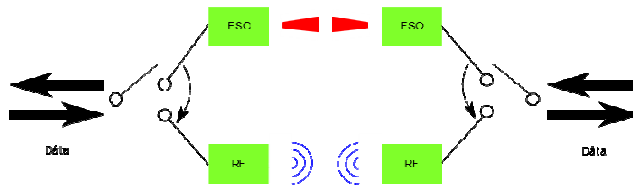
Pre zlepšenie výkonu linky FSO a zvýšenie spoľahlivosti pri rôznych poveternostných podmienkach je možné kombinovať so systémami RF, aby bolo možné vytvoriť hybridný systém. Ten má dve základné konfigurácie prepínania medzi linkami, a to:

- tvrdé prepínanie,
- mäkké prepínanie.

1) Tvrdé prepínanie

Pri konfigurácii s tvrdým prepínaním vysielača a prijímača strana spolu vyberú jeden z kanálov, ktorý bude dáta odosielať, teda buď FSO alebo RF linku. Komunikácia pri tomto druhu prepínania prebieha vždy len po jednej prenosovej trase. Primárne je však realizované spojenie pomocou FSO linky, pričom kanál RF tvorí iba záložnú linku. V tomto prípade dochádza k spätnej väzbe medzi prijímačom a

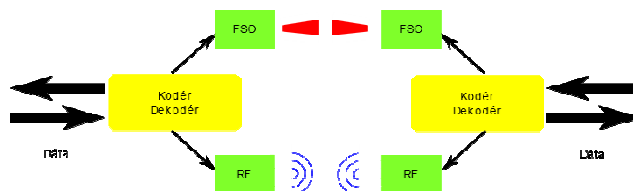
vysielačom, čím dochádza ku koordinácii prepínania liniek na oboch stranách. Jednou z najväčších nevýhod tvrdého prepínania je zložitosť výpočtov v dôsledku synchronizácie a potreby spätnej väzby. Ako funguje tvrdé prepínanie, je zobrazené na Obr. 3 [43, 56].



Obr. 3 Spôsob tvrdého prepínania hybridného FSO/RF systému.

2) Mäkké prepínanie

Hlavnou myšlienkou mäkkého prepínania, zobrazeného na Obr. 4, je potreba prekonania nevýhody tvrdého prepínania. Preto dochádza v tomto spôsobe ku koordinácii prenosu údajov na oboch linkách pomocou kódovania kanálov. Prenášané dáta sú kódované pomocou LDPC (low-density parity check) kódu s časťou kódového slova, ktorá je rozdelená pre FSO a RF linku a samotná rýchlosť sa nastaví pomocou punktovania v závislosti od aktuálnych podmienok na kanáli. Konfigurácia s mäkkým prepínaním dokáže poskytnúť lepší výkon ako je to v prípade s tvrdým prepínaním, najmä s nízkym pomerom SNR. Na druhej strane, prevádzka takéhoto systému je oveľa zložitejšia, pretože si vyžaduje kombináciu viacerých techník na koncovej strane. Je však preukázané, že pri hybridnom systéme by mal byť zlepšený výkon, avšak len pri priamej viditeľnosti spojenia (LOS) medzi oboma stranami a na vzdialenosť niekoľkých km. Aj keď dochádza k zlepšovaniu techník mäkkého prepínania, podmienky na strane prijímača i vysielača musia byť vopred známe a v dôsledku toho sa vyžaduje zložité dekódovanie, čo je pri vysokých prenosových rýchlostiach FSO systému náročné [43, 56].



Obr. 4 Princíp fungovania mäkkého prepínania v hybridnom FSO/RF systéme.

C. Metóda tvrdého prepínania v hybridnom FSO/RF systéme pomocou strojového učenia

Strojové učenie je dôležitou súčasťou rastúcej oblasti dátovej vedy. Použitím štatistických metód sú algoritmy vyškolené na vytváranie klasifikácií alebo predpovedí. Pretože hlboké učenie a strojové učenie sa často zamieňajú, za zmienku však stoja menšie rozdiely medzi nimi. Strojové učenie, hlboké vzdelávanie a neurónové siete sú podoblasti umelej inteligencie. Hlboké vzdelávanie je však v skutočnosti podoblasťou strojového učenia a neurónové siete sú podoblasťou hlbokého učenia. Hlboké a strojové učenie sa

líši v tom, ako sa jednotlivé algoritmy učia. Hlboké učenie automatizuje veľkú časť procesu extrakcie funkcií, eliminuje niektoré ručné zásahy človeka a umožňuje použitie väčších množín údajov. Hlboké učenie si môžete predstaviť ako „škálovateľné strojové učenie“ [64].

Klasické alebo „nehlbkové“ strojové učenie je pri učení viac závislé od ľudského zásahu. Humánni experti určujú súbor funkcií na pochopenie rozdielov medzi dátovými vstupmi, ktoré si spravidla vyžadujú učenie sa štruktúrovanejších dát [64].

„Hlboké“ strojové učenie môže na označenie svojho algoritmu využiť označené súbory údajov, známe tiež ako učenie pod dohľadom, ale nevyhnutne nevyžaduje označený súbor údajov. Môže prijímať neštruktúrované údaje v surovej forme (napr. Text, obrázky) a môže automaticky určovať súbor funkcií, ktoré navzájom odlišujú rôzne kategórie údajov. Na rozdiel od strojového učenia nevyžaduje spracovanie údajov ľudským zásahom, čo nám umožňuje škálovať strojové učenie zaujímavejšími spôsobmi [64].

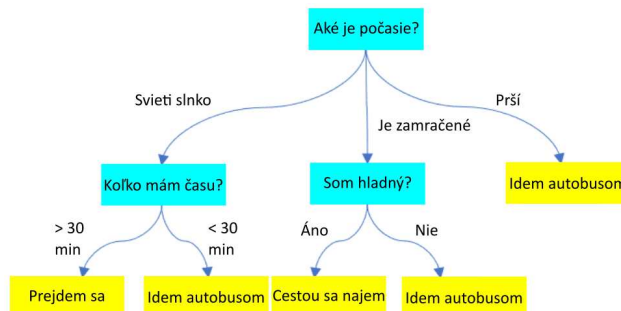
Základný koncept strojového učenia v oblasti vedy o údajoch zahŕňa použitie metód štatistického učenia a optimalizácie, ktoré umožňujú počítačom analyzovať množiny údajov a identifikovať vzory. Techniky strojového učenia využívajú dolovanie údajov na identifikáciu historických trendov a informovanie budúcich modelov [64].

- **Rozhodovací proces:** Na predikciu alebo klasifikáciu sa vo všeobecnosti používajú algoritmy strojového učenia. Na základe niektorých vstupných údajov, ktoré môžu byť označené alebo neoznačené, algoritmus vytvorí odhad vzoru v údajoch [64].
- **Chybová funkcia:** Chybová funkcia slúži na vyhodnotenie predikcie modelu. Ak sú známe príklady, chybová funkcia môže vykonať porovnanie na posúdenie presnosti modelu [64].
- **Proces optimalizácie modelu:** Ak model lepšie vyhovuje údajovým bodom v tréningovom súbore, váhy sa upravujú tak, aby sa znížil nesúlad medzi známym príkladom a odhadom modelu. Algoritmus zopakuje tento proces vyhodnotenia a optimalizácie a automaticky aktualizuje váhy, kým nie je dosiahnutý prah presnosti [64].

1) Rozhodovacie stromy

Medzi najjednoduchšie a najpoužívanejšie algoritmy strojového učenia patrí práve metóda rozhodovacích stromov. Tie sa využívajú v riešení klasifikácie a regresie. Rozhodovací strom predstavuje štruktúru stromu, kde každý uzol prezentuje testovanie danej určitej vlastnosti vstupných dát. Prvotný uzol predstavuje koreňový uzol, kde listy stromu ďalej reprezentujú výsledné hodnoty danej triedy alebo skupiny vstupných dát. Najčastejšie sa využívajú pri klasifikácii dát. Pri vytváraní stromu sú vstupné dáta pomaly testované v každom uzle, kým sa nedostanú do daného listu, kde sa vstupný bod dát nachádza. Medzi najčastejšie výhody rozhodovacích stromov patria jednoduchosť a interpretovateľnosť. Dokážu pracovať ako s kategorickými tak aj so spojivými dátami. Ďalšou výhodou je možnosť identifikovať a následne odstrániť

nepotrebné vlastnosti, ktoré na výstup nemajú zásadný vplyv. Hlavnou nevýhodou je možnosť pretrénovania, kedy stromy nedokážu správne predikovať výsledné dáta na základe prílišného prispôsobenia sa modelu k tréningovým dátam. Príklad rozhodovacieho stromu je zobrazený na Obr. 5. V tomto prípade ide o cestovanie do práce [65, 66].



Obr. 5 Príklad rozhodovacieho stromu.

2) Indikátor intenzity prijatého signálu

Indikátor intenzity prijatého signálu (z angl. Received Signal Strength Indicator - RSSI) je považovaný za jeden z najdôležitejších ukazovateľov FSO prepojenia z pohľadu dostupnosti. V rámci tvrdého prepínania hybridného FSO/RF systému je RSSI zároveň hlavným parametrom. Jednotky pre hodnoty RSSI sú určené na dBm (decibelmili) a typicky sú zobrazené v záporných hodnotách v rozsahu od 0 dBm do -110 dBm pričom 0 dBm je zobrazenie vynikajúceho signálu a -110 dBm predstavuje hodnotu extrémne zlého signálu [68].

RSSI je teda meranie toho, ako kvalitne dokáže zariadenie detegovať a prijímať signál z prístupového bodu alebo konkrétneho smerovača. Na hodnote RSSI je dobré to, že dokáže určiť, či je prenášaný signál dostatočný na vytvorenie bezdrôtového spojenia. Ak sa vysielajú na väčšiu vzdialenosť, dôjde k zoslabeniu signálu a samotná šírka pásma bezdrôtového spojenia sa spomalí. Tento úbytok vedie k slabšiemu výstupu údajov. Pre bežného používateľa je tento parameter neviditeľný, ale keďže sa intenzita prenášaného signálu často líši a ovplyvňuje kvalitu prepojenia, niektoré zariadenia tento parameter sprístupňujú používateľovi, aby bolo možné zistiť, kde a kedy nastal pokles prenášaného signálu [43].

III. DOSIAHNUTÉ VÝSLEDKY

Jedným z hlavných prínosov je návrh správneho algoritmu pre predikciu parametra RSSI pre potreby tvrdého prepínania v hybridnom FSO/RF systéme. Na to, aby bolo možné tento parameter správne predikovať, je potrebné brať do úvahy aj správanie sa prenosového prostredia, teda prenosového kanála, na čo nám slúži monitorovacia stanica poveternostných podmienok. Následne po zaznamenaní dát je možné navrhnuť algoritmus strojového učenia, na základe ktorého bude môcť hybridný FSO/RF systém správne predikovať hodnotu RSSI a tak vopred určiť, ktorá z liniek tohto systému bude v tom čase vysielateľ.

A. Monitorovanie poveternostných podmienok pre potreby predikcie parametra RSSI

Monitorovanie poveternostných podmienok na prenosovej trase medzi oboma stranami hybridného FSO/RF systému je dôležité z hľadiska určovania spoľahlivosti a dostupnosti tohto systému. Keďže FSO systém potrebuje na svoje fungovanie priamu viditeľnosť, práve monitorovanie atmosférického kanála pomáha pri určovaní parametra RSSI, ktorý je dôležitý pri metódach tvrdého či mäkkého prepínania. Prahová hodnota, kedy FSO systém ešte dokáže komunikovať je -42 dBm.

Monitorovacia stanica je zostavená z viacerých senzorov, ako sú senzor teploty, vlhkosti, tlaku, koncentrácie častíc, viditeľnosti a anemometer na meranie rýchlosti vetra. Sensory sú umiestnené na mikropočítači Raspberry Pi 3 B. Ten slúži iba na zber dát, ktoré sú následne odosielané cloudové úložisko ThingSpeak. Ten ponúka možnosť vytvorenia vlastných aplikácií pre IOT a tak isto umožňuje virtualizáciu dát.

B. Metóda rozhodovacích stromov

Výsledkom využitia rozhodovacieho stromu je vzdelávací model, ktorý je možné použiť na predpovedanie hodnoty alebo triedy výslednej premennej pomocou spôsobu učenia sa jednotlivých pravidiel rozhodovania vyvedených zo školiacich údajov. Tieto pravidlá vo všeobecnosti predstavujú tréovanie na konkrétnom súbore dát, ktoré systém už pozná. Je potrebné uviesť, že na tréovanie tohto algoritmu je potrebná množina tréovacích dát. Tá je vytvorená z viacerých prípadov $P = \{P_1, P_2, \dots, P_n\}$, pričom každý z prípadov je vopred definovaný k-ticou testovaných vlastností, teda výstupom premenných $X = \{X_1, X_2, \dots, X_n\}$, ktoré môžu byť vyjadrené buď číslom alebo slovom. Rozhodovací strom je stromový štruktúrovaný klasifikátor, kde interné uzly predstavujú vlastnosti súboru údajov, vetvy predstavujú rozhodovacie pravidlá a každý listový uzol predstavuje výsledok. V rozhodovacom strome sú dva uzly, ktorými sú rozhodovací uzol a listový uzol. Rozhodovacie uzly sa používajú na akékoľvek rozhodnutie a majú viac vetiev, zatiaľ čo listové uzly sú výstupom týchto rozhodnutí a neobsahujú žiadne ďalšie vetvy. Rozhodnutia alebo test sa vykonávajú na základe vlastností daného súboru údajov [65, 69]. Pri vytváraní algoritmu boli použité dve metódy, a to metóda rozhodovacích stromov a metóda rozhodovacích stromov s využitím AdaBoost regresora. Obe metódy boli tréované a testované v rôznych pomeroch, a to 70:30, 80:20 a 90:10. Tento pomer určuje rozloženie tréovacích a testovacích dát. Dôležité bolo určiť, koľko vstupných dát potrebujeme na dosiahnutie správnych výsledkov. Pri vytváraní rozhodovacieho stromu sa množina vstupných tréovacích dát rozdelí na podmnožiny, ktoré pri ďalších krokoch charakterizujú vstupné premenné. V procese delenia sa rekurzívne využíva dovtedy, kým nie je splnená podmienka ukončenia. Pri rozbere atmosférického kanála pre hybridný FSO/RF systém bol rozsiahlo navrhnutý systém vstupných prípadov (premenných) X spolu s prislúchajúcimi premennými y , ktoré sú na výstupe. Pre hybridné FSO/RF systémy je výstupná premenná y hodnota

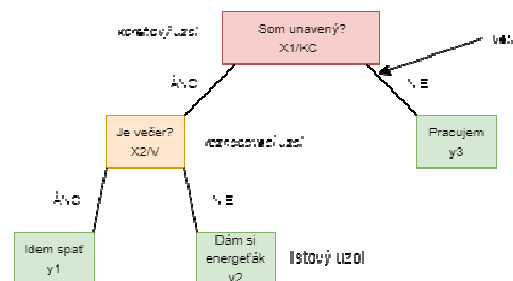
RSSI, na základe ktorej dochádza k prepínaniu liniek. Tréovacia matica vstupných premenných je zobrazená nižšie.

$$X = \begin{bmatrix} TL_1 & TE_1 & KC_1 & V_1 & \dots & RV_1 \\ TL_2 & TE_2 & KC_2 & V_2 & \dots & RV_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ TL_n & TE_n & KC_n & V_n & \dots & RV_n \end{bmatrix} \quad (1)$$

Jednotlivé stĺpce predstavujú namerané hodnoty pomocou monitorovacej stanice poveternostných podmienok. TL je barometrický tlak [hPa], VV predstavuje vlhkosť prostredia [%], TE reprezentuje teplotu ovzdušia [°C], RV rýchlosť vetra [m/s], KC popisuje koncentráciu častíc [mg/m³] a V je viditeľnosť [m]. Matica výstupnej premennej y (target) reprezentuje prijatý optický výkon.

$$y = [y_{RSSI.1} \ y_{RSSI.2} \ \dots \ y_{RSSI.n}]^T \quad (2)$$

Obr. 6 zobrazuje príklad generovania algoritmu rozhodovacieho stromu. Ak to aplikujeme na riešenie problematiky prepínania hybridného FSO/RF systému, tak ten pozostáva z množiny prípadov na určenie stavu, či je človek schopný pracovať alebo nie. V tomto prípade koreňový uzol $X1$ predstavuje stav únavy a rozhodovací uzol $X2$ dobu počas dňa. Pre FSO/RF systém teda ide o $X1$ – koncentráciu častíc (KC) a $X2$ – viditeľnosť (V).



Obr. 6 Príklad fungovania rozhodovacieho stromu.

Pre výpočet y_d by rovnica vyzerala nasledovne.

$$y_d = [y_{idem spať} \ y_{dám si energetik} \ y_{pracujem}]^T \quad (3)$$

1) Výpočet korelačného koeficientu

Keďže oba algoritmy pracujú s veľkou množinou vstupných údajov, je potrebné vziať do úvahy vplyvy jednotlivých parametrov prenosového kanála na predikciu parametra RSSI. V dôsledku toho je potrebné zistiť korelačný koeficient, ktorý sa využíva práve na zistenie vzájomného vplyvu (vzťahu) medzi jednotlivými parametrami a parametrom RSSI. Výsledky korelácie sú v rozmedzí od -1 do +1, pričom:

- -1 – predstavuje záporný vplyv medzi vybranými parametrami,
- +1 – predstavuje kladný vplyv medzi vybranými parametrami,

- 0 – predstavuje hodnotu, kedy nie je možné nájsť medzi vybranými parametrami žiadny vplyv.

Medzi najčastejšie používané metódy určenia korelácie sa využíva Pearsonov korelačný koeficient, ktorý funguje na princípe kovariancie. Znamená to, že tento korelačný koeficient vo výsledku vracia informáciu a korelácii medzi skúmané premenné. Boli ale otestované viaceré korelačné metódy na výpočet korelačného koeficientu, a to:

- Pearsonov korelačný koeficient,
- Kendall tau-b korelačný koeficient,
- Spearmanov korelačný koeficient.

TABULKA 1 Výsledky výpočtov korelačných koeficientov.

Hodnota Y	Spearmanov korelačný koeficient	Kondall tau korelačný koeficient	Pearsonov korelačný koeficient
RSSI	1,000	1,000	1,000
Teplota 1	-0,276	-0,234	0,145
Tlak	0,131	0,091	0,053
Teplota 2	-0,286	-0,242	0,130
Vlhkosť	-0,130	-0,106	-0,007
Teplota 3	-0,271	-0,230	0,167
Rýchlosť vetra	0,158	0,120	-0,010
Sila vetra	0,158	0,120	-0,010
Koncentrácia častíc	-0,041	-0,039	0,177
Viditeľnosť	-0,081	-0,072	0,264

Výsledky výpočtov korelačných koeficientov sú zobrazené v TABULKA 1. Ak sa pozrieme na výsledky, je zrejmé, že najlepšie výsledné hodnoty dosahuje práve Pearsonov

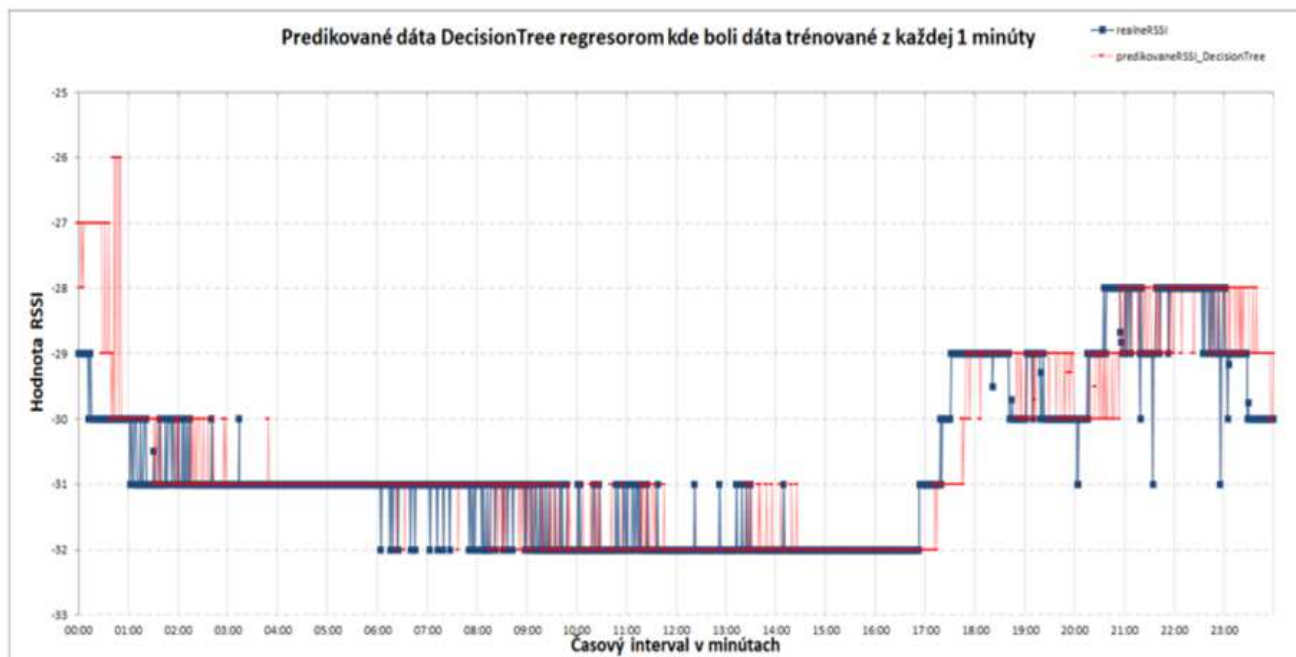
korelačný koeficient, ktorým sa určil najväčší vplyv vstupných parametrov na parameter RSSI. Najväčší vplyv majú: teplota1, teplota2, teplota3, koncentrácia častíc a viditeľnosť.

2) *Analýza výsledkov predikcie pomocou metódy Decision Tree a Decision Tree s AdaBoost regresorom*

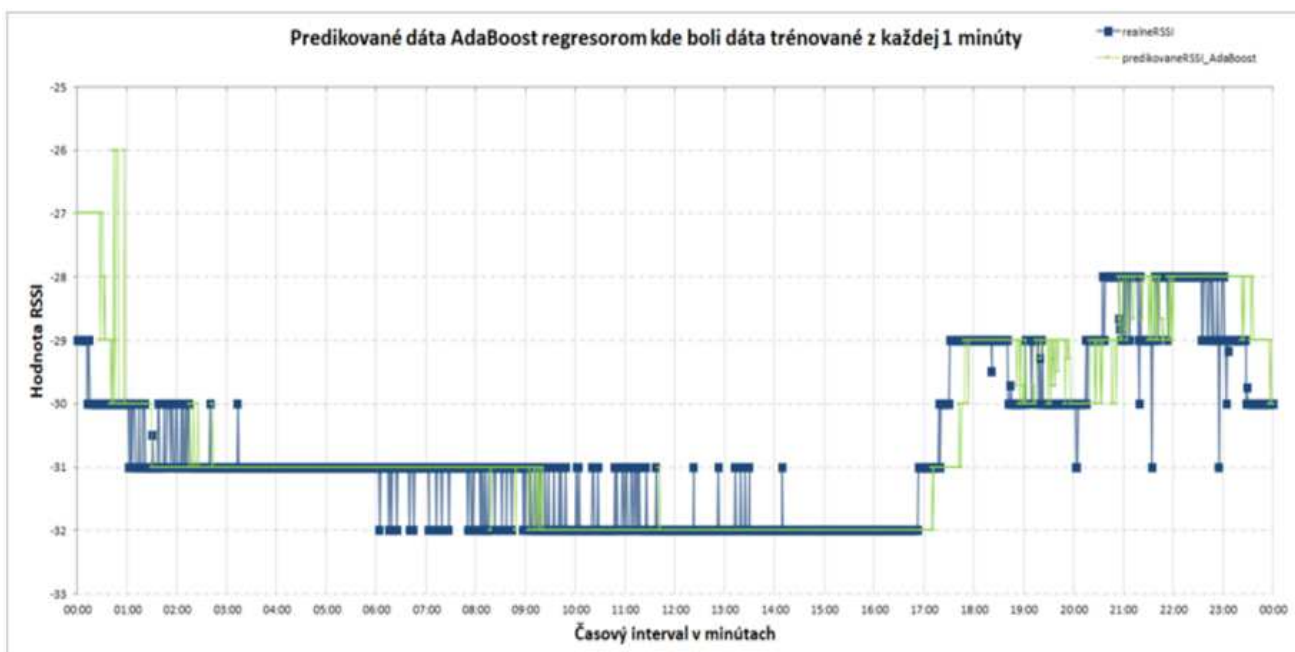
Algoritmus bol navrhnutý na rozloženie daného problému na menšie časti, aby bolo jednoduchšie dosiahnuť požadované výsledky. Dôležitou súčasťou je vstupná množina dát. Tá bola získavaná pomocou monitorovacej stanice poveternostných podmienok zo servera ThingSpeak. Aktuálne obsahuje okolo 120 000 záznamov. V programe je vytvorený aj ďalší súbor .csv, ktorý je prázdny a slúži ako pomocný súbor s dátumom a časom predikcie. V tomto návrhu bol program situovaný na predikciu parametra RSSI o minútu dopredu. Keďže nevieme vopred určiť namerané hodnoty z meteorologickej stanice, tento program pracoval so spriemerovaním 50 posledných nameraných hodnôt. Testovalo sa viacero variant spriemerovania, ako bol súbor posledných 40, 30 a 20 hodnôt, ale najuspokojivejšie výsledky dosahovala práve hodnota 50 údajov. Keďže ide o premenlivé dáta, nie v každom prípade bola predikcia 100%, čo je možné vidieť v tabuľke nižšie. Samotný algoritmus bol trébovaný na viacerých množinách vstupných dát.

TABULKA 2 Úspešnosť predikcie oboch použitých metód strojového učenia v pomere trébovacích a testovacích dát 70:30.

Výber dát na trébovanie z každej:	Úspešnosť predikcie metódou Decision Tree	Úspešnosť predikcie metódou Decision Tree s AdaBoost Regresorom
1. minúty	77%	87%
2. minúty	68%	83%
5. minúty	47%	72%
10. minúty	27%	59%



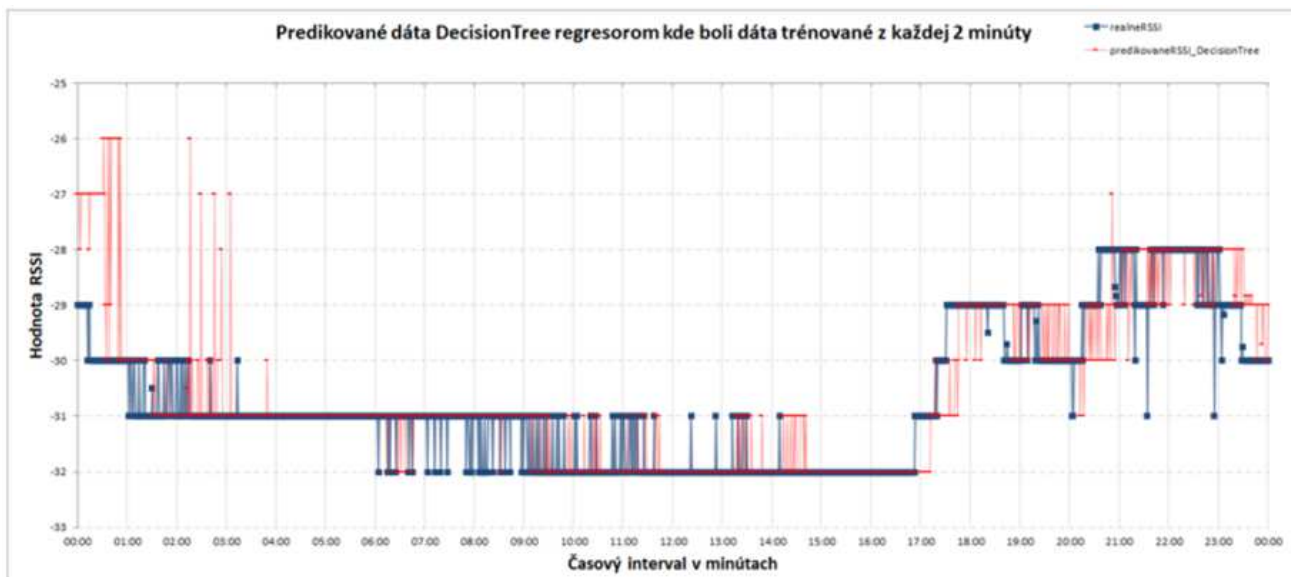
Obr. 7 Porovnanie reálnych a predikovaných hodnôt RSSI.



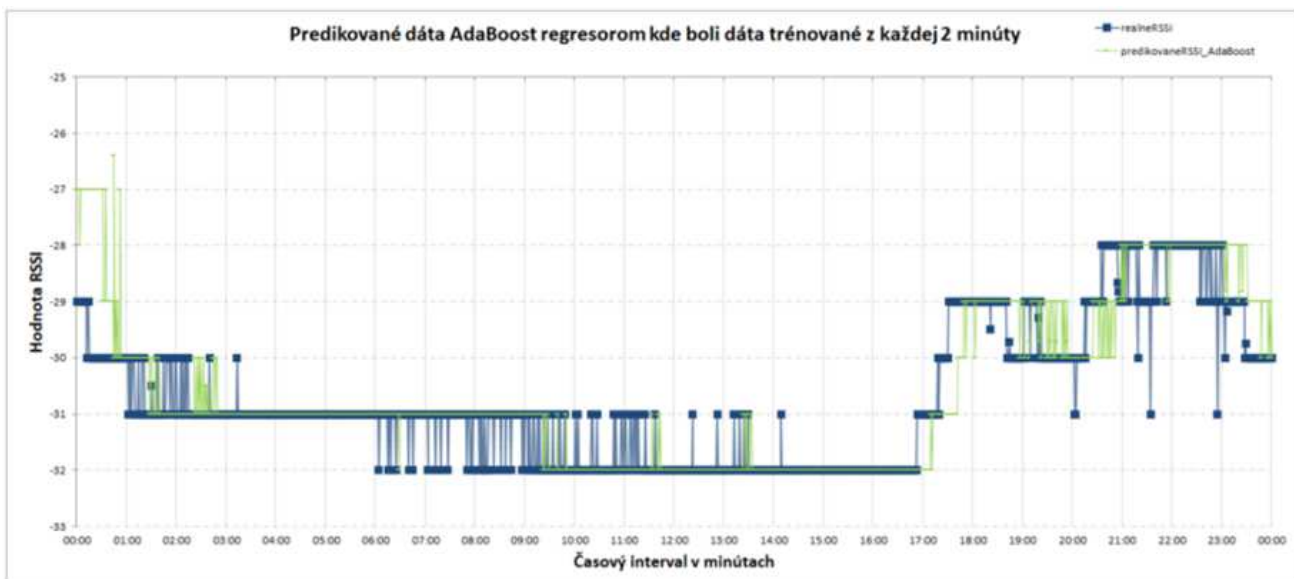
Obr. 8 Porovnanie reálnych a predikovaných hodnôt RSSI.

Ak sa pozrieme na metódu Decision Tree s AdaBoost regressorom, tento algoritmus dosiahol lepšiu úspešnosť predikcie. Ak si porovnáme grafy, na výkyvy predikovaných hodnôt majú vplyv hlavne premenlivé parametre. Z grafov je taktiež možné vidieť, že v krátkom časovom intervale mali grafy medzi sebou priveľké odchýlky, no ak sa na to pozrieme, je vidieť, že predikcia parametra RSSI bola

s takmer úplnou presnosťou. Ďalej sa skúmali aj ďalšie možnosti, ako bola zmena vstupných dát či už výberom parametrov z každej 2., 5. a 10. minúty a zároveň aj parametre s najväčším korelačným vplyvom. Ide o množinu dát obsahujúcu teplotu z 3 senzorov, koncentráciu častíc a viditeľnosť.



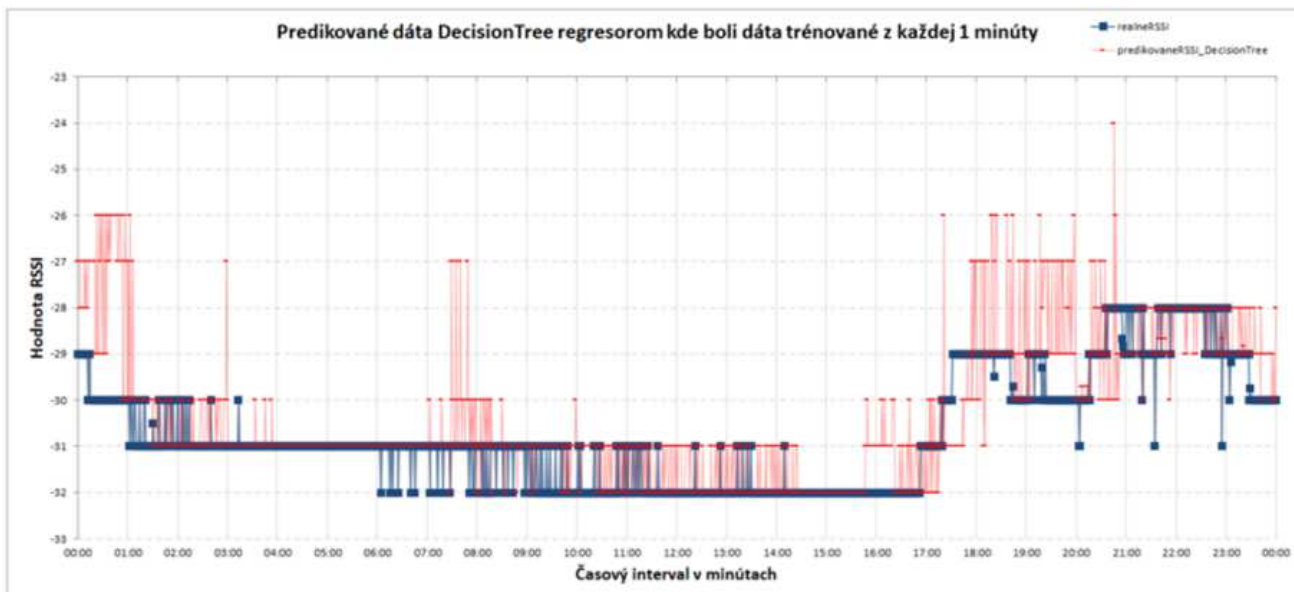
Obr. 9 Porovnanie reálnych a predikovaných hodnôt RSSI.



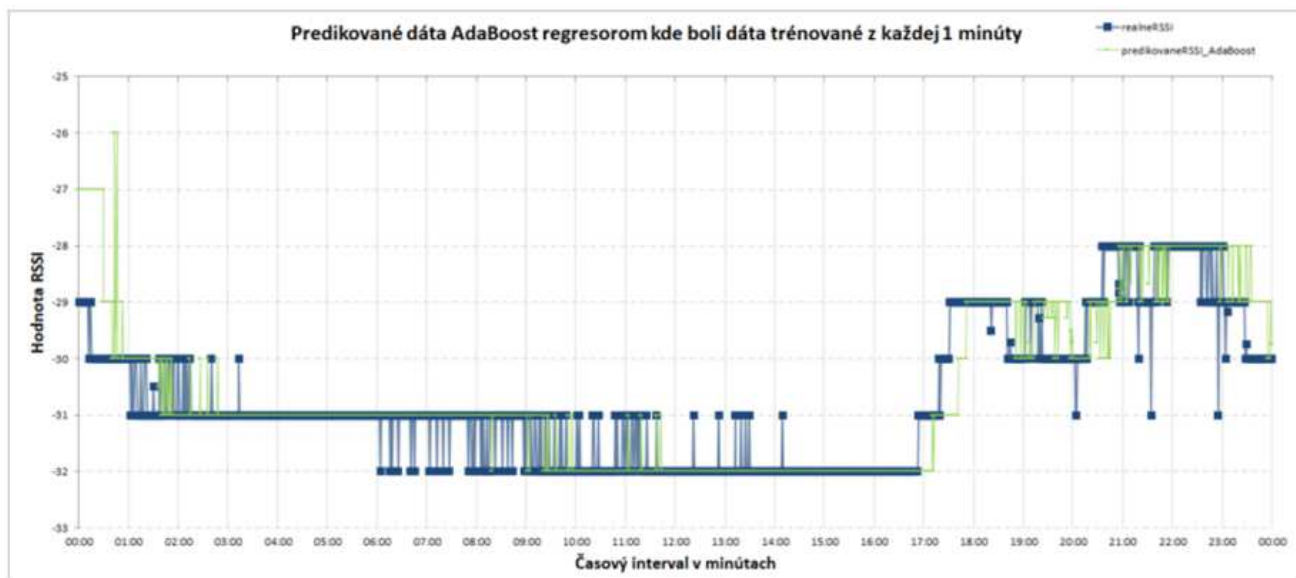
Obr. 10 Porovnanie reálnych a predikovaných hodnôt RSSI.

Z týchto výsledkov je zrejmé, že pri dátach z každej 2. minúty je algoritmus ešte schopný predikovať s dobrou úspešnosťou. V ďalšom prípade nastala zmena vo vstupných dátach, avšak nie len z časového horizontu ale tak isto aj čo sa týka výberu vstupných parametrov. Z korelácií vyplynulo, že

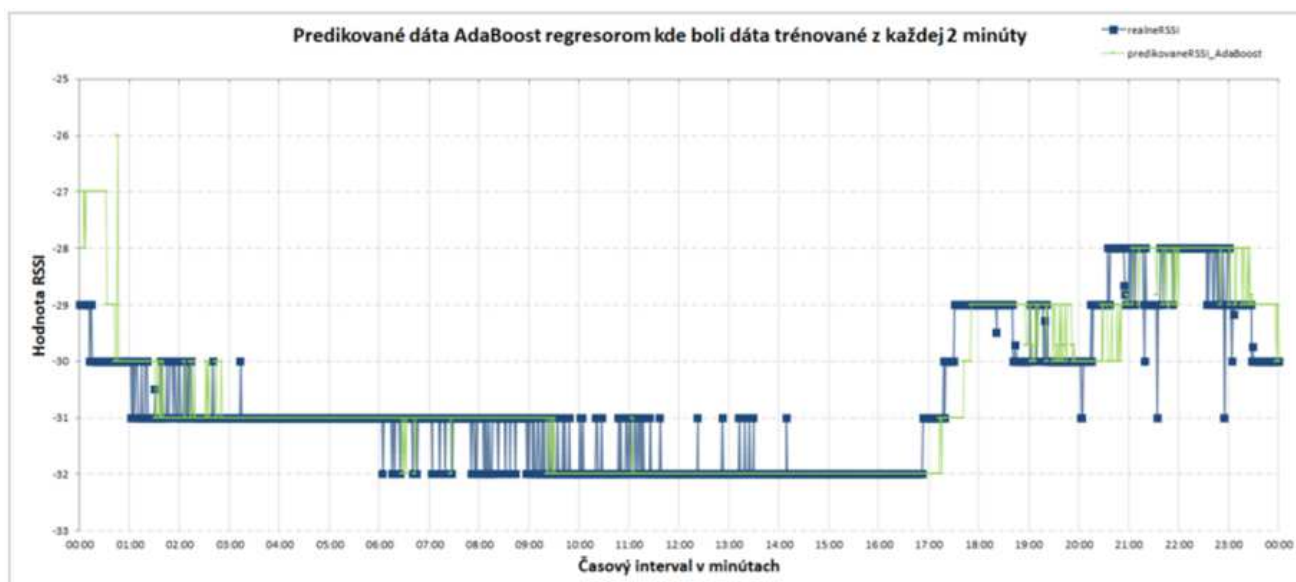
najväčší vplyv na parameter RSSI má teplota, viditeľnosť a koncentrácia častíc. Práve tento prípad je zobrazený v nasledujúcich výstupných grafoch. Opäť sú porovnané obe metódy strojového učenia bez použitia a s použitím AdaBoost regresora.



Obr. 11 Porovnanie reálnych a predikovaných hodnôt RSSI.



Obr. 12 Porovnanie reálnych a predikovaných hodnôt RSSI.



Obr. 13 Porovnanie reálnych a predikovaných hodnôt RSSI.

Z výsledných grafov je zrejmé, že algoritmus Decision Tree nedokáže veľmi presne predpovedať predikciu parametra RSSI. Naopak, lepšie sa javí algoritmus Decision Tree s AdaBoost regresorom, ktorý dokáže presnejšie predpovedať parameter RSSI aj s menším objemom dát.

C. Návrh záložného RF systému pre potreby hybridného FSO/RF systému

Neoddeliteľnou súčasťou hybridného FSO/RF systému je záložná RF linka, ktorá plní funkciu sekundárnej linky v prípade poklesu prijatého výkonu na strane FSO hlavice pod kritickú hodnotu. Tá je stanovená na hodnote -42 dBm. Ak teda hodnota parametra RSSI klesne pod prahovú hodnotu, systém sa na základe tvrdého prepínania prepne do módu

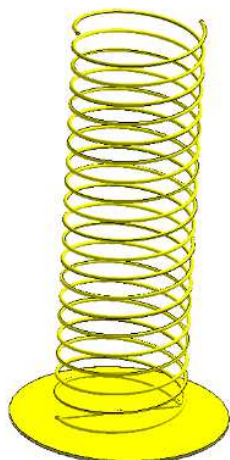
odosielania/prijímania dát pomocou záložnej linky, v tomto prípade RF kanála. V dôsledku toho je potrebné navrhnuť a správne vyladiť sekundárnu RF linku.

1) Návrh a simulácia špirálovej antény ako sekundárnej linky v hybridnom FSO/RF systéme

Navrhnutá špirálová anténa obsahuje dve špirály, ktoré sú v opačnej fáze, teda je bifilar, ako je zobrazené na Obr. 14. Znamená to, že doska plošných spojov, ktorá je umiestnená pod špirálami, slúži ako reflektor a súčasne sa tam vytvára fázový posun.

Rozmery špirálovej antény sú popísané v TABUĽKA 3. V tabuľke sa nachádzajú rozmery zemnacej plochy, ktoré popisujú rozmery medenej vrstvy v tejto doske plošných spojov. Ako reflektor teda slúži samotná medená vrstva.

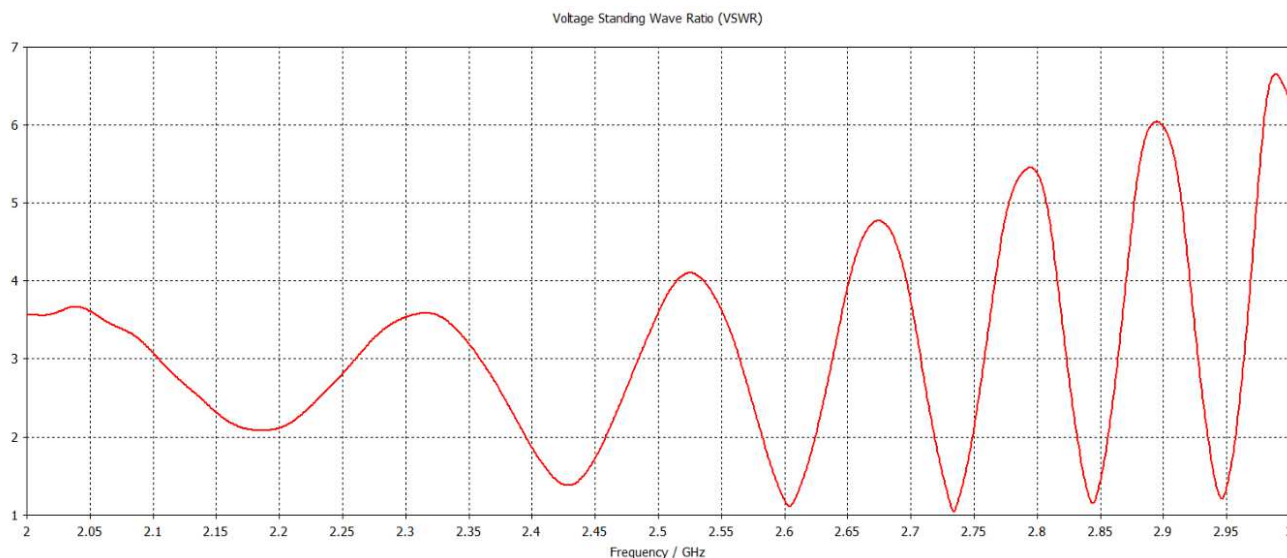
Špirály sú vytvorené pomocou medeného vodiča s hrúbkou 1,3 mm. Jadro tejto dosky je tvorené materiálom FR-04.



Obr. 14 3D zobrazenie špirálovej antény s pracovnou frekvenciou 2,4 GHz.

TABUĽKA 3 Parametre simulovanej špirálovej antény s pracovnou frekvenciou 2,4 GHz.

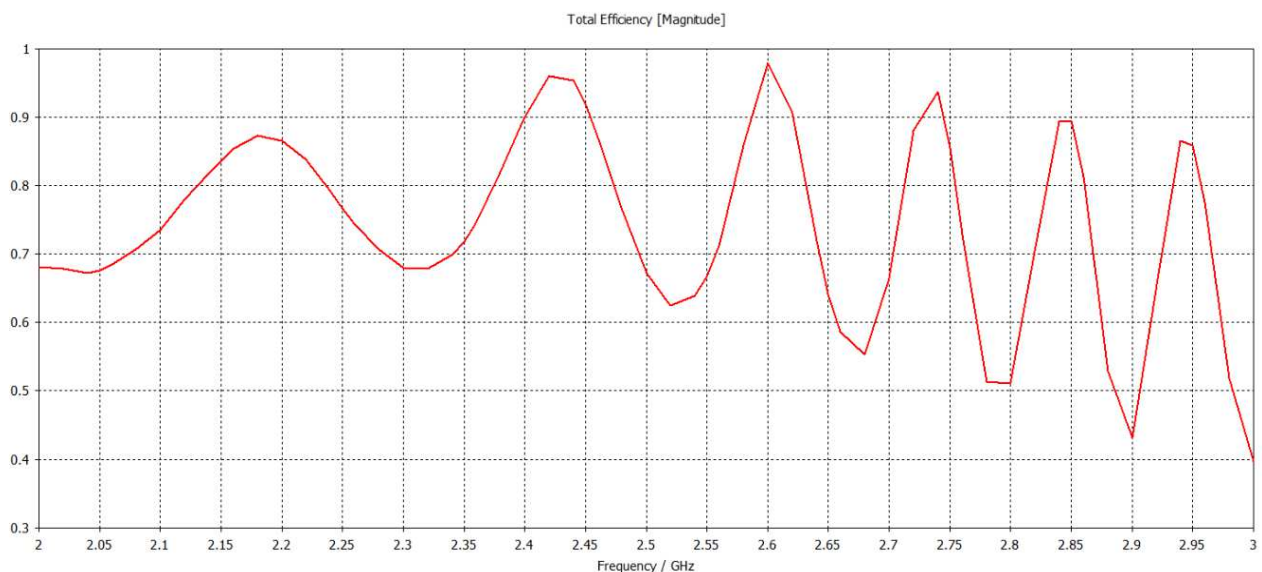
Parameter	Popis
Typ antény	Špirálová anténa
Pracovná frekvencia	2,4 GHz
Dĺžka špirál	144 mm
Priemer špirál	44 mm
Počet závitov	10
Smer závitov	Pravotočivá závitnica
Rozmery PCB	75 x 1 mm (priemer x hrúbka)
Rozmery zemniacej plochy	75 x 0,035 mm (priemer x hrúbka)
Použitie médium	Meď, FR-04
Typ napájania	SMA konektor pre pripojenie koaxiálnym vodičom



Obr. 15 Hodnoty VSWR špirálovej antény s pracovnou frekvenciou 2,4 GHz zobrazené v lineárnom grafe.

Táto bifilar špirálová anténa s pracovnou frekvenciou 2,4 GHz má dobré impedančné prispôsobenie. Je však prispôsobená na veľmi úzke pásmo, pretože hodnoty VSWR sú v rozmedzí 2,39 – 2,46 GHz pod úrovňou 2, čo predstavuje nelicencované pásmo, kde bude táto anténa vysielať. V tomto nelicencovanom pásme sa minimum nachádza na úrovni frekvencie 2,427 GHz, kde hodnota VSWR je na úrovni 1,38.

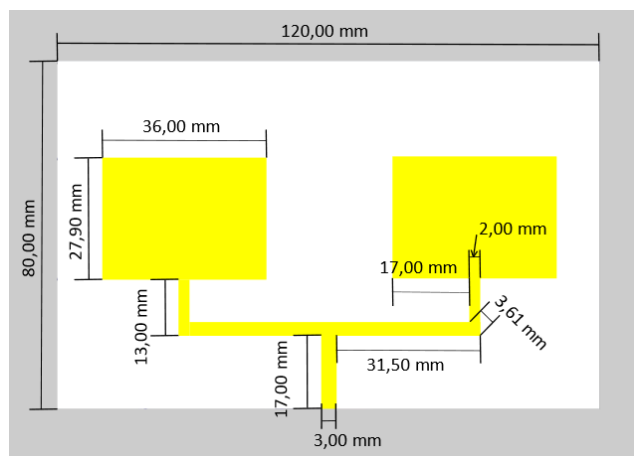
Z grafu je možné vidieť, že prepadliny sú na viacerých frekvenciách, ktoré sú pre túto anténu rezonančné. Na Obr. 16 vidíme účinnosť simulovanej špirálovej antény, ktorá je v rozmedzí 2,4 – 2,45 GHz veľmi vysoká. Hodnota účinnosti predstavuje viac ako 90%. Účinnosť antény koreluje s grafom na Obr. 15, kde hodnoty VSWR pre dané frekvencie sú pod úrovňou 2, čo značí vysokú úroveň účinnosti.



Obr. 16 Účinnosť špirálovej antény s pracovnou frekvenciou 2,4 GHz.

2) Návrh mikropásikového anténneho poľa s pracovnou frekvenciou 2,4 GHz

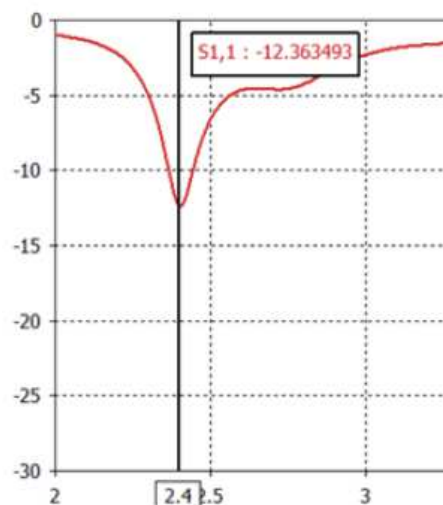
Navrhané pole mikropásikových antén pracuje vo frekvenčnom nelicencovanom pásme 2,4GHz. Ako substrát bol zvolený materiál FR-04, ktorý patrí k najpoužívanejším materiálom vo výrobe plošných spojov. Materiál FR-04 má vysokú dielektrickú pevnosť a dosahuje nízke straty. Práve v dôsledku toho je vhodný pre substrát v mikropásikových anténach. Ako materiál zemniacej plochy je v tomto prípade zvolená meď a to v dôsledku vodivosti a uzemneniu antény. Samotný návrh antény je zobrazený na Obr. 17.



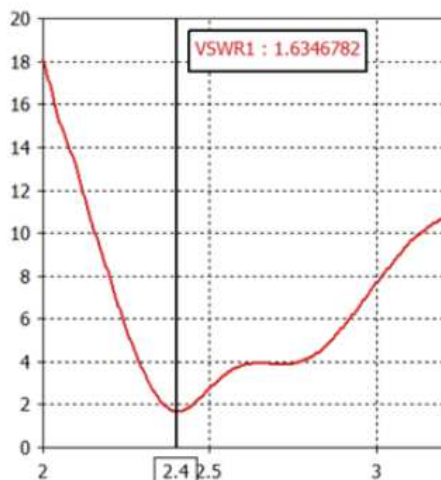
Obr. 17 Návrh mikropásikového poľa antén s pracovnou frekvenciou 2,4 GHz.

Pre tento navrhnutý model dosiahla hodnota S-parametra úroveň -12,36 dB, ako je možné vidieť na Obr. 18. V tomto prípade to pre pole mikropásikových antén znamená, že vstupná energia, ktorá smeruje do antény sa následne odráža späť a to s pomerom 0,175, čo vo VSWR predstavuje hodnotu

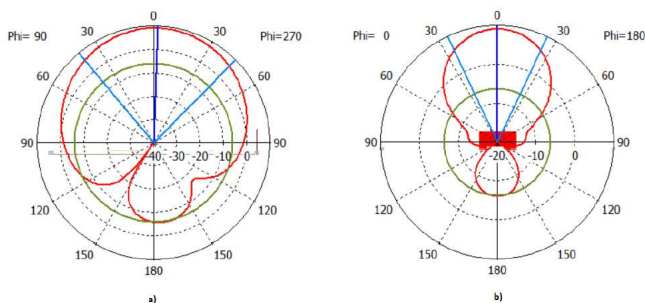
1,63, ako je to zobrazené na Obr. 19. Táto hodnota vyjadruje mieru strát signálu v dôsledku odrazov v anténe. Zároveň predstavuje pomer maximálnej a minimálnej hodnoty napätia na zdroja pripojeného signálu. Pri návrhu mikropásikových antén je dôležité sledovať a dodržiavať nízku hodnotu VSWR kvôli zabezpečeniu optimálneho fungovania a prevádzky prenosového systému. Radiačná účinnosť antény dosiahla hodnotu 0,42, čo predstavuje 42% vstupnej energie, ktorá sa naozaj vyžiarí v podobe elektromagnetických vln.



Obr. 18 S-parameter mikropásikového poľa s pracovnou frekvenciou 2,4 GHz.



Obr. 19 Hodnota VSWR pre mikropásikové pole antén s pracovnou frekvenciou 2,4 GHz.



Obr. 20 Vyžarovaná charakteristika mikropásikového poľa antén s pracovnou frekvenciou 2,4 GHz.

Obr. 20 zobrazuje zmenu vyžarovacej charakteristiky. Na ľavej strane (a) ide o rez v uhle $\varnothing = 90^\circ$ a na ľavej strane (b)) sa jedná o rez v uhle $\varnothing = 0^\circ$. Vo všeobecnosti platí, že rez v 0° sa využíva na určenie charakteristík antény v hlavnej rovine, pričom sa očakáva väčšia intenzita signálu kvôli hlavnej lúčovej energii antény. Ak hovoríme o reze v 90° , tak ten sa naopak využíva na meranie charakteristík kolmo na hlavnú rovinu antény. Tento návrh antény dosiahol hodnotu zisku na úrovni 5,26 dBi a hodnota smerovosti sa dostala na úroveň 9,02 dBi.

IV. ZÁVER

Tento príspevok bol zameraný na hybridné FSO/RF systémy. Kvôli potrebe dosiahnuť takmer 100% dostupnosť je potrebné vhodne zvoliť metódu strojového učenia, ktorá

dokáže správne predikovať parameter RSSI a tak zabezpečiť správny chod systému. Na oba tieto systémy majú vplyv rozličné poveternostné javy. V dôsledku toho je nutné monitorovať poveternostné podmienky systému, ktoré sú následne využité na predikciu parametra RSSI. Ako je možné vidieť vo výsledkoch, systém predikcie je schopný správne predikovať parameter RSSI. Ďalšou dôležitou súčasťou je návrh sekundárnej RF linky pre potreby hybridného FSO/RF systému. Boli vybrané 2 typy antén, z ktorých obe môžu byť využité pre budúcu realizáciu vo výrobe.

POĎAKOVANIE

Tento výskum bol podporovaný z projektu VEGA 1/0260/23 Umelá inteligencia, robustné MANET a viacpreskokové D2D siete integrované do siete 6. generácie.

LITERATÚRA

- [1] H. Khalid, S. S. Muhammad, a spol., "Performance Analysis of Hard-Switching Based Hybrid FSO/RF System over Turbulence Channels," *Computation*, 2019, roč. 7, zv. 2, str. 1 - 10, DOI: 10.3390/computation7020028.
- [2] F. Nadeem, V. Kvicera, a spol., "Weather Effects on Hybrid FSO/RF communication Link," *IEEE journal on selected areas in communications*, 2009, roč. 27, zv. 9, str. 1687 - 1697, DOI: 10.1109/JSAC.2009.091218.
- [3] H. Willebrand, J. Schuster a spol., "Understanding the performance of free-space optics," *Journal of Optical Networking*, 2003, roč. 2, zv. 6, DOI: 10.1364/JON.2.000178.
- [4] M. Lapčák, "Návrh tvrdého prepínania pre hybridný fso/rf prenosový systém na báze predikcie parametra sily prijatého signálu," diplomová práca, Technická Univerzita v Košiciach, 74 s., 2020.
- [5] H. Mainor, S. Arnon, "Performance of an optical wireless communication system as a function of wavelength," *Applied Optics*, 2003, roč. 42, zv. 21, DOI: 10.1364/AO.42.004285.
- [6] M. Usman, Y. Hong-Chuan, M. S. Alouini, "Practical Switching-Based Hybrid FSO/RF Transmission and Its Performance Analysis," *IEEE Photonics Journal*, 2014, roč. 6, zv. 5, str. 1 - 13, DOI: 10.1109/JPHOT.2014.2352629.
- [7] IBM Cloud Education, "What is Machine Learning?," 2020, cit. 20.05.2023, dostupné na: <https://www.ibm.com/cloud/learn/machine-learning>.
- [8] S. Shalew-Shwartz, S. Ben-David, "Understanding Machine Learning: From Theory to Algorithms," Cambridge University Press, 2014, ISBN: 1107057132.
- [9] O. Theobald, *Machine Learning for Absolute Beginners: A Plain English Introduction (Third Edition)*, Scatterplot Press, 2021, ISBN: 1913666522.
- [10] S. M. Navidpour, M. Uysal, M. Kavehrad, "BER performance of free space optical transmission with spatial diversity," *IEEE Transactions on Wireless Communications*, 2007, roč. 6, zv. 8, DOI: 10.1109/TWC.2007.06109.

Riadenie viacportového aktívneho mostíka

Adrián Marcinek

Katedra elektrotechniky a mechatroniky
Fakulta elektrotechniky a informatiky, TUKE
Košice, Slovensko
Email: adrian.marcinek@tuke.sk

Marek Pástor

Katedra elektrotechniky a mechatroniky
Fakulta elektrotechniky a informatiky, TUKE
Košice, Slovensko
Email: marek.pastor@tuke.sk

Abstrakt—Príspevok sa zaoberá matematickým opisom multiport meniča pri neharmonických priebehoch veličín. V úvode príspevku je v skratke opísaný aktuálny stav problematiky a sú opísané topológie vhodné pre prelievanie energie. V ďalších kapitolách je zostavený model meniča a navrhnutý regulátor výstupného napätia. Tento menič je potom simulačne verifikovaný a v závere sú prezentované aj výsledky merania na prototyp multiport meniča

Kľúčové slová — *prelievanie energie, multiport menič, riadenie multiport meniča, planárny transformátor, viacportový menič, MAB menič, riadenie MAB meniča*

Abstract—The paper presents mathematical description of a multiport converter with non-harmonic waveforms. In the beginning of the paper, the current state of the topic is briefly described and topologies suitable for energy exchange are described. In the next chapters, the converter model is assembled and the output voltage regulator is designed. This converter is then verified by simulation, and at the end, the measurement results on the prototype of the multiport converter are also presented.

Key words — *energy exchange, multiport converter, multiport converter control, planar transformer, MAB converter, MAB converter control*

I. ÚVOD

Významnú časť emisií dnešného sveta produkuje priemysel, vykurovanie budov a celosvetový transport, kde majú svoj nezanedbateľný podiel aj nákladné či osobné vozidlá. Existuje niekoľko ciest, ako tieto emisie redukovat'. Jednou z nich je využívanie obnoviteľných zdrojov energie, ako sú palivové články, fotovoltika alebo batériové systémy na ukladanie energie. Hlavnou požiadavkou na všetky tieto zdroje je spoľahlivá premena hnacej energie na energiu elektrickú. Proti tejto premene však ide fakt, že v praxi existuje nespočetné množstvo zariadení s rôznymi napäťovými úrovňami, ktoré bez elektrickej energie nie sú schopné fungovať. Najviac používanou metódou na prepájanie zariadení, resp. zdrojov a spotrebičov s rôznymi napäťovými úrovňami sa preto stali DC/DC alebo DC/AC meniče s jednosmerným medziobvodom, resp. DC zbernicou. Hlavnou nevýhodou tejto topológie je väčší počet konverzných krokov, čo prispieva k menšej efektívnosti. Používaním veľkého počtu DC/DC meničov tak rastú systémové výdavky, zväčšujú sa rozmery systému a klesá výkonová hustota celého systému[1]-[4].

Nevýhody topológií s DC zbernicou je možné vyriešiť použitím multiport meničov. Tieto meniče sú schopné prenášať

energii medzi portami v jednom konverznom kroku. Konštrukčne sú založené na topológiách základných jednosmerných meničov, kde porty multiport meniča sú vzájomne prepojené magnetickou alebo kapacitnou väzbou. V prípade magnetickej väzby je možné hovoriť aj o neizolovaných a izolovaných portoch. Porty v multiport meničoch sú už vo veľkej miere navrhované pre obojsmerný tok energie. Cez obojsmerný port meniča je tak možné dodávať, ale aj odoberať elektrickú energiu. Tieto meniče sa vyznačujú nízkym počtom konverzných krokov, menšími rozmermi celého systému, vysokou výkonovou hustotou a vysokou modularitou meniča. Takýto multiport menič so všetkými jeho výhodami je potom možné použiť na prepájanie obnoviteľných zdrojov energie, ale aj v elektromobilitě na prelievanie energie medzi jednotlivými zdrojmi/spotrebičmi[5] – [11].

V príspevku je predstavený návrh, matematický opis, simulačné overenie riadenia a experimentálne meranie pre nami prezentovanú topológiu viacportového aktívneho mostíka. Príspevok predkladá matematický opis viacportového meniča pracujúceho s neharmonickými veličinami a rozoberá možnosti riadenia a regulácie takéhoto meniča. V závere práce sú porovnané výsledky simulácie z dvoch simulačných modelov a multiport menič je podrobený praktickému meraniu pre overenie teórie riadenia.

Druhá kapitola práce sa teda venuje úvodu do problematiky a matematickému opisu DAB meniča, vytvoreniu jeho modelu a možnostiam riadenia. Tretia kapitola opisuje matematický model MAB meniča, metódy jeho riadenia a výzvy spojené s jeho riadením. Štvrtá kapitola opisuje simulačné výsledky trojitého aktívneho mostíka bez regulátora a s regulátorom. Piata kapitola opisuje merania na prototyp trojportového aktívneho mostíka.

II. TOPOLOGIE VHODNÉ PRE PRELIEVANIE ENERGIE

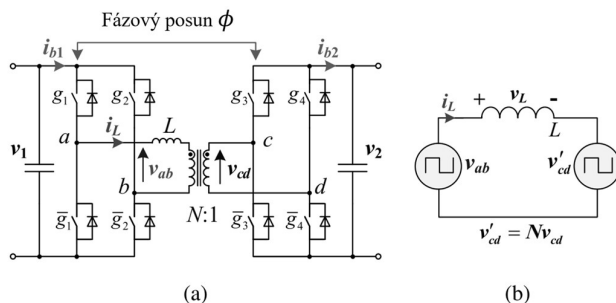
Jedným z hlavných problémov v systémoch s viacerými zdrojmi energie je architektúra napájania, t. j. rozhrania pre pripojenie záťaží, zdrojov a prvkov pre ukladanie energie. Mnoho aplikácií je založených na architektúrach so spoločnou jednosmernou zbernicou, kde sú všetky prvky prepojené cez jednosmerné meniče [12], vrátane prvkov na ukladanie energie, ako sú batérie a superkondenzátory. Hlavnou požiadavkou pre takéto DC/DC meniče je možnosť obojsmerného toku energie a galvanické odizolovanie jednotlivých portov. Preto je vďaka svojím vlastnostiam a účinnosti pre takéto aplikácie vhodný menič s dvojitým aktívnym mostíkom, tzv. „Dual Active

Bridge (DAB)“. DAB menič budeme považovať ako základ pre rozvoj viacportového meniča. Keďže DAB menič umožňuje obojsmerný tok energie medzi dvoma portami, túto vlastnosť využijeme aj pri viacportovom meniči, kde budeme dbať na vysokú modularitu a veľkú výkonovú hustotu. Práve pre túto požiadavku boli skúmané modulárne topológie výkonových meničov [13][14][15], čo umožňuje veľkú flexibilitu celého systému. Napríklad v elektrických vozidlách môžu rôzne jazdné režimy vyžadovať rôznu veľkosť úložných prvkov, pričom takéto modulárne topológie umožňujú optimalizovať prevádzkovú stratégiu. Takýto modulárny systém pre prenos energie má široké možnosti použitia, pretože s rovnakou základnou jednotkou, ktorá tvorí jeden port meniča, je možné pokryť množstvo rôznych požiadaviek na prenos výkonu. Viacportové systémy založené na DC/DC meničoch s názvom „Multi Active Bridge (MAB)“ boli opísané v [16] a [17]. Menič opísaný v [17] je založený na jednom transformátore s viacerými vinutiami, ktorý spája všetky porty. Akonáhle je však systém navrhnutý a postavený, jeho architektúru nie je možné výrazne meniť. Ďalej teda bude analyzovaný MAB menič, ktorý je odvodený od DAB meniča a je založený na prepojení niekoľkých modulov. Každý port je reprezentovaný plným mostíkom, transformátorom s prevodom a sériovo radenou indukčnosťou, ktorá predstavuje rozptylovú indukčnosť transformátora. Elektrická časť topológie MAB meniča je ekvivalentná tej v [17], ale implementácia magnetickej väzby medzi rôznymi mostíkmi je dosiahnutá pomocou fyzicky nezávislých transformátorov, resp. pomocou elektrického prepojenia sekundárných vinutí týchto transformátorov. Vysoká flexibilita je tak garantovaná možnosťou prídania alebo odstránenia polovodičového mostíka a zodpovedajúceho transformátora s jeho rozptylovou indukčnosťou. Analýza MAB meniča v [17] je založená na interakcii každého páru portov. V [18] a [19] je tiež predstavený a modelovaný podobný MAB menič.

V nasledujúcich podkapitolách je najprv uvedený krátky opis DAB meniča, kde sa so zameraním na ustálený stav analyzuje vzťah medzi strednou hodnotou vstupného prúdu každého modulu a fázovými posunmi medzi týmito modulmi. Ďalej je uvedený model DAB meniča a možnosti jeho riadenia.

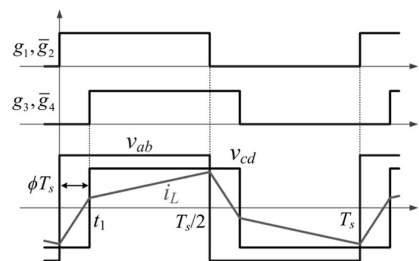
A. Modelovanie DAB meniča

Modelovanie je reprezentácia fyzikálnych javov za pomoci matematických prostriedkov [20]. Modelovanie DAB meniča je v porovnaní s modelovaním konvenčných jednosmerných meničov náročnejšie, pretože jedna zo stavových premenných, prúd cievky i_L , je striedavá a teda jeho stredná hodnota je rovná 0. Topológia DAB meniča a jeho náhradná schéma je na Obr. 1, priebehy spínania tranzistorov a priebeh prúdu i_L potom na Obr. 2.



Obr. 1 a) Topológia DAB meniča a b) jeho náhradná schéma

V nasledujúcich podkapitolách sú opísané niektoré metódy modelovania DAB meniča, ktoré bude neskôr možné aplikovať aj na viacportový menič. Metódy modelovania opisujú DAB menič riadený pomocou fázového posunu, tzv. „single phase shift (SPS)“ modulácie (Obr. 2), ktorú využijeme aj pri opise viacportového meniča. Tieto metódy modelovania je možné ďalej rozšíriť na DAB menič riadený pomocou dvojitého fázového posunu, tzv. „dual phase shift (DPS)“ modulácia alebo trojitým fázovým posunom, tzv. „triple phase shift (TPS)“ modulácia [21].



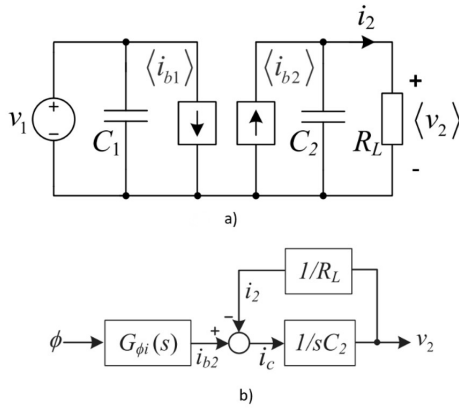
Obr. 2 Priebehy DAB meniča pri SPS modulácii

1) Model s redukovaným rádom

Jedným zo spôsobov, ako DAB menič modelovať, je ignorovať dynamiku prúdu i_L . Takýto model sa nazýva model redukovaného rádu [22]–[24], resp. v zahraničnej literatúre ako „reduced-order model“. V tomto modeli sa na popis prúdu používajú stredné hodnoty vstupných a výstupných prúdov počas jedného spínacieho cyklu (alebo polovice cyklu). Takýmto spôsobom sa DAB menič zjednoduší na systém prvého rádu, čo je znázornené na Obr. 3 a), kde $\langle i_{b1} \rangle$ a $\langle i_{b2} \rangle$ sú stredné hodnoty prúdov i_{b1} a i_{b2} počas spínacieho cyklu.

Výstupný výkon P_0 DAB modulovaného pomocou SPS vyjadríme ako [25]

$$P_0 = \frac{v_1 v_2 N \phi (1 - 2|\phi|)}{f_s L} \quad (1)$$



Obr. 3 a) Zjednodušený model DAB meniča a b) blokový diagram redukovaného modelu

kde N, v_1, v_2 a L sú zobrazené na Obr. 1 a Obr. 2 a f_s je spínacia frekvencia DAB meniča. Keďže platí $P_o = v_2 \langle i_{b2} \rangle$, pričom $\langle i_{b2} \rangle$ je stredná hodnota prúdu, pre $\langle i_{b2} \rangle$ a kladný tok výkonu potom platí:

$$\langle i_{b2} \rangle = \frac{v_1 N \phi (1-2\phi)}{f_s L} \quad (2)$$

Zavedením poruchy $\phi = \Phi + \hat{\phi}$ a $\langle i_{b2} \rangle = I_{b2} + \hat{i}_{b2}$ získame prenosovú funkciu z ϕ na i_{b2} v tvare

$$G_{i\phi}(s) = \frac{\hat{i}_{b2}}{\hat{\phi}} = \frac{d\langle i_{b2} \rangle}{dt} = \frac{V_1 N \phi (1-2\phi)}{f_s L} \quad (3)$$

kde V_1, V_2, ϕ a I_{b2} sú ustálené hodnoty v_1, v_2, ϕ a i_{b2} . Vzťah medzi ϕ a I_{b2} potom môžeme definovať ako

$$\phi = \begin{cases} \frac{1}{4} - \sqrt{\frac{1}{16} - \frac{f_s L I_2}{2 N V_1}} & I_2 \geq 0 \\ -\frac{1}{4} + \sqrt{\frac{1}{16} + \frac{f_s L I_2}{2 N V_1}} & I_2 < 0 \end{cases} \quad (4)$$

$$I_{b2} = I_2 = \frac{V_2}{R_L} \quad (5)$$

Na základe Obr. 2 a) a rovnice (3) je na Obr. 2 b) znázornená bloková schéma riadenia v_2 v závislosti od ϕ . Prenosová funkcia takéhoto riadenia, kde ako vstup uvažujeme ϕ a ako výstup v_2 má potom tvar

$$G_{v\phi}(s) = \frac{\hat{v}_2}{\hat{\phi}} = \frac{V_1 N \phi (1-2\phi)}{f_s L} \frac{R_L}{R_L C_2 s + 1} \quad (6)$$

Na základe vyššie uvedenej analýzy môžeme povedať, že model s redukovaným rádom DAB meniča predstavuje systém prvého rádu. Pri opise meniča sme uvažovali s SPS moduláciou, avšak aj v prípade použitia DPS a TPS modulácie môže byť model DAB meniča reprezentovaný obvodom na Obr. 2.

Hlavný rozdiel medzi týmito modelmi by potom reprezentovalo zosilnenie $G_{\phi i}(s)$, pretože výpočet zosilnenia pre DAB menič s DPS resp. TPS moduláciou je omnoho náročnejší ako výpočet s SPS moduláciou. Model DAB meniča s redukovaným rádom a TPS moduláciou je uvedený v [26]. Okrem toho je možné pri zostavovaní modelu uvažovať aj parazitné parametre obvodu (parazitné odpory, indukčnosti a kapacity). Takýto model je popísaný v [27] a zohľadnenie týchto v niektorých prípadoch bezvýznamných parametrov prináša značnú komplikáciu pri zostavovaní modelu. Tento model použijeme aj v prípade modelovania viacportového meniča, kde bude zo simulačných výsledkov zrejme, že model je dostatočne presný aj bez zohľadnenia parazitných parametrov.

B. Riadenie DAB meniča

Táto časť sa zaoberá dostupnými metódami riadenia a regulácie výstupného napätia, resp. prúdu. V kapitole sú opísané možnosti a implementácia spätnoväzobného regulátora. Na efektívne vyhodnotenie metód regulácie výstupného napätia sú odvodené prenosové funkcie z riadenia na výstup $G_{ro}(s)$ a prenosová funkcia výstupnej impedancie meniča $Z_o(s)$. Metóda je potom implementovaná na viacportový menič a simulačne overená.

1) Spätnoväzobná regulácia výstupného napätia

Spätnoväzobná regulácia pomocou proporcionálno-integrálneho (PI) regulátora je najjednoduchší spôsob regulácie výstupného napätia. Na Obr. 4 je znázornený blokový diagram DAB meniča spolu s regulátorom, ktorý upravuje fázový posun ϕ medzi mostíkmi meniča v závislosti od chyby výstupného napätia vzhľadom na žiadanú hodnotu [28][29]. Prenos PI regulátora je zadefinovaný ako $G_{c1} = (k_p + k_i/s)$, pričom tento typ regulátora zabezpečí nulovú regulačnú odchýlku, a teda minimálnu chybu v ustálenom stave. Ak vezmeme do úvahy aj oneskorenie pri implementácii regulátora do digitálneho signálového procesora, prenosová funkcia uzavretej slučky z riadenia na výstup odvodená na základe Obr. 4 bude mať tvar

$$G_{ro}(s) = \frac{v_2}{V_2^*} = \frac{G_{c1} G_{i\phi} Z_L e^{-1.5T_s s}}{1 + G_{c1} G_{i\phi} Z_L e^{-1.5T_s s}} \quad (7)$$

kde $e^{-1.5T_s s}$ je oneskorenie spôsobené vzorkovaním a vykonaním akčného zásahu. Oneskorenie odberu vzoriek predstavuje čas T_s a oneskorenie riadenia čas $0.5T_s$. Oneskorenie digitálneho riadenia $0.5T_s$ je spôsobené tým, že riadiaca premenná ϕ sa aktualizuje raz za spínací cyklus. Tento proces možno v simulačných nástrojoch modelovať ako tzv. „zero order hold (ZOH)“, ktorého oneskorenie je $0.5T_s$. Zátťaž meniča $Z_L(s)$ je zobrazená na Obr. 4 a jej prenos je vyjadrený ako

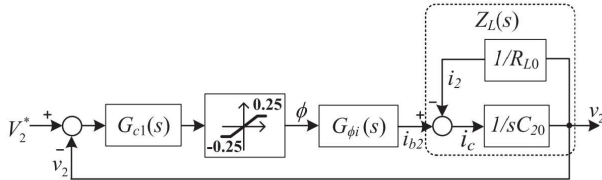
$$Z_L(s) = \frac{R_L}{1 + s C_2 R_L} \quad (8)$$

Z vyššie uvedeného je taktiež možné odvodiť výstupnú impedanciu. Na odvodenie výstupnej impedancie $Z_o(s)$ je

odporová záťaž R_L nahradená zdrojom prúdu i_2 v obvode DAB, výstupná impedancia je potom

$$Z_o(s) = \frac{v_2}{i_2} = \frac{1/(sC_2)}{1 + \frac{G_{v1}G_i\phi e^{-1,5T_s s}}{sC_2}} \quad (9)$$

Uvedený postup návrhu PI regulátora dáva dobrý základ pre jeho použitie pre viacportový menič. PI regulátor je jednoduchý z hľadiska návrhu a implementácie a vo veľkom množstve prípadov dostatočný aj vzhľadom na žiadanú dynamiku.



Obr. 4 Bloková schéma regulácie výstupného napätia

III. MAB MENIČ

Z doposiaľ uvedených informácií je zrejme, že meniče, ktoré sú schopné prelievať energiu medzi dvoma portami sú vo výkonovej elektronike rozšírené a ich modelovanie a spôsoby regulácie sú značne rozvinuté. Počas obdobia doktorandského štúdia zameraného na aktuálny stav problematiky sme pracovali s výrazom „multiport converter“, čo však pri hľadaní relevantných podkladov v zahraničnej literatúre nedávalo uspokojivé výsledky. Správny pojem, ktorý v zahraničnej literatúre popisuje topológiu totožnú s našou je „Multi Active Bridge“ (MAB) a je odvodený od pomenovania dvojportového meniča známeho pod názvom „Dual Active Bridge“. Používanie správnej terminológie rozšírilo zdroje informácií v rámci danej problematiky a prispelo k lepším výstupom tejto práce. Pre nesúlad zahraničnej a slovenskej terminológie si teda zadefinujeme pojem „viacportový aktívny mostík“, resp. „viacportový menič“, skratkou MAB, ktorý budeme používať v nasledujúcich kapitolách a tiež pojmy ako „port“, ktorý predstavuje fyzické rozhranie pre pripojenia meniča a pojem „modul“, ktorý zahŕňa riadiacu a výkonovú časť mostíkového striedača a k nemu pripojený transformátor. Niekoľko takýchto multiaktívnych, resp. multiportových aktívnych mostíkov odvodených od DAB je uvedených v [30]–[35]. Na jednotlivé porty takéhoto meniča sú zväčša pripojené zariadenia ako zdroj energie, záťaž, prípadne zariadenie pre ukladanie energie (batéria, kondenzátor). Pre prenos energie medzi portami je vo vyššie spomenutých príspevkoch zvyčajne použitý viacvinuťový transformátor. Táto topológia je vhodná pre menší počet portov, pretože pri veľkom počte portov sa môže návrh transformátora výrazne skomplikovať. Navyše, ak je už transformátor pre konkrétny typ viacportového aktívneho mostíka s konkrétnymi parametrami navrhnutý, je ťažké pridať meniču ďalšie porty. V takom prípade je potrebné upraviť transformátor a brať do úvahy maximálny povolený prenášaný výkon transformátora.

V tejto kapitole bude opísaný viacportový aktívny mostík s viacerými dvojinuťovými transformátormi, ktorých

sekundárne vinutia sú vzájomne elektricky prepojené, takýto menič je zobrazený na Obr. 5. Tento typ meniča bol tiež analyzovaný a predstavený v [36]. Použitie nezávislých dvojinuťových transformátorov je označované ako tzv. multitransformátorový prístup. Prezentovaná topológia je elektricky ekvivalentná aj iným viacportovým aktívnym mostíkom ako napr. „quad-active-bridge“ (QAB) [37] alebo ine topológie s jedným viacvinuťovým transformátorom [38]. Hlavnou výhodou našej topológie je, že je plne modulárna a škálovateľná. Moduly je možné pripojiť a odpojiť kedykoľvek a počet portov je tak možné prispôsobiť rôznym potrebám prenosu výkonu. Ako príklad môžeme uviesť použitie viacportového aktívneho mostíka v elektromobilitě resp. elektrickej trakkii, kde vzhľadom na rôzne jazdné režimy je možné takýto menič kedykoľvek rozšíriť o ďalšie úložisko energie, pričom môže ísť o dlhodobé úložisko energie tvorené batériou alebo krátkodobé úložisko energie tvorené kondenzátorom. Nie je však potreba výmeny celého meniča ako to je u iných topológií. Z topológie viacportového aktívneho mostíka vyplýva aj výhoda z hľadiska výroby, kedy môžu byť vyrábané moduly s rôznym výkonom a k už existujúcemu viacportovému meniču budú tieto moduly pripojené prostredníctvom sekundárnej strany transformátorov. To prispieva k zníženiu výrobných nákladov, keďže nie je potrebné meniť celý menič, ale iba modifikovať jeho časť. V [39] je popísaný trojportový menič zostavený z viacerých transformátorov, ktorého dva porty sú napájané jednofázovým striedačom a tretí port je napájaný trojfázovým striedačom alebo predstavuje záťažový port. V [40] je zase popísaný viacportový aktívny mostík s nezávislými transformátormi, ktorých sekundárne vinutia sú zapojené do série. Náročnosť riadenia takýchto viacportových aktívnych mostíkov je popísaná v [41]–[43]. Základným princípom riadenia viacportových aktívnych mostíkov je zmena fázového posunu medzi jednotlivými mostíkmi. Je však náročné exaktne definovať výpočet fázových posunov podľa žiadaného rozloženia výkonu, keďže výkon každého portu je závislý na všetkých ostatných fázových posunoch. V kapitole je ďalej opísaná topológia MAB meniča, jeho riadenie a modelovanie. V rámci topológie je vyjadrená rovnica pre napätie medziobvodu v ustálenom stave a stredné hodnoty prúdov jednotlivých portov. V časti zaoberajúcej sa riadením sú analyzované možnosti riadenia takéhoto meniča aj s regulátorom. Pre viacportový aktívny menič je tiež zostavený model redukovaného rádu, ktorý je v závere kapitoly simulácie overený a porovnaný so spínaným modelom pri rôznych podmienkach dodávaného výkonu.

A. Topológia MAB meniča

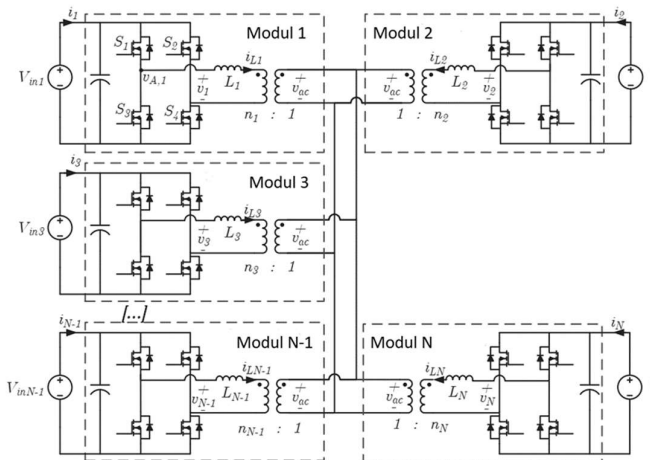
Topológia MAB meniča je zobrazená na Obr. 4 a skladá sa z mostíkového striedača, indukčnosti, ktorá predstavuje rozptylovú indukčnosť transformátora a transformátora, ktorý je pri matematickom popise vyjadrený iba ako prevod medzi primárnou a sekundárnou stranou transformátora. Sekundárne vinutia transformátorov sú paralelne prepojené. Náhradná schéma MAB meniča pre N portov je zobrazená na Obr. 6 b), kde indukčnosť L_k/n_k^2 , ak $k = 1, \dots, N$ predstavuje rozptylovú indukčnosť prepočítanú na sekundárnu stranu transformátora. Sekundárnu stranu transformátora môžeme definovať ako striedavú zbernicu, v zahraničnej literatúre pomenovanú ako „ac bus“ resp. jednoducho nazvať medziobvod a napätie tohto medziobvodu označiť ako v_{ac} . Transformačný pomer

transformátora v k -tom porte môžeme označiť ako n_k a prúd na primárnej strane transformátora, teda prúd indukčnosti, ako i_{Lk} . Výkon prenášaný medzi portami je potom určený vzájomným pôsobením napätia konkrétneho mostíka v_k a napätia medziobvodu v_{ac} cez rozptylovú indukčnosť L_k . Pre získanie matematického vyjadrenia napätia medziobvodu v_{ac} musíme aplikovať tzv. Millmanovu vetu, ktorá napätie medziobvodu definuje ako (10), pričom platí (11). Na základe toho môžeme tvrdiť, že veľkosť a smer toku výkonu závisí na napätí jednotlivých mostíkov v_k , ktoré je určené vstupným napätím každého portu V_{ink} , fázovým posunom medzi jednotlivými mostíkmi, rozptylovou indukčnosťou L_k a transformačným pomerom n_k daného transformátora. Vzhľadom na vyššie uvedenú definíciu môžeme napätie v_{ac} vyjadriť ako

$$v_{ac} = \frac{\sum_{k=1}^N v_k \frac{n_k}{L_k}}{\frac{1}{L_{eq}}} \quad (10)$$

kde

$$L_{eq} = \left(\sum_{k=1}^N \frac{n_k^2}{L_k} \right)^{-1} \quad (11)$$



Obr. 5 Topológia MAB meniča pre N portov

1) Napätie medziobvodu MAB meniča v ustálenom stave

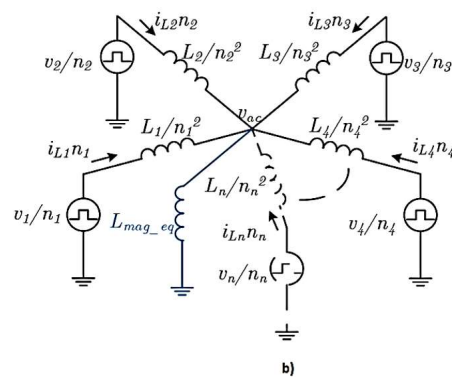
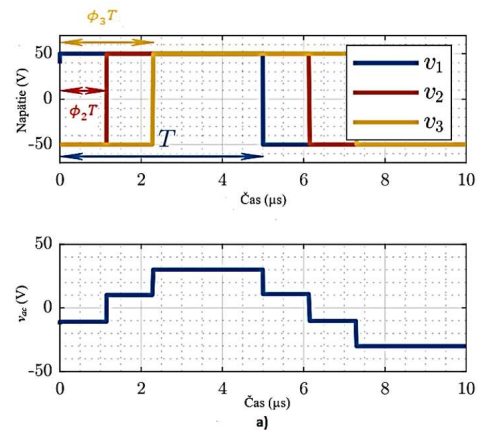
Pre riadenie MAB meniča je možné využiť rôzne typy modulácie podobne ako u DAB meniča. V [44] je popísané riadenie, ktoré pre prenos výkonu využíva fázový posun medzi napätiami mostíkov v_1, v_2, \dots, v_n , čo je zobrazené na Obr. 6 a). Čas posunu každého mostíka je pritom definovaný ako súčin medzi riadiacou veličinou ϕ_k a časovým intervalom T , kde riadiaca veličina ϕ_k je z rozsahu od -1 do 1 a T je reprezentuje polovicu spínacej periódy T_{sw} . Veličina ϕ_k môže byť tiež vyjadrená ako uhol v stupňoch v rozsahu -180° do 180° . Pre MAB menič s troma portami sú napätia mostíkov v_1, v_2 a v_3 a napätie medziobvodu zobrazené na Obr. 6 a). Pri tomto popise je prvý port meniča zvolený ako referenčný a teda fázový posun prvého mostíka ϕ_1 je rovný nule. Analyzovaním vplyvu

magnetizačnej indukčnosti na menič sme náhradnú schému meniča upravili ako na Obr. 6 b). Výpočet napätia medziobvodu ostáva v tomto prípade totožný podľa (5.39), avšak zmení sa výpočet ekvivalentnej indukčnosti L_{eq} , kde $L_{mag,k}$ je magnetizačná indukčnosť k - teho portu vzhľadom na sekundárnu stranu. Rovnica pre L_{eq} potom bude vyzerat'

$$L_{eq} = \left(\sum_{k=1}^N \left(\frac{n_k^2}{L_k} + \frac{1}{L_{mag,k}} \right) \right)^{-1} \quad (12)$$

Ak sú parametre všetkých portov rovnaké, je rovnaký aj ich vplyv na medziobvod z hľadiska sekundárnej strany transformátora. Potom môžeme porty nazvať ekvivalentnými a bude platiť (13)

$$\begin{aligned} \frac{V_{ink}}{n_k} &= V_0, \\ \frac{L_k}{n_k^2} &= L_0 \end{aligned} \quad (13)$$



Obr. 6 a) Fázový posun jednotlivých mostíkov (hore) a napätie medziobvodu (dole);
b) náhradná schéma MAB meniča

Za predpokladu, že porty MAB meniča sú ekvivalentné a že teda magnetizačné indukčnosti sú vo všetkých portoch rovnaké, potom môžeme prepísať rovnicu (10) na rovnicu (14), kde L_{mag}

je magnetizačná indukčnosť jedného z portov prepočítaná na sekundárnu stranu.

$$v_{ac} = \frac{1}{N} \left(\frac{\sum_{i=1}^N v_k}{\frac{1}{L_0} + \frac{1}{L_{mag}}} \right) \quad (14)$$

Maximálnu hodnotu napätia medziobvodu je potom možné vyjadriť rovnicou (15)

$$V_{ac\ max} = \left(\frac{V_0}{\frac{1}{L_0} + \frac{1}{L_{mag}}} \right) \quad (15)$$

Z uvedeného vyplýva, že amplitúda napätia medziobvodu v_{ac} závisí od hodnoty L_{mag} a za určitých podmienok uvedených v (16) môžeme tvrdiť, že úbytok napätia v_{ac} vzhľadom na hodnotu magnetizačnej indukčnosti bude 1% z ideálnej hodnoty napätia medziobvodu v_{ac} , pre rôzny počet portov.

$$L_{mag, \text{ primar}} > 100L_k \quad (16)$$

2) Stredná hodnota prúdu

Strednú hodnotu prúdu k – teho portu v závislosti na porte j , môžeme označiť ako i_{kj} a je definovaná v rovnici (17). Rovnica je založená na popise DAB meniča [44]:

$$i_{kj} = K_{kj} V_{inj} (\phi_j - \phi_k) (1 - |\phi_j - \phi_k|) \quad (17)$$

kde

$$K_{kj} = T \frac{n_k n_j}{L_k L_j} \quad (18)$$

V prípade že chceme vyjadriť celkový prúd k – teho portu, musíme zobrať do úvahy príspevky prúdu všetkých portov ako v rovnici (19)

$$i_k = \sum_{j=1}^N K_{kj} V_{inj} (\phi_j - \phi_k) (1 - 2|\phi_j - \phi_k|) \quad (19)$$

Výsledná rovnica pre celkový prúd jednotlivých portov je potom vyjadrená v maticovom tvare (20). pričom platí, že súčet všetkých výkonov na jednotlivých portoch je rovný nule (21)

$$\sum_{k=1}^M i_k V_{ink} = 0 \quad (21)$$

B. Riadenie MAB meniča

Prvým cieľom z pohľadu riadenie MAB meniča je nájsť hodnoty fázových posunov jednotlivých mostíkov ϕ_k , ktoré zabezpečia požadovaný tok výkonu pre všetky porty meniča. Ak sú napätia portu V_{ink} známe, výzvou ostáva správne nastaviť vstupné prúdy I_k jednotlivých portov. Pretože vstupný prúd každého portu závisí od radiacích veličín ϕ týchto portov, MAB menič môžeme považovať za systém s viacerými vstupmi

a viacerými výstupmi (MIMO), kde cieľom je určiť hodnoty N fázových posunov podľa N požadovaných prúdov, resp. výkonov. Vzťah medzi prúdmi a fázovými posunmi mostíkov však v prípade MAB meniča nie je lineárny, a nie je tak možné odvodiť rovnice pre fázové posuny (20) v celom pracovnom rozsahu meniča.

1) Metódy odstránenia nelineárnej väzby

Tento problém je analyzovaný v rôznych odborných príspevkoch, kde napríklad v [45] autori navrhujú použitie algoritmu založeného na Newton-Raphsonovej metóde hľadania zodpovedajúcich fázových posunov na získanie požadovaných hodnôt vstupných prúdov alebo výkonov. V príspevku [46] je prezentovaná problematika popisu trojfázového TAB meniča, kde autori ako výsledný opis meniča uvádzajú set kvartických polynómov. Tieto polynómy obsahujú fázové posuny, ktoré je pre požadovaný výkon portov potrebné nastaviť na požadovanú hodnotu. Riešenie týchto polynómov je možné za pomoci algoritmu predstaveného od Ferrari v [47]. Prístup od Ferrariho prezentovaný v [47] vyžaduje aj riešenie kubických rovnic napríklad Cardanovou metódou, čo spôsobuje potrebu značného matematického aparátu. Preto sú numerické prístupy založené na Newtonovej metóde v porovnaní s inými metódami pre výpočet fázových posunov v praxi uskutočniteľnejšie. Napríklad pre k – ty port meniča je možné pomocou Newtonovej iteračnej metódy linearizovať rovnicu vstupného prúdu portu v závislosti na fázových posunoch ostatných portov ϕ_k a nájsť tak vzťah medzi zmenou prúdu portu Δi_k a zmenou fázového posunu $\Delta \phi_k$. Použitím tejto metódy je odvodená matica \mathbf{A} , ktorá v sebe zahŕňa parametre obvodu a hodnoty fázových posunov. Finálne rovnice pre implementáciu takejto metódy spolu s rovnicou (20) sú (22)(23)[31]

$$\begin{bmatrix} \Delta i_1 \\ \Delta i_2 \\ \vdots \\ \Delta i_N \end{bmatrix} = \mathbf{A} \begin{bmatrix} \Delta \phi_1 \\ \Delta \phi_2 \\ \vdots \\ \Delta \phi_N \end{bmatrix} \rightarrow \begin{bmatrix} \Delta \phi_1 \\ \Delta \phi_2 \\ \vdots \\ \Delta \phi_N \end{bmatrix} = \mathbf{A}^{-1} \begin{bmatrix} \Delta i_1 \\ \Delta i_2 \\ \vdots \\ \Delta i_N \end{bmatrix} \quad (22)$$

a pre maticu \mathbf{A} platí (23). Ďalej budú uvedené dva rôzne prístupy pre odstránenie nelineárnej väzby medzi portami MAB meniča.

a) Numerické metódy odstránenia nelineárnej väzby

Základom Newtonovej iteračnej metódy je funkcia $f(x)$ rozšírená Taylorovým rádom v bode x , kde používa rovnicu (24) pre iteračnú aproximáciu koreňa funkcie $f(x) = 0$.

$$x_k = x_{k-1} - \frac{f(x_{k-1})}{f'(x_{k-1})} \quad (24)$$

kde $k = 1, 2, \dots$ označuje iteračný krok a x je aproximovaný výsledok. Newtonova metóda vždy iteruje od začiatku intervalu smerom k väčším fázovým posunom (väčším uhlom zaťaženia) a teda aproximuje k prvému koreňu funkcie v rámci daného

$$\begin{bmatrix} i_1 \\ i_2 \\ \vdots \\ i_M \end{bmatrix} = \begin{bmatrix} 0 & K_{12}(\phi_2 - \phi_1)(1 - 2|\phi_2 - \phi_1|) & \cdots & K_{1M}(\phi_M - \phi_1)(1 - 2|\phi_M - \phi_1|) \\ K_{21}(\phi_1 - \phi_2)(1 - 2|\phi_1 - \phi_2|) & 0 & \cdots & K_{2M}(\phi_M - \phi_2)(1 - 2|\phi_M - \phi_2|) \\ \vdots & \vdots & \ddots & \vdots \\ K_{M1}(\phi_1 - \phi_M)(1 - 2|\phi_1 - \phi_M|) & K_{M2}(\phi_2 - \phi_M)(1 - 2|\phi_2 - \phi_M|) & \cdots & 0 \end{bmatrix} \cdot \begin{bmatrix} V_{in1} \\ \vdots \\ V_{inM} \end{bmatrix}$$

(20)

$$\mathbf{A} = \begin{bmatrix} -\sum_{k \neq 1} V_{ik} K_{1k} (1 - 2|\phi_k - \phi_1|) & V_{i2} K_{12} (1 - 2|\phi_1 - \phi_2|) & \cdots & V_{iN} K_{N1} (1 - 2|\phi_1 - \phi_N|) \\ -V_{i1} K_{12} (1 - 2|\phi_2 - \phi_1|) & -\sum_{k \neq 2} V_{ik} K_{2k} (1 - 2|\phi_k - \phi_2|) & \cdots & V_{iN} K_{N2} (1 - 2|\phi_2 - \phi_N|) \\ \vdots & \vdots & \ddots & \vdots \\ -V_{i1} K_{N1} (1 - 2|\phi_N - \phi_1|) & -V_{i2} K_{N2} (1 - 2|\phi_2 - \phi_N|) & \cdots & -\sum_{k \neq N} V_{ik} K_{Nk} (1 - 2|\phi_k - \phi_N|) \end{bmatrix}$$

(23)

intervalu. Výsledok aproximácie zodpovedá najmenšiemu fázovému posunu potrebnému pre prenos požadovaného výkonu.

Ak nie je známy výsledný interval fázových posunov pre požadované výkony, Newtonova metóda musí byť opakovane aplikovaná na každý polynóm pre dosiahnutie požadovaných výkonových pomerov v meniči. V dôsledku kvadratickej konvergencie Newtonovej metódy sa absolútna chyba znižuje rýchlo so zvyšujúcim sa počtom iterácií.

Ako bolo uvedené vyššie, pre riešenie nelineárnych závislostí v rovniciach popisujúcich MAB menič je použitie Newton-Raphsonovej metódy veľmi výhodné vzhľadom na náročnosť výpočtu a možnosť implementácie do aktuálne dostupného hardvéru. Newton-Raphsonova metóda linearizuje nelineárne rovnice a používa definovaný počet iterácií pre dosiahnutie požadovaného riešenia. Táto metóda sa tiež vyznačuje rýchlou konvergenciou k výsledku a nevyžaduje vysoký výpočtový výkon, čo umožňuje už spomínanú implementáciu do digitálne signálového procesora alebo FPGA obvodu. Rýchlosť konvergencie je závislá aj na počiatočných predpokladaných fázových posunoch. V niektorých špecifických prípadoch sú použiteľné aj iné metódy ako je Gauss-Seidelova iterácia a tzv. „Fast Decoupling Method“ [49]. Pre výpočet toku energie je za cenu vyššieho výpočtového výkonu tiež možné použiť rôzne výpočtové nástroje ako je Matpower [50]. Ďalej bude uvedený popis Newton-Raphsonovej metódy prispôbenej na analýzu toku energie v MAB meniči. Uvedený algoritmus je tiež programovo implementovaný do prostredia MATLAB, čím môže byť ďalej využitý pri riadení MAB meniča. Algoritmus požaduje nasledovné vstupy:

- parametre siete vrátane indukčností vetiev, magnetizačných indukčností, kapacít vetiev a odporov vo vetvách,
- požadovaný výkon jednotlivých portov a amplitúda napätia,
- amplitúda napätia referenčného portu,
- počiatočné predpokladané fázové posuny jednotlivých portov.

Samotný iteračný krok metódy je potom vyjadrený rovnicami:

$$\begin{bmatrix} \Delta\phi_1 \\ \vdots \\ \Delta\phi_{n-1} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial \phi_1} & \cdots & \frac{\partial f_1}{\partial \phi_{n-1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{n-1}}{\partial \phi_1} & \cdots & \frac{\partial f_{n-1}}{\partial \phi_{n-1}} \end{bmatrix}^{-1} \begin{bmatrix} \Delta P_1 \\ \vdots \\ \Delta P_{n-1} \end{bmatrix} \quad (25)$$

$$\phi_k = \phi_{k-1} - \Delta\phi \quad (26)$$

$$\frac{\partial f_i}{\partial \phi_j} = \begin{cases} \frac{V_i V_j}{\omega L_{ij}} \left(\frac{2|\Phi(\phi_{ij})|}{\pi} - 1 \right) & j \neq i \\ \sum_{k \neq i} \frac{V_i V_k}{\omega L_{ik}} \left(1 - \frac{2|\Phi(\phi_{ik})|}{\pi} \right) & j = i \end{cases} \quad (27)$$

Pri vhodne zvolených počiatočných fázových posunoch je možné výsledné fázové posuny jednotlivých portov nájsť v priebehu niekoľkých iteračných krokov. Ak sa jednotlivé porty meniča realizujú ako mostíkové striedače, potom tieto porty môžeme označiť ako P s riadením iba činnou zložkou výkonu. Za

predpokladu, že celkový počet portov je n , jeden port je považovaný za referenčný s fázovým posunom $\phi = 0$ a zvyšný počet portov je m , jeden iteračný krok je potom vyjadrený ako (28) a (29)

$$\begin{bmatrix} \Delta V_1 \\ \vdots \\ \Delta V_m \\ \Delta \phi_1 \\ \vdots \\ \Delta \phi_{n-1} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial v_1} & \dots & \frac{\partial f_1}{\partial v_m} & \frac{\partial f_1}{\partial \phi_1} & \dots & \frac{\partial f_1}{\partial \phi_{n-1}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{n-1}}{\partial v_1} & \dots & \frac{\partial f_{n-1}}{\partial v_m} & \frac{\partial f_{n-1}}{\partial \phi_1} & \dots & \frac{\partial f_{n-1}}{\partial \phi_{n-1}} \end{bmatrix}^{-1} \begin{bmatrix} \Delta P_1 \\ \Delta P_2 \\ \vdots \\ \Delta P_{n-1} \end{bmatrix} \quad (28)$$

$$V_k = V_{k-1} - \Delta V, \quad \phi_k = \phi_{k-1} - \Delta \phi \quad (29)$$

Z uvedeného vyplýva, že existuje $n - 1 + m$ neznámych premenných a $n - 1$ nelineárnych rovníc pričom Jacobiánska matica nie je štvorcovou maticou. Preto pre výpočet veľkosti iteračného kroku je potrebné použiť Moore-Penroseovu pseudoinverziu Jacobiánskej matice [47].

Ako bolo opísané v kapitole vyššie, nie je jednoduché odvzäbiť riadiacu veličinu, ktorou je fázový posun pre dosiahnutie požadovaného vstupného prúdu na konkrétnom porte. Preto bol najprv teoreticky opísaný iteračný algoritmus, ktorý je na základe správnych vstupných parametrov a počiatočných fázových posunov schopný dospieť k fázovým posunom potrebným pre požadovaný prúd počas niekoľkých iterácií. Pre implementáciu do digitálneho signálového procesora alebo FPGA obvodu je však vhodné mať tento algoritmus vo forme vývojového diagramu, prípadne uvedený postup krokov algoritmu. Vývojový diagram je zobrazený na Obr. 7. Sekvencia krokov potom je:

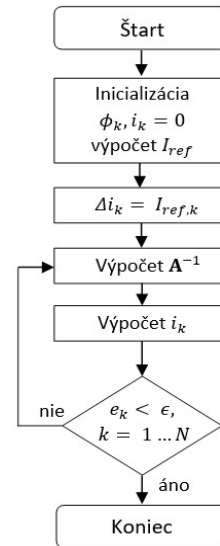
1. Inicializácia riadiacich veličín a vstupných prúdov (štandarde rovných nule) a výpočet referenčných vstupných prúdov na základe $I_{ref,k} = \frac{P_{ref,k}}{V_{in,k}}$ v každom k - tom porte, od portu k_1 po port k_{n-1} . Tento krok môže byť vynechaný v prípade, že žiadanými hodnotami sú prúdy jednotlivých portov. Ak počiatočný bod riadiacej veličiny je nastavený na nulu, vypočítané fázové posuny budú tiež nulové.
2. V prvej iterácii je odchýlka vstupného prúdu konkrétneho portu rovná žiadanej hodnote prúdu tohto portu.
3. Výpočet inverznej matice \mathbf{A} , vektora odchýliek riadiacej veličiny $\Delta \phi$ a nové hodnoty riadiacich veličín $\phi[m] = \phi[m - 1] + \Delta \phi[m]$, kde $[m]$ predstavuje m - tú iteráciu.
4. S novými hodnotami riadiacej veličiny $\phi[m]$ vypočítanými v predchádzajúcom kroku môžu byť vypočítané nové hodnoty vstupných prúdov i_1, \dots, i_n .
5. Výpočet absolútnej chyby medzi žiadaným a vypočítaným vstupným prúdom $e_k = I_{re} - i_k$. Ak sú všetky chyby menšie ako určená tolerancia,

algoritmus končí výpočet a riadiace premenné ϕ_k sú nastavené na MAB meniči.

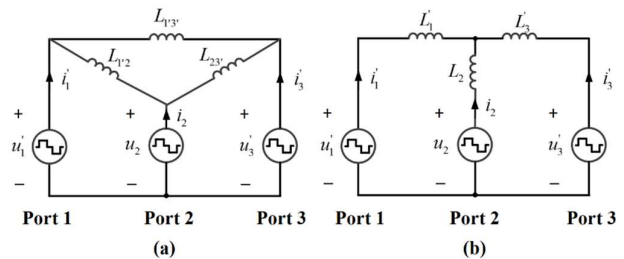
Pri štandardnej prevádzke meniča tento algoritmus beží nepretržite a v ideálnom prípade by jedna iterácia mala byť vykonaná za menej ako spínaciu periódu pre minimalizovanie časového oneskorenia. Taktó časovo optimálny výpočet však nie je možné vždy dosiahnuť a oneskorenie je potom potrebné zohľadniť pri návrhu ďalšieho regulátora s uzavretou slučkou.

a) *Hardvérové odstránenie nelineárnej väzby*

Vzhľadom na analýzu uvedenú vyššie je zjavné, že rozptylová indukčnosť transformátora predstavuje významný komponent pre prenos výkonu medzi portami. Náhradná schéma MAB meniča môže byť opísaná buď v zapojení do hviezdy, alebo v zapojení do trojuholníka. Na Obr. 8 je náhradná schéma trojportového meniča v zapojení do trojuholníka aj do hviezdy, čo zodpovedá aj pôvodnej analýze v kapitole a Obr. 6 b).



Obr. 7 Vývojový diagram iteračného algoritmu



Obr. 8 Náhradná schéma trojportového meniča v zapojení a) do trojuholníka b) do hviezdy

Z tejto náhradnej schémy je potom možné vyjadriť nové hodnoty indukčností ako

$$\begin{cases} L_{1'2} = L_1' + L_2 + \frac{L_1' L_2}{L_3'} \\ L_{1'3'} = L_1' + L_3' + \frac{L_1' L_3'}{L_2} \\ L_{23'} = L_2 + L_3' + \frac{L_2 L_3'}{L_1'} \end{cases} \quad (30)$$

Pri opise ďalej uvažujeme tzv. „single phase shift (SPS)“ riadenie, kde u_1', u_2 a u_3' sú napätia s pravouhlým priebehom a striedou 50%. Tiež predpokladáme, že:

- napätie u_1' predbieha napätie u_2 o fázový posun ϕ_{12} ,
- napätie u_2 predbieha napätie u_3' o fázový posun ϕ_{23} ,
- napätie u_1' predbieha napätie u_3' o fázový posun ϕ_{13} .

Prenos výkonu trojportového meniča potom môže byť definovaný nasledujúcimi rovnicami:

$$\begin{cases} P_1 = P_{12} + P_{13} \\ P_2 = -P_{12} + P_{23} \\ P_3 = -P_{13} - P_{23} \end{cases} \quad (31)$$

pričom

$$\begin{cases} P_{12} = \frac{U_1' U_2 T_s \phi_{12} (1 - |\phi_{12}|)}{2L_{1'2}} \\ P_{23} = \frac{U_2 U_3' T_s \phi_{12} (1 - |\phi_{23}|)}{2L_{23'}} \\ P_{13} = \frac{U_1' U_3' T_s \phi_{13} (1 - |\phi_{13}|)}{2L_{1'3'}} \\ \phi_{13} = \phi_{12} + \phi_{23} \end{cases} \quad (32)$$

Z vyššie uvedených rovníc je zrejmé, že vzťah medzi prenosom výkonu medzi portami P_1, P_2, P_3 a fázovými posunmi $\phi_{12}, \phi_{23}, \phi_{13}$ medzi týmito portami je veľmi komplikovaný, pretože medzi nimi existujú silné nelineárne väzby. Ako lineárne väzby môžeme definovať väzby medzi výkonmi P_1, P_2, P_3 a P_{12}, P_{23}, P_{13} , kvadratické nelineárne väzby sú medzi P_{12}, P_{23}, P_{13} a $\phi_{12}, \phi_{23}, \phi_{13}$ a lineárne väzby aj medzi ϕ_{12}, ϕ_{23} a ϕ_{13} . Ak uvažujeme s počiatočnými podmienkami, že $\phi_{12}, \phi_{23} \in [0, \frac{1}{2}]$, potom z rovníc (31) a (32) môžeme odvodiť vzťah medzi fázovým posunom ϕ_{23} a výkonmi P_1 a P_2 portu 1 a portu 2 ako:

$$\begin{aligned} & [(ab + bc + ac)^2 - 4abc^2] \phi_{23}^4 + \\ & [2b(ab + bc + ac)(a + c) - 4abc^2] \phi_{23}^3 + \\ & [P_2(4ac^2 - 2(a - c)(ab + bc + ac)) - 2aP_1(ab + bc + \\ & \quad ac) + a^2c^2 - b^2(a + c)^2] \phi_{23}^2 + \\ & [P_2(4abc + a^2b + bc^2) + P_1(2abc + 2a^2b)] \phi_{23} - \\ & \quad a^2P_1^2 - (a^2 + 2ac)P_1P_2 - (a + c)^2P_2^2 = 0 \end{aligned} \quad (33)$$

kde

$$a = \frac{U_1' U_2 T_s}{2L_{1'2}}, b = \frac{U_2 U_3' T_s}{2L_{23'}}, c = \frac{U_1' U_3' T_s}{2L_{1'3'}} \quad (34)$$

Z rovnice (5.68) vyplýva, že vzťah medzi fázovým posunom ϕ_{23} a výkonmi P_1 a P_2 je kvartický, čo bolo popísané aj v úvode kapitoly. Aj keď je možné vypočítať koreň kvartickej rovnice, riešenie takejto rovnice je v praxi často veľmi zložitá nato, aby ho bolo možné použiť. Veľká miera nelinearity ale spôsobuje ťažkosti pri analýze prúdového namáhania a efektivity meniča a tiež pri realizácii vhodného riadenia, preto je odstránenie resp. zníženie nelinearity a redukovanie väzieb medzi portami potrebné. Ako vyplýva z uvedeného vyššie, zmenou fázového posunu portu 1 zabezpečíme tok výkonu P_{12} , a P_{13} z portu 1 do portu 2 a portu 3. Keďže prenosová cesta medzi dvoma portami je ovplyvnená silnými väzbami a nelinearitami, tieto väzby musia byť zmiernené, resp. odstránené spôsobom redukcie prenosových ciest výkonu. Napríklad odstránením indukčnosti $L_{1'3'}$ z náhradnej schémy bude nová náhradná schéma ako na Obr. 9 a) a rovnice definujúce výkon zmenia svoj tvar na

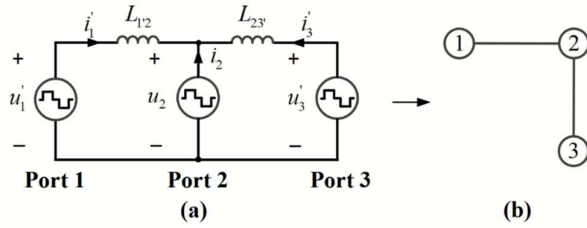
$$\begin{cases} P_1 = P_{12} = \frac{U_1' U_2 T_s \phi_{12} (1 - |\phi_{12}|)}{2L_{1'2}} \\ P_2 = -P_1 - P_3 \\ P_3 = -P_{23} = \frac{U_1' U_3' T_s \phi_{13} (1 - |\phi_{13}|)}{2L_{1'3'}} \end{cases} \quad (35)$$

Z rovníc (35) je zrejmé, že väzby medzi výkonmi P_1 a P_3 súvisiace s fázovými posunmi ϕ_{12}, ϕ_{23} sú po odstránení indukčnosti oslabené. Kvartická rovnica sa tak zmení na kvadratickú, ktorú je možné riešiť funkciou odmocniny. Aby sme dosiahli odväzbenie výkonu na porte 2 ako je znázornené na Obr. 9 a), musí byť splnená nasledujúca podmienka:

$$L_2 \ll L_1', L_3' \quad (36)$$

Táto metóda môže byť rozšírená aj pre väčší počet portov, pričom samotná metóda sa opiera o teóriu grafov. Základom je zjednodušiť náhradnú schému meniča na topologický graf, kde jeden port meniča predstavuje vrchol a cesty pre prenos výkonu (indukčnosť) medzi portami predstavujú ohodnotené hrany medzi týmito vrcholmi. Platí pritom pravidlo, že ak existuje n vrcholov (portov) a m hrán (ciest pre prenos výkonu/indukčností) v grafe G , pričom uvažujeme, že náhradná schéma MAB meniča je ekvivalentná s grafom G , potom požiadavka odväzbenia je splnená a zároveň neexistujú žiadne cykly (slučky) v grafe G . Definícia takéhoto stromového grafu v diskretnéj matematike je fakticky vyjadrená ako: G je stromový graf, kde je $m = n - 1$, čo znamená, že počet potrebných indukčností sa rovná počtu portov mínus jeden [51]. Jedna indukčnosť navyše spôsobí, že port nebude odväzbený a naopak, jedna indukčnosť menej spôsobí, že nebude splnená požiadavka prenosu výkonu medzi všetkými portami. Platí tiež, že graf G má nasledujúce vlastnosti: existuje jedna a iba jedna cesta medzi každým párom vrcholov, teda cesta medzi dvomi portami je jedinečná. Dosiahnutie vyššie uvedených podmienok je v praxi možné iba hardvérovou úpravou už existujúceho meniča alebo hardvérovo už počas návrhu meniča. Podmienky tejto metódy je možné dosiahnuť aj zmenou topológie a teda reorganizáciou portov, avšak menič by musel

po takejto zmene vyhovovať požiadavkám na prenos výkonu, čo nie je vždy splniteľné.



Obr. 9 a) náhradná schéma po aplikovaní metódy odväzbenia b) trojportový menič premietnutý do teórie grafov

2) Výpočet fázového posunu v závislosti od žiadaného prúdu

Vzhľadom na vyššie popísanú problematiku spojenú s nelineárnymi väzbami v matematickom popise MAB meniča a po naštudovaní problematiky spojenej s možnosťami odstránenia týchto nelinearit, sme pre náš trojportový aktívny mostík odvodili rovnicu fázového posunu v závislosti od žiadaného prúdu z rovnice (17). Nato, aby bolo možné štandardnými numerickými metódami vyjadriť rovnicu pre fázový posun, je potrebné najprv upraviť rovnicu prúdu (17) a zároveň definovať, v ktorých prípadoch bude naše odvodenie platiť. Ak uvažujeme trojportový aktívny mostík, potom prípady, kedy bude daná rovnica platiť sú:

- ak jeden z troch portov predstavuje záťažový port a ostatné dva porty meniča pracujú ako zdrojové porty s rovnakým fázovým posunom,

- ak dva z troch portov predstavujú záťažové porty s rovnakým fázovým posunom a tretí port predstavuje zdrojový port s fázovým posunom rovným 0.

Ďalej uvedieme rovnicu prúdu a aj samotné odvodenie rovnice pre ϕ_k v niekoľkých krokoch:

$$i_k = K_{kj} \cdot V_{inj} \cdot (\phi_j - \phi_k) \cdot (1 - 2(\phi_j - \phi_k)) + K_{kl} \cdot V_{inl} \cdot (\phi_l - \phi_k) \cdot (1 - 2(\phi_l - \phi_k)) \quad (37)$$

ak zavedieme poruchy pre fázový posun $\phi = \Phi + \hat{\phi}$ a prúd $i_k = I_k + \hat{i}_k$, kde Φ a I_k sú ustálené hodnoty, môžeme rovnicu prepísať ako

$$I_k = K_{kj} \cdot V_{inj} \cdot (\phi_j - \phi_k) \cdot (1 - 2(\phi_j - \phi_k)) + K_{kl} \cdot V_{inl} \cdot (\phi_l - \phi_k) \cdot (1 - 2(\phi_l - \phi_k)) \quad (38)$$

pričom už pri porovnaní rovnice (37) s rovnicou (17) je zrejme, že absolútna hodnota bola z výrazu odstránená. Ďalej sme postupovali štandardnými numerickými metódami. Roznásobením a presunutím prúdu na druhú stranu rovnice

$$4 \cdot K_{kj} \cdot V_{inj} \cdot \phi_k \cdot \phi_j + K_{kj} \cdot V_{inj} \cdot \phi_j - 2 \cdot K_{kj} \cdot V_{inj} \cdot \phi_j^2 - K_{kj} \cdot V_{inj} \cdot \phi_k - 2 \cdot K_{kj} \cdot V_{inj} \cdot \phi_k^2 + 4 \cdot K_{kl} \cdot V_{inl} \cdot \phi_k \cdot \phi_l + K_{kl} \cdot V_{inl} \cdot \phi_l - 2 \cdot K_{kl} \cdot V_{inl} \cdot \phi_l^2 - K_{kl} \cdot V_{inl} \cdot \phi_k - 2 \cdot K_{kl} \cdot V_{inl} \cdot \phi_k^2 - I_k = 0$$

a prepísaním do štandardnej formy $ax^2 + bx + c = 0$ dostávame

$$-(2K_{kj}V_{inj} + 2K_{kl}V_{inl})\phi_k^2 + (4K_{kj}V_{inj}\phi_j - K_{kj}V_{inj} + 4K_{kl}V_{inl}\phi_l - K_{kl}V_{inl})\phi_k + K_{kj}V_{inj}\phi_j - 2K_{kj}V_{inj}\phi_j^2 + K_{kl}V_{inl}\phi_l - 2K_{kl}V_{inl}\phi_l^2 - I_k = 0 \quad (40)$$

a ďalej riešime ako kvadratickú rovnicu s výsledkom:

$$\phi_k = -\frac{-4K_{kj}V_{inj}\phi_j + K_{kj}V_{inj} - 4K_{kl}V_{inl}\phi_l - \sqrt{(-16K_{kj}V_{inj}K_{kl}V_{inl}\phi_j^2 + 32K_{kj}V_{inj}K_{kl}V_{inl}\phi_j\phi_l + K_{kj}^2 \cdot V_{inj}^2 + K_{kl}^2 \cdot V_{inl}^2 + 2 \cdot K_{kj} \cdot V_{kj} \cdot K_{kl} \cdot V_{kl} - 8 \cdot K_{kj} \cdot V_{kj} \cdot \phi_j - 8 \cdot K_{kl} \cdot V_{kl} \cdot \phi_l - 16 \cdot K_{kj} \cdot V_{kj} \cdot K_{kl} \cdot V_{kl} \cdot \phi_l^2)}}{2 \cdot (-2 \cdot K_{kj} \cdot V_{kj} - 2 \cdot K_{kl} \cdot V_{kl})} \quad (41)$$

kde $\phi_j = 0$ a $\phi_l = 0$. Takto definovanú rovnicu ďalej použijeme pri simulačnom overení trojportového aktívneho mostíka.

Vyššie uvedené podmienky síce značne zužujú široké pracovné spektrum výmeny energie medzi portami meniča, avšak spomínaná Newton-Raphsodyho metóda nerieši problém s odstránením nelineárnej väzby ako taký, ponúka riešenie, ako sa pomocou niekoľkých iteračných krokov dostať k požadovaným fázovým posunom vzhľadom na výkony žiadané na portoch meniča. Pri tomto iteračnom algoritme je nutné poznamenať, že samotný výpočet by mal byť časovo optimalizovaný tak, aby ho hardvér stihol vykonať počas jedného spínacieho cyklu, čo nie je vždy možné. Výrazne sa to môže prejaviť pri väčšom počte portov meniča. Najvhodnejšie, aj keď v určitých situáciách použiteľné, nie je ani hardvérové odväzbenie založené na podmienke, že hodnota jednej indukčnosti v obvode bude oveľa menšia ako ostatných indukčností. Pri splnení takejto podmienky už nie je možné označiť MAB menič ako vysoko modulárny, keďže pri prípadnej zmene topológie, tzn. pridaní/odobratí portov je nutné brať do úvahy túto podmienku a tomu prispôsobiť aj indukčnosti nachádzajúce sa v novej topológii.

IV. SIMULAČNÉ VÝSLEDKY

Simulačné overenie trojportového meniča prebiehalo v programe PLECS Blockset, ktorý je priamo prepojený s pracovným prostredím softvéru MATLAB, čo umožňovalo efektívne spracovanie výsledkov. Na verifikáciu poznatkov uvedených v predchádzajúcich kapitolách sme využili trojportový výkonový menič. V tomto prípade je vhodnejšie použiť terminológiu bližšiu anglickému pomenovaniu MAB – Multi Active Bridge, resp. v našom prípade TAB – Triple

Active Bridge a teda „trojitý aktívny mostík“. Meniče, kde sa kladie dôraz na ich vysokú modularitu a teda jednoduché pridávanie/odoberanie portov je potom ešte vhodnejšie označovať terminológiou MMAB – Modular Multi Active Bridge. Toto označenie je tiež vhodné aj v našom prípade.

Každý modul v simulačnej schéme pozostáva z aktívneho mostíka napájaného ideálnym napäťovým zdrojom, ktorý ďalej odovzdáva energiu cez planárny transformátor. Planárny transformátor je reprezentovaný modelom magnetického jadra s rozptylovou indukčnosťou a odpormi jednotlivých vinutí. Všetky tri transformátory majú rovnaké parametre.

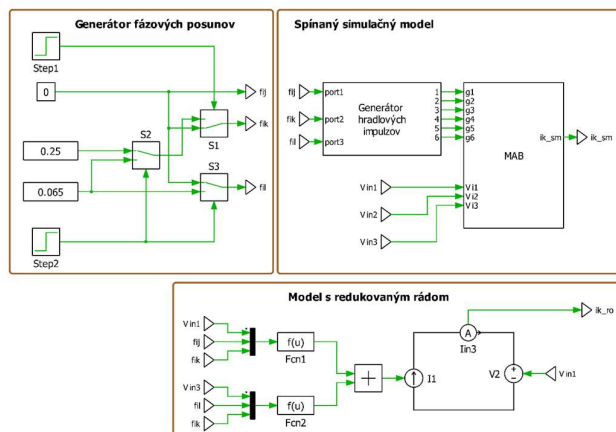
Pre takto zostavený trojportový aktívny mostík bol odvodený aj model s redukovaným rádom podľa [48]-[50]. Model bol zostavený pre situáciu, ak jeden z troch portov predstavuje záťažový port a ostatné dva porty meniča pracujú ako zdrojové porty s rovnakým fázovým posunom. Samotný model vychádzal z rovnice (19) a jeho hlavným cieľom je porovnať zostavený simulačný model s magnetickým jadrom transformátora s modelom redukovaného rádu, ktorý je založený na matematickom popise. Simulačné výsledky sú rozdelené do dvoch častí. V prvej časti je trojitý aktívny mostík simulovaný bez regulácie a teda pre požadovaný výkon na jednotlivých portoch je dopočítaný potrebný fázový posun podľa rovnice (41). Napájacie napätia mostíkov sú pre všetky porty rovnaké a konštantné, zdrojom napätia je pre všetky porty ideálny napäťový zdroj. Sledované sú prúdy jednotlivých portov. V druhej časti simulácie je ideálny napäťový zdroj na porte 2 nahradený záťažou reprezentovanou filtračnou kapacitou a premenlivou odporovou záťažou. Simulácia je doplnená o regulátor výstupného napätia, ktorý zabezpečuje zmenu fázového posunu vzhľadom na zmeny záťaže alebo na zmenu požadovaného výstupného napätia. Simulácia viacportového meniča prebiehala súčasne na spínanom simulačnom modeli, ako aj na modeli redukovaného rádu. Preto sú niektoré priebehy označené indexom „SM“ ako „simulačný model“ a niektoré ako „RM“ a teda „redukovaný model“. Orientácia prúdov počas simulácie bola stanovená tak, že prúd vtekajúci do portu (smerom zo zdroja do striedača) má kladný smer (kladné hodnoty - zdroj) a prúd vytekajúci z portu (smerom zo striedača do zdroja) má záporný smer (záporné hodnoty - spotrebič).

A. Simulácia trojitého aktívneho mostíka bez regulátora

Simulácia bola vykonaná ako pre spínaný simulačný model, tak aj pre model s redukovaným rádom. Výsledky oboch modelov sú porovnané v priebehoch. Celkový simulačný model je zobrazený na Obr. 10 a je rozdelený do troch celkov: generátor fázových posunov, spínaný simulačný model a model s redukovaným rádom. Generátor fázových posunov zabezpečuje zmenu fázových posunov na hodnoty vypočítané podľa nižšie popísaného postupu. Spínaný simulačný model obsahuje generovanie hradlových impulzov podľa fázových posunov z generátora a kompletný spínaný simulačný model. Model s redukovaným rádom bol reprezentovaný blokom „Function“, tieto bloky sú v simulačnej schéme označené ako „Fcn1“ a „Fcn2“ a obsahujú matematický popis – rovnicu modelu redukovaného rádu trojportového aktívneho mostíka.

Model je do dvoch rovníc rozdelený preto, že prúd k – teho mostíka je zložený zo súčtu prúdov l -teho a j -teho mostíka.

Vstupné napätia jednotlivých portov v simulačnom modeli označené ako V_{in1} , V_{in2} a V_{in3} zodpovedali napätiu primárneho vinutia transformátora, a teda 24 V. Porty trojitého aktívneho mostíka boli počas simulácie označené ako P1, P2 a P3. Simulácia prebiehala v troch časových intervaloch, kde v každom intervale bol nastavený iný pomer odoberaných / dodávaných prúdov jednotlivých portov. Fázové posuny a prúdy sú zobrazené na Obr. 11. Pomery v obvode počas

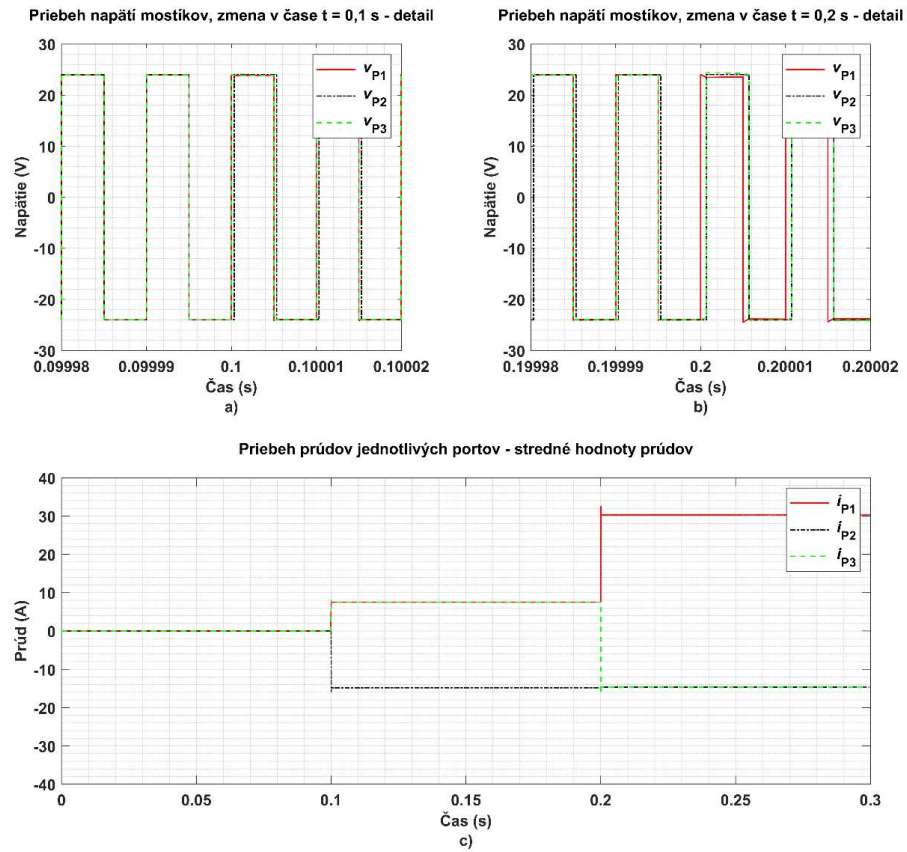


Obr. 10 Celková simulačná schéma trojportového meniča bez regulátora v softvéri PLECS

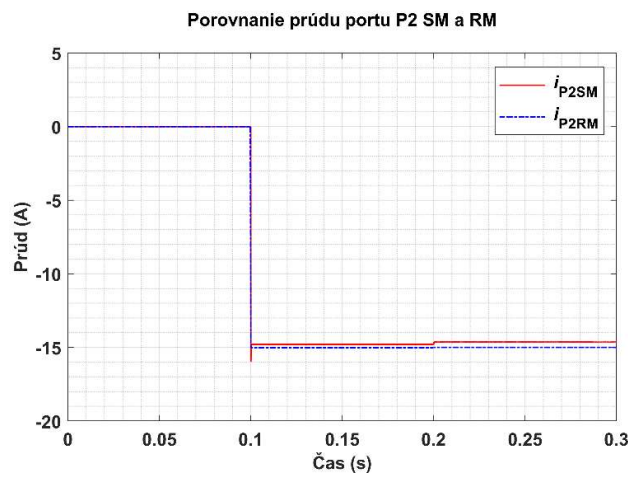
simulácie boli nasledovné:

- V čase $t = 0$ až $0,1$ s, bol dodávaný/odoberaný prúd všetkých portov nastavený na hodnotu 0 A. Táto požiadavka bola dosiahnutá nastavením fázových posunov ako $\phi_1 = \phi_2 = \phi_3 = 0$. Medzi portami meniča netiekol žiaden prúd a teda sa neprelieval žiaden výkon.
- V čase $t = 0,1$ až $0,2$ s, port P2 bol zvolený ako záťažný s odoberaným prúdom 15 A. Fázový posun pre port P2, označený ako ϕ_2 bol prepočítaný rovnicou (41) na hodnotu $\phi_2 = 0,0321$. Porty P1 a P3 predstavovali zdrojové porty, dodávali teda prúd do portu P2 a tento prúd bol medzi porty P1 a P3 rozdelený presne na polovicu.
- V čase $t = 0,2$ až $0,4$ s, port P1 bol zvolený ako zdrojový port a porty P2 a P3 boli zvolené ako záťažové porty s prúdom záťaže 15 A pre každý z portov. Fázové posuny pre porty P2 a P3 označené ako ϕ_2 a ϕ_3 , boli prepočítané rovnicou (41) pre celkový odoberaný prúd 30 A, čo znamená 15 A pre port P2 a 15 A pre port P3. Fázové posuny pre porty P2 a P3 potom boli $\phi_2 = \phi_3 = 0,0697$.

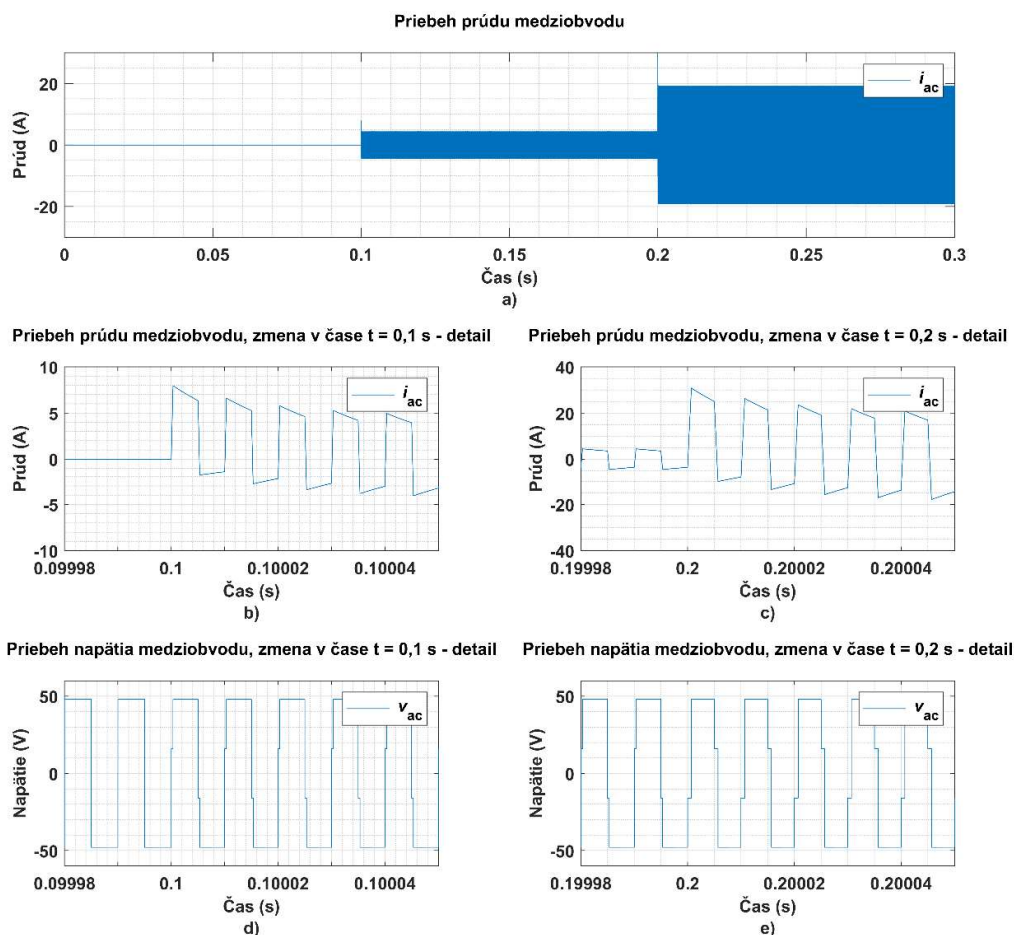
Na priebehoch môžeme pozorovať malé rozdiely v priebehoch spínaného modelu a priebehoch redukovaného modelu. Spínaný model je síce iba simulačný, avšak transformátor je modelovaný prostredníctvom parametrov



Obr. 11 a) Zmena fázového posunu striedačov v čase $t = 0,1$ s; b) zmena fázového posunu striedačov v čase $t = 0,2$ s; c) priebehy prúdov jednotlivých portov počas simulácie



Obr. 12 Porovnanie prúdu portu 2 spínaného modelu meniča a modelu s redukovaným rádom



Obr. 13 a) Priebeh prúdu medziobvodu, b) a c) detaily zmeny tohto prúdu pri zmene fázového posunu; d) a e) detail zmeny napätia medziobvodu pri zmene fázového posunu

skutočného jadra transformátora a mostíky zo súčiastok s určitými prechodovými odporami a teda aj stratami. Redukovaný model je založený na matematickom popise MAB meniča a je vyjadrený iba prostredníctvom rovníc. Treba však poznamenať, že ak simulačný model umožňuje implementovanie strát jednotlivých polovodičových súčiastok, resp. modelovanie magnetických obvodov, sú výsledky týchto simulácií bližšie parametrom nameraným na reálnom meniči.

B. Simulácia trojitého aktívneho mostíka s regulátorom

Na základe návrhu regulátora pre DAB sme rovnako pre trojportový aktívny mostík navrhli PI regulátor výstupného napätia druhého portu. Tento regulátor zabezpečí správne nastavenie fázového posunu mostíka portu 2 vzhľadom na požadované výstupné napätie a premenlivú záťaž na danom porte. Bloková schéma, z ktorej sme vychádzali pri návrhu regulátora pre trojportový aktívny mostík je zobrazená na Obr. 4. Regulátor typu PI bol navrhnutý podľa [52][53], kde

hlavnými parametrami pre návrh regulátora boli požadované tlmenia a požadovaný čas ustálenia

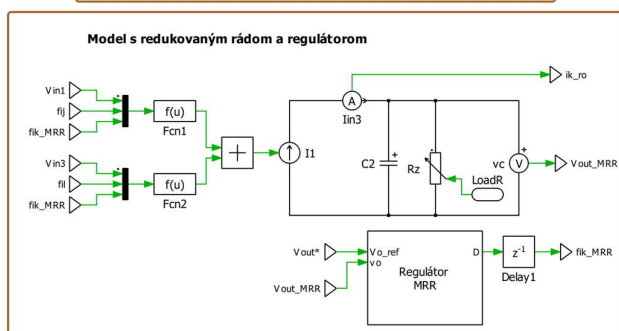
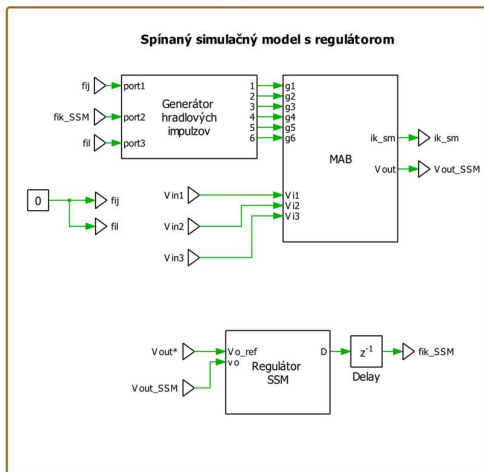
Výkonové pomery jednotlivých portov meniča boli počas simulácie definované tak, že porty P1 a P3 dodávali energiu do medziobvodu, predstavovali teda zdrojové porty s rovnakým fázovým posunom $\phi_{1,3} = 0$ a port dva túto energiu odoberal, bola na ňom pripojená záťaž. Vstupom PI regulátora bola žiadaná hodnota napätia na porte 2 a výstup regulátora, teda akčný zásah predstavoval fázový posun portu 2 vzhľadom na referenčné porty P1 a P3. Samotná záťaž na porte dva bola reprezentovaná filtračným kondenzátorom o kapacite 1 mF a premenlivou odporovou záťažou. Simulačný model meniča s regulátorom je na Obr. 14.

Parametre filtračného kondenzátora a premenlivej odporovej záťaže boli zohľadnené aj pri návrhu regulátora, kde bolo potrebné rozhodnúť, pre aký nominálny prúd portu bude regulátor navrhnutý. Tento prúd bol stanovený na hodnotu $i_2 = 10\text{ A}$, čo predstavuje pri nominálnom napätí portu V_{P3} hodnotu odporovej záťaže $R = 2,4\ \Omega$. Hodnota tohto prúdu ovplyvnila

prostredníctvom fázového posunu výsledný prenos samotného meniča a tiež aj výsledný prenos uzavretého obvodu pre návrh regulátora. Hodnota filtračnej kapacity bola konštantná. Vstupné napätia portov P1 a P3 boli konštantné s hodnotou $V_{P1}, V_{P3} = 24 V$. Simulácia prebiehala v štyroch časových

napätia zareagoval s požadovanou dynamikou a prostredníctvom zväčšenia fázového posunu portu dva od referenčných portov zvýšil napätie na porte dva na požadovanú hodnotu. Keďže záťaž je odporová a ostala bez zmeny, teda na hodnote $R = 1,2 \Omega$ so zväčšením napätia na porte dva sa zvýšil aj prúd odoberaný týmto portom. Ustálený stav interval trval až do času $t = 0,3 s$.

- V čase $t = 0,3 s$ bola aktuálna záťaž $R = 1,2 \Omega$ zväčšená na pôvodnú hodnotu $R = 2,4 \Omega$, čo spôsobilo zmenšenie odoberaného prúdu meniča na porte dva, ale taktiež skokový nárast napätia na porte dva v dôsledku odľahčenia obvodu výmeny energie. Regulátor znova zareagoval s očakávanou dynamikou a zmenšil fázový posun mostíka portu dva tak, aby bola na porte dva žiadaná hodnota výstupného napätia. Tento stav trval do času $t = 0,4 s$, kedy bola simulácia ukončená.



Obr. 14 Celkový simulačný model meniča s regulátorom

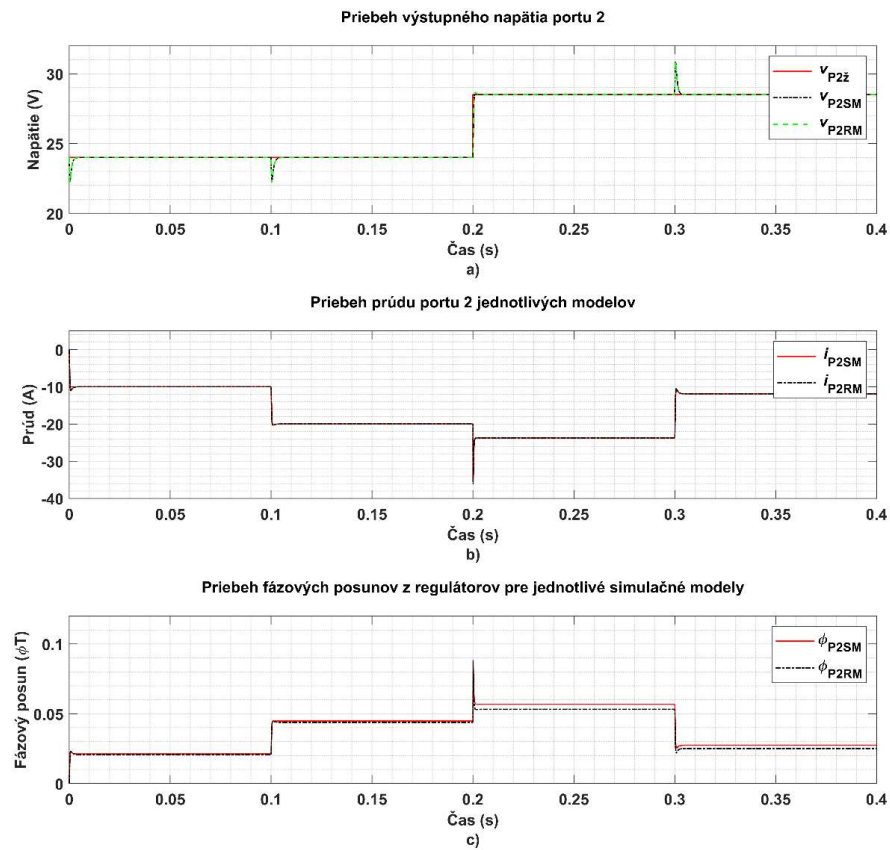
intervaloch, kde v každom intervale nastala zmena buď na záťaži, alebo zmena žiadanej hodnoty výstupného napätia. Pomery v obvode počas simulácie boli nasledovné:

- V intervale $t = 0$ až $0,1 s$ bola žiadaná hodnota výstupného napätia na porte dva $V_{P2} = 24 V$. Prúd dodávaný do tohto portu bol nastavený na hodnotu $I_2 = 10 A$. Regulátor snímal výstupné napätie portu dva a týmto podmienkam prispôbil fázový posun. Prúd odoberaný portom 2 bol dodávaný portom 1 a portom 3, kde hodnoty týchto prúdov boli rovnaké a teda polovičné ako na porte 2. (Obr. 15)
- V čase $t = 0,1 s$ nastala zmena v hodnote premenlivej záťaže z $R = 2,4 \Omega$ na hodnotu $R = 1,2 \Omega$, čo zvýšilo prúd odoberaný portom dva z $I_2 = 10 A$ na $I_2 = 20 A$. Zvýšenie odoberaného prúdu síce spôsobilo pokles napätia na porte dva, avšak regulátor upravil fázový posun tak, aby napätie na porte dva bolo rovnaké ako žiadaná hodnota tohto napätia a teda $24 V$. Ustálený stav potom trval do času $t = 0,2 s$.
- V čase $t = 0,2 s$ sa zmenila žiadaná hodnota výstupného napätia portu dva z hodnoty $V_{P2} = 24 V$ na hodnotu $V_{P2} = 28,5 V$. Regulátor výstupného

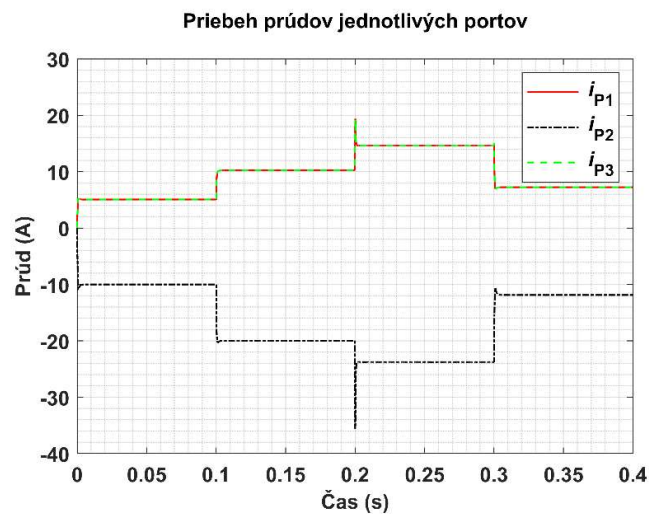
Na priebehoch na Obr. 15 a) a b) môžeme pozorovať, že regulátor pracoval správne v prípade oboch modelov a odchýlku výstupného napätia vyreguloval na požadovanú hodnotu aj pri zmene záťaže, aj pri zmene žiadanej hodnoty výstupného napätia. Na Obr. 15 c) je vidieť vyššie spomínaný rozdiel medzi spínaným modelom a redukovaným modelom, kedy spínaný model pracoval počas zaťaženia s väčším fázovým posunom ako redukovaný model. Z priebehov na Obr. 17 je zjavné, že podobne ako pri simulácii meniča bez regulátora, aj v tomto prípade je ovplyvnený priebeh prúdu a napätia medziobvodu zmenou fázového posunu striedača portu 2. Z priebehov na Obr. 17 b) c) môžeme sledovať, ako sa pri zväčšovaní fázového posunu zväčšuje aj prúd medziobvodu, rovnako tak na Obr. 17 e) f) a g) sú zjavné zmeny priebehu napätia medziobvodu pri zmene fázového posunu. Obr. 16 iba potvrdzuje vyššie popísanú teóriu, že prúd odoberaný portom 2 je zložený z prúdu dodávaného portom 1 a prúdu dodávaného portom 3.

V. MERANIE TROJPORTOVÉHO AKTÍVNEHO MOSTÍKA

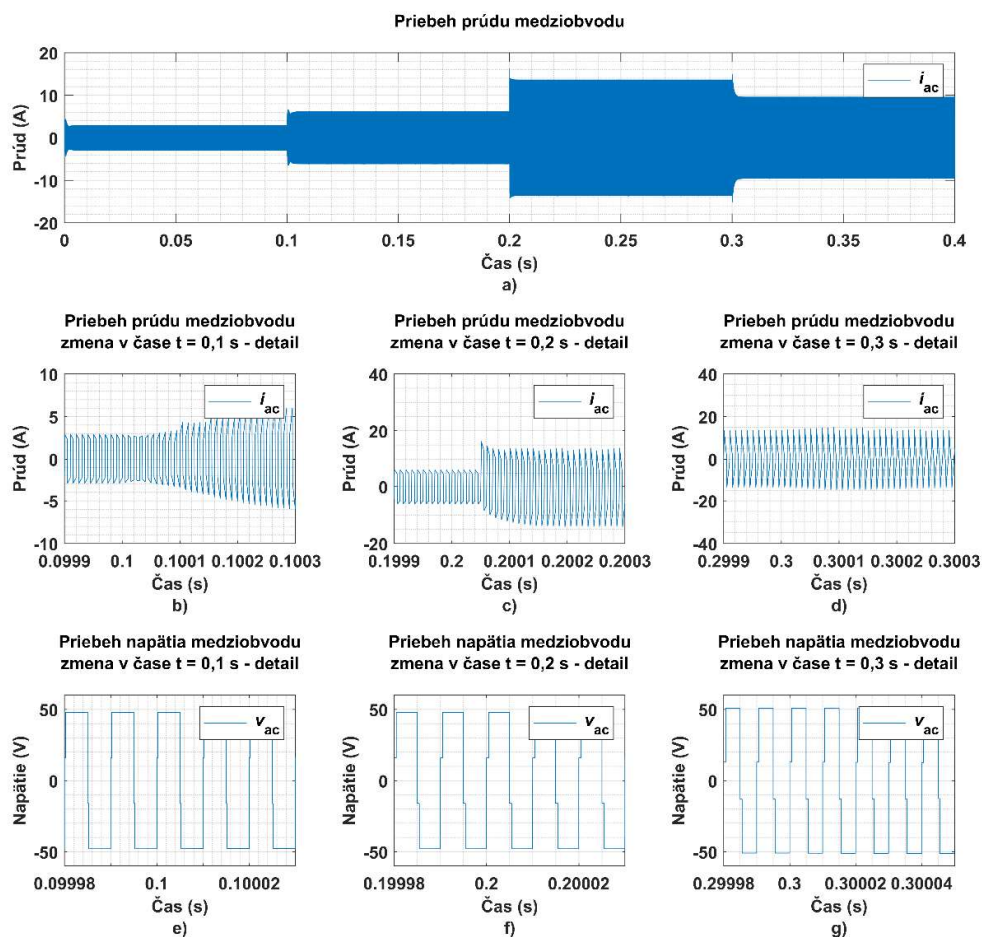
Prototyp trojportového aktívneho mostíka bol zostavený na základe striedačov dostupných na pracovisku KEM FEI TUKE a transformátorov, ktoré boli navrhnuté v rámci dizertačnej práce. Ide o trojfázové striedače pôvodne určené pre riadenie elektrických pohonov. Z týchto trojfázových striedačov boli použité iba dve fázy, ktoré generovali pravouhlé napätie pre primárne strany jednotlivých transformátorov. Striedače sú osadené digitálnym signálovým procesorom od spoločnosti Texas Instruments, konkrétne ide o model TMS320F28375D, ktorého výpočtový výkon a periférie sú plne dostupné používateľovi cez programovacie prostredie Code Composer Studio. Každý z týchto striedačov bolo potrebné naprogramovať tak, aby spĺňal požiadavky potrebné pre požadovaný prenos výkonu. Každý zo striedačov predstavuje nezávislý modul, ktorý s okolitým svetom komunikuje len v rámci periférií, ktoré mu nastaví programátor. Keďže sme od začiatku výskumu kládli dôraz na dodržanie vysokej modularity skúmanej topológie, chceli sme koncept *striedač + transformátor = port* zachovať. Pri takto zachovanom koncepte



Obr. 15 a) Priebeh žiadanej a skutočnej hodnoty napätia portu 2 pre simulačný a redukovaný model; b) priebeh prúdu portu 2 oboch modelov meniča; c) priebeh fázových posunov portu dva oboch modelov meniča



Obr. 16 Priebehy prúdov jednotlivých portov počas simulácie



Obr. 17 a) Priebeh prúdu medziobvodu počas simulácie; b) c) d) detail zmeny prúdu medziobvodu pri zmene záťaže resp. žiadanej hodnoty výstupného napätia; e) f) g) detail zmeny napätia medziobvodu pri zmene fázových posunov

je totiž možné prostredníctvom medziobvodu pripojiť teoreticky neobmedzené množstvo portov do topológie meniča. Výzvou však vždy bola výmena dát medzi jednotlivými strieďačmi a synchronizácia modulov PWM, najmä ak strieďač pracuje s vysokou spínacou frekvenciou, v našom prípade to je 100 kHz.

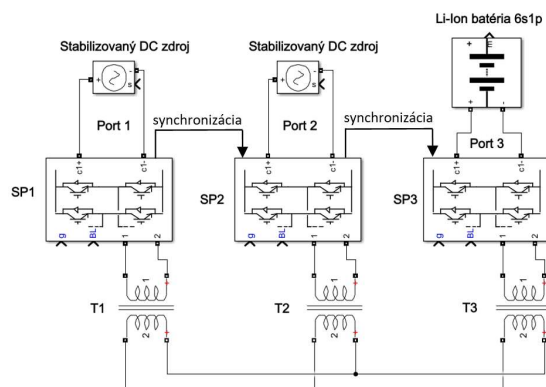
Takýto model riadenia, resp. takéto riadenie sa nazýva decentralizované, kedy každý zo strieďačov vykonáva svoj program, od ostatných zbiera dáta a v prípade nutnosti vykonáva akčný zásah. Pre takéto decentralizované systémy potom existujú rôzne na to prispôbené komunikačné zbernice bežiacie v reálnom čase, prostredníctvom ktorých je možné zdieľať aj synchronizačný signál hodín pre PWM periférie.

Označme strieďače podľa portov, pre prvý port to bude strieďač SP1, pre druhý port to bude strieďač SP2 a pre tretí port to bude strieďač SP3. Prvotná koncepcia distribúcie synchronizačného impulzu a výmeny dát počítala so zbernicou CAN, ktorej hardvérové rozhranie je dostupné na doskách strieďačov, avšak softvérovú výbavu bolo potrebné naprogramovať. Z dôvodu

krátkosti času a obtiažnosti implementácie CAN zbernice v reálnom čase sa od tejto myšlienky upustilo a synchronizačný signál bol medzi jednotlivými strieďačmi zdieľaný prostredníctvom štandardného vstupno/výstupného pinu, ktorý bol taktovaný frekvenciou 100 kHz priamo PWM perifériou referenčného strieďača. Keďže pri prvotných experimentoch išlo hlavne o potvrdenie teórie riadenia, nebol implementovaný ani regulátor výstupného napätia, a teda nebolo potrebné medzi strieďačmi zdieľať dáta. Každý strieďač však umožňuje pomocou tlačidiel na doske meniť fázový posun v oboch smeroch. Rozsah zmeny fázového posunu je však ohraničený, najmä kvôli ochrane zariadení pred vysokým prúdom cez polovodiče. Fázový posun jednotlivých mostíkov bol počas merania sledovaný osciloskopicky a podľa toho aj nastavený. Princiáľna schéma merania je zobrazená na Obr. 18.

Spustenie a synchronizácia všetkých portov trojportového meniča prebiehala v niekoľkých krokoch:

- Na striedač SP1 (port 1) sa priviedlo napätie stabilizovaného zdroja $V_{p1} = 24\text{ V}$, jeho transformátor však nebol budovaný ihneď. Po spustení programu v SP1 sa tlačidlom na doske striedača spustili impulzy pre budiče tranzistorov, striedač začal napájať transformátor portu 1. Prostredníctvom medziobvodu a ostatných transformátorov napájal aj striedače SP2 a SP3 cez nulové diódy ich tranzistorov.
- To znamená, že energia sa preniesla zo zdroja pripojeného na SP1 do zostávajúcich portov meniča a teda na port 2 a port 3. Takto bolo privedené napätie na port 2 a port 3 meniča a zároveň sa spustil program v DSP pre oba tieto striedače.
- Ak už striedač na porte 2 a striedač na porte 3 boli napájané prostredníctvom portu 1, bolo možné synchronizovať ich fázový posun. Hlavný synchronizačný impulz bol generovaný zo striedača portu 1, ktorý slúži aj ako referenčný. Synchronizácia striedačov na porte 2 a porte 3 bola vyriešená softvérovo, používateľ dal iba pokyn na synchronizáciu prostredníctvom tlačidla na doske striedača – platí pre všetky porty.
- Takto synchronizovaný trojportový menič však stále nemal na porte 2 a porte 3 pripojený žiadny zdroj, pretože bol napájaný prostredníctvom portu 1 a jeho zdroja. Po synchronizácii všetkých portov sme tak pristúpili k pripojeniu zdrojov k ostatným dvom portom P2 a P3. K portu 2 sme rovnako ako k portu 1 pripojili stabilizovaný zdroj s napätím $V_{p2} = 24\text{ V}$ a portu 3 sme pripojili Li-Ion batériu zloženú z článkov 18650 v konfigurácii 6s1p. Jednotlivé porty meniča je možné synchronizovať aj s pripojeným zdrojom energie, pri experimente sme si však chceli potvrdiť teóriu napájania ostatných striedačov prostredníctvom SP1.



Obr. 18 Principiálna schéma merania

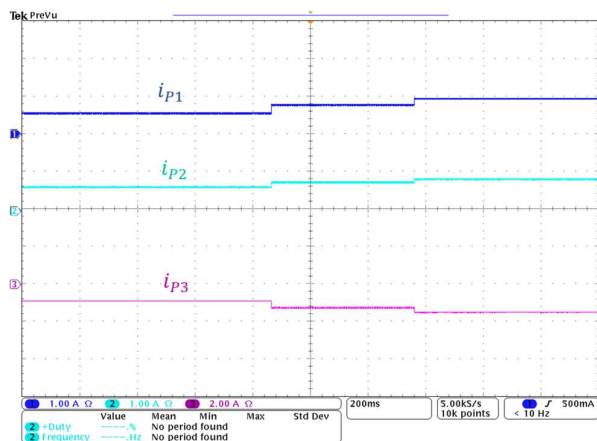
Samotný postup merania potom zodpovedal simulácii, kde dva porty predstavovali zdroj energie a tretí port predstavoval spotrebič. Ako bolo spomenuté v jednotlivých krokoch spustenia meniča, v prípade merania boli dva zdrojové porty napájané stabilizovaným nastaviteľným sieťovým zdrojom

s napätím nastaveným na hodnotu 24 V a k tretiemu záťažovému portu bola pripojená Li-Ion batéria v konfigurácii 6s1p, čo nám pri plnom nabití článkov umožňuje ísť do maximálneho napätia $25,2\text{ V}$ na svorkách batérie. Napätie batérie na začiatku merania bolo $23,4\text{ V}$ a teda pre prenos energie do batérie stačil aj menší fázový posun medzi SP3 a SP1 a SP2 kvôli nerovnosti vstupných napätí portov meniča. Hodnoty fázových posunov pre meranie neboli deterministicky dané, daný bol však najmenší možný krok fázového posunu. Hodnoty fázových posunov budú preto odčítané z oscilogramov.

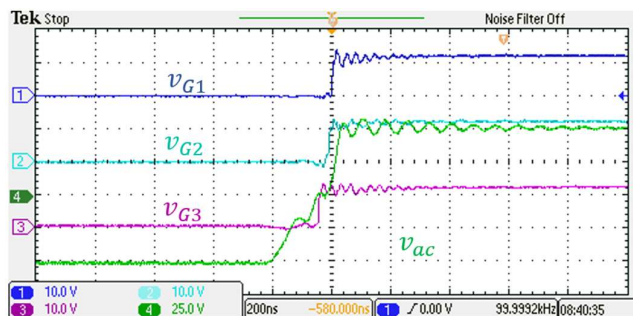
Ďalej budú popísané oscilogramy zaznamenané počas merania. Orientácia prúdov všetkých portov bola taká, že prúd vtekajúci do portu (od zdroja smerom k striedaču) je orientovaný kladne (kladné znamienko) a prúd vytekajúci z portu (od striedača do zdroja) má zápornú orientáciu (záporné znamienko). Fázové posuny jednotlivých striedačov boli osciloskopicky presledované na hradlách rovnakých tranzistorov v rámci mostíkového striedača, pretože priebehy merané hoci aj diferenciálnou sondou priamo na svorkách transformátorov obsahovali veľké rušenie, ktoré ovplyvnilo správne rozlíšenie počiatku priebehu. Prúdy jednotlivých portov sú označené ako i_{pk} , napätia na hradlách tranzistorov ako v_{Gk} , napätie v medziobvode ako v_{ac} a prúd v medziobvode ako i_{ac} , pričom $k = 1,2,3$.

Na Obr. 19 sú zobrazené priebehy prúdov všetkých portov počas dvojnásobnej zmeny fázového posunu na porte 3. Ak zoberieme do úvahy rozsah nastavené pre jednotlivé prúdy je zrejme, že prúd dodávaný do portu 3, označený ako i_{p3} je súčtom prúdov i_{p1} a i_{p2} , čo prakticky potvrdzuje teóriu prelievania energie v MAB meniči predstavenú v predošlých kapitolách. Na Obr. 20, Obr. 21 a Obr. 22 sú zobrazené priebehy napätí na hradlách rovnakých tranzistorov v rámci striedačov jednotlivých portov a tiež napätie medziobvodu. Fázové posuny korešpondujú so zmenami prúdov na Obr. 19, avšak kvôli potrebe merania veľkého počtu veličín nebolo možné zaznamenať ich priebehy na jeden oscilogram. Preto sú prúdy meniča a fázové posuny zaznamenané na rôznych osciloskopoch a zobrazené na rôznych obrázkoch. Na Obr. 20 je zobrazený prvý fázový posun tretieho portu nastavený počas merania. Z obrázku je zrejme, že striedač tretieho portu predbieha ostatné dva striedače, a tak vyvstáva otázka, či by port 3 nemal dodávať energiu do ostatných dvoch portov. To však v tomto prípade neplatí, pretože napätie batérie, ktorá je pripojená k portu 3 je menšie ako napätie stabilizovaných zdrojov na portoch 1 a 2. A keďže platí rovnica (21) o výkonových a teda aj prúdových pomeroch v MAB meniči, hoci fázovým posunom striedač portu 3 predbieha ostatné dva striedače, príspevok jeho energie do obvodu je stále menší, ako príspevok súčtu portu 1 a portu 2. Je to viditeľné aj na hodnotách prúdov na Obr. 19, kedy prúd tečúci do portu 3 je v časovom intervale prislúchajúcom k danému fázovému posunu najmenší. S ďalším zväčšovaním fázového posunu prúd tečúci do portu 3 rastie, čo je tiež vidieť aj na priebehu prúdu medziobvodu Obr. 23, Obr. 24 a Obr. 25. Ako bolo spomínané vyššie, rozsah fázového posunu bol počas experimentu

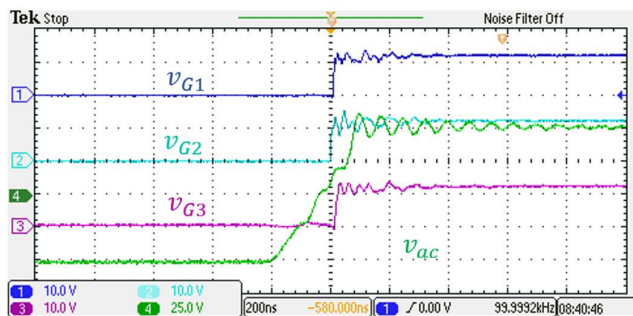
softvérovo obmedzený, aby nedošlo k nadprúdu v niektorej časti MAB meniča. Pri zmene fázového posunu sledujeme taktiež zmenu priebehu napätia medziobvodu v_{ac} . V tomto prípade sa taktiež potvrdzuje opísaná teória a to, že napätie medziobvodu je ovplyvňované fázovými posunmi takým spôsobom, že v jeho priebehu vznikajú „schody“. To vyjadruje, že napätie medziobvodu je zložené z napätí jednotlivých striedačov a pri fázovom posune medzi portami tieto „schody“ vznikajú. Veľmi dobre to je viditeľné na simulačných výsledkoch. Keďže v experimente je oveľa nižšia strmosť nárastu napätia, tieto „schody“ sú viditeľné menej.



Obr. 19 Priebehy prúdov jednotlivých portov počas merania

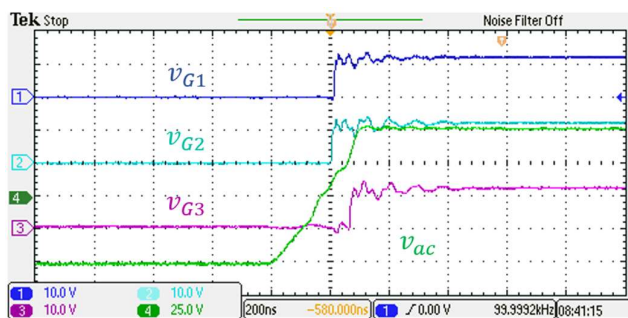


Obr. 20 Priebehy hradlových impulzov tranzistorov a priebehu napätia medziobvodu pre fázový posun $\phi_2 = -0,008$

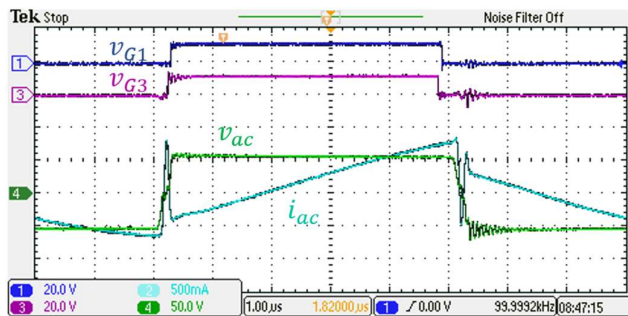


Obr. 21 Priebehy hradlových impulzov tranzistorov a priebehu napätia medziobvodu pre fázový posun $\phi_2 = 0,004$

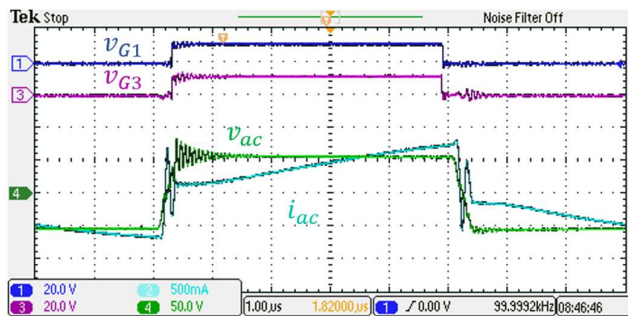
Na uvedených obrázkoch sú zobrazené priebehy napätí na hradlách tranzistorov portu 1 ako portu, ktorý dodáva energiu do meniča a portu 3 ako portu, ktorý energiu z meniča odoberá. Vzhľadom na ich fázové posuny boli preskúvané priebehy napätia a prúdu medziobvodu označované ako v_{ac} a i_{ac} . Obrázky sú zoradené tak, ako bol nastavovaný fázový posun na Obr. 19. Z priebehu prúdu medziobvodu i_{ac} je zrejmé, že zvyšovaním prúdu dodávaného do portu 3 sa mení priebeh prúdu medziobvodu z kvázi sinusového na pravouhlý. Na priebehu prúdu je taktiež vidieť výrazné kmity v mieste komutácie tranzistorov vo vetvách mostíkového striedača a tiež fázového posunu medzi portami meniča.



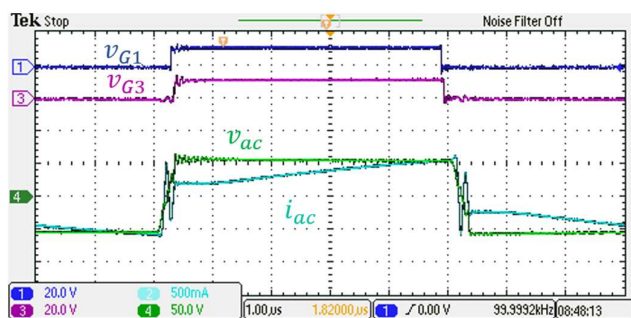
Obr. 22 Priebehy hradlových impulzov tranzistorov a priebehu napätia medziobvodu pre fázový posun $\phi_2 = 0,0016$



Obr. 23 Priebehy hradlových impulzov tranzistorov a prúdu a napätia medziobvodu pre fázový posun $\phi_2 = -0,008$



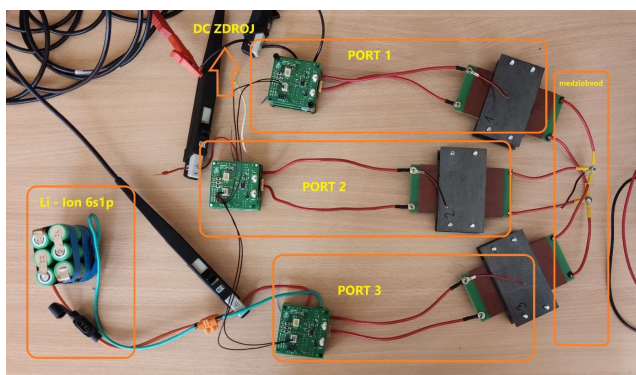
Obr. 24 Priebehy hradlových impulzov tranzistorov a prúdu a napätia medziobvodu pre fázový posun $\phi_2 = 0,004$



Experiment verifikoval uvedenú teóriu riadenia MAB meniča a poukázal na možnosti ďalšieho rozvoja problematiky napríklad pri overení podmienok pre spínanie v nule napätie (ZVS), prípadne pri skúmaní možností modifikácie topológie

Obr. 26 Priebehy hradlových impulzov tranzistorov a prúdu a napätia medziobvodu pre fázový posun $\phi_2 = 0,0016$

a taktiež pri skúmaní možností riadenie takýchto meničov s pomocou neurónových sietí a umelej inteligencie. Problematika MAB meničov tiež zahŕňa skúmanie možností decentralizovaného riadenia, výmeny dát a synchronizačného impulzu jednotlivých portov MAB meniča prostredníctvom komunikačných rozhraní pracujúcich v reálnom čase, ako napríklad CAN, EtherCAT, atď. Na Obr. 26 je fotka prototypu trojportového aktívneho mostíka, na ktorom bolo meranie vykonané.



Obr. 25 Prototyp trojportového aktívneho mostíka zostaveného pre experimentálne meranie

ZÁVER

Príspevok opisuje topológiu, modelovanie a návrh regulátorov MAB meniča kde veličiny v obvode meniča nie sú harmonické. Značná väčšina matematických popisov takýchto meničov vychádza z DAB meniča, ktorý predstavuje najjednoduchší model viacportového meniča – dvojportový menič. Uvedené sú postupy pre zhotovenie redukovaného modelu a tiež regulátora výstupného napätia, ktorý je možné na takýto typ meniča v určitých podmienkach nasadiť.

Na základe matematického popisu DAB meniča bol potom opísaný aj viacportový aktívny mostík resp. menič, označený aj ako MAB, pre ktorý boli zadefinované základné rovnice.

Z týchto rovníc bolo zrejmé, že tok výkonu v takomto type meniča je riadený na základe fázových posunov jednotlivých mostíkov a že väzby medzi jednotlivými portami MAB meniča sú nelineárne. Tieto nelinearity boli analyzované a v práci sú uvedené možné riešenia odväzbenia portov tak, aby bolo možné riadiť tok výkonu pre každý port meniča nezávisle. Pre MAB menič je tiež zhotovený model redukovaného rádu a navrhnutý spätnoväzobný regulátor výstupného napätia. Takýto MAB menič je potom simulačne verifikovaný v softvéri PLECS Blockset najprv bez regulátora a potom aj s regulátorom výstupného napätia.

Experimentálne meranie prebehlo na prototyp meniča, kde jeden jeho port, resp. modul sa skladal z transformátora a z mostíkového striedača dostupného na pracovisku KEM FEI TUKE. Meranie prebehlo v režime, kedy dva porty meniča boli pripojené na stabilizovaný zdroj a energiu dodávali a tretí port bol pripojený na Li-Ion batériu v konfigurácii 6s1p a energiu odoberal. Experiment potvrdil teoretický základ pre prelievanie energie medzi portami a ukázal možnosti ďalšieho rozvoja topológie meniča a topológie komunikačnej siete pre decentralizované riadenie takýchto meničov.

Nie kompletne preskúmanou oblasťou ostáva oblasť možností odväzbenia portov a teda odstránenia nelinearit z obvodu MAB meniča. Existuje niekoľko ďalších spôsobov, okrem tých predstavených v práci, avšak je nutné ich aspoň simulačne verifikovať a zhodnotiť výpočtovú náročnosť daného algoritmu a možnosti jeho implementácie v rámci dostupných hardvérových riešení.

REFERENCIE

- [1] H. Mahmood, D. Michaelson and J. Jiang, "A Power Management Strategy for PV/Battery Hybrid Systems in Islanded Microgrids," in IEEE Journal of Emerging and Selected Topics in Power Electronics, vol. 2, no. 4, pp. 870-882, Dec. 2014.
- [2] B. Wang, M. Sechilariu and F. Locment, "Intelligent DC Microgrid With Smart Grid Communications: Control Strategy Consideration and Design," in IEEE Transactions on Smart Grid, vol. 3, no. 4, pp. 2148-2156, Dec. 2012.
- [3] H. Zhou, T. Bhattacharya, D. Tran, T. S. T. Siew and A. M. Khambadkone, "Composite Energy Storage System Involving Battery and Ultracapacitor With Dynamic Energy Management in Microgrid Applications," in IEEE Transactions on Power Electronics, vol. 26, no. 3, pp. 923-930, March 2011.
- [4] K. Sun, L. Zhang, Y. Xing and J. M. Guerrero, "A Distributed Control Strategy Based on DC Bus Signaling for Modular Photovoltaic Generation Systems With Battery Energy Storage," in IEEE Transactions on Power Electronics, vol. 26, no. 10, pp. 3032-3045, Oct. 2011.
- [5] Hongfei Wu, Yan Xing, Yanbing Xia and Kai Sun, "A family of non-isolated three-port converters for stand-alone renewable power system," IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society, Melbourne, VIC, 2011, pp. 1030-1035.
- [6] M. Amirabadi, H. A. Toliyat and W. C. Alexander, "A Multiport AC Link PV Inverter With Reduced Size and Weight for Stand-Alone Application," in IEEE Transactions on Industry Applications, vol. 49, no. 5, pp. 2217-2228, Sept.-Oct. 2013.
- [7] H. Kim, B. Parkhideh, T. D. Bongers and H. Gao, "Reconfigurable Solar Converter: A Single-Stage Power Conversion PV-Battery System," in IEEE Transactions on Power Electronics, vol. 28, no. 8, pp. 3788-3797, Aug. 2013.
- [8] Z. Qian, O. Abdel-Rahman, H. Al-Atrash and I. Batarseh, "Modeling and Control of Three-Port DC/DC Converter Interface for Satellite

- Applications," in *IEEE Transactions on Power Electronics*, vol. 25, no. 3, pp. 637-649, March 2010.
- [9] H. Wu, J. Zhang, X. Qin, T. Mu and Y. Xing, "Secondary-Side Regulated Soft-Switching Full-Bridge Three-Port Converter Based on Bridgeless Boost Rectifier and Bidirectional Converter for Multiple Energy Interface," in *IEEE Transactions on Power Electronics*, vol. 31, no. 7, pp. 4847-4860, July 2016.
- [10] Dalton Honorio, Luiz Henrique S. C. Barreto, Andrew Foote, Burak Ozpineci, João Onofre Pereira Pinto, "Modular transformer in isolated multiport power converters" in *IEEE Southern Power Electronics Conference (SPEC)*, Puerto Varas, 4-7 Dec. 2017.
- [11] Ferroxcube, „ Design of planar power transformers – application note“, online.
- [12] J. A. Oliver *et al.*, "High level decision methodology for the selection of a fuel cell based power distribution architecture for an aircraft application," *2009 IEEE Energy Conversion Congress and Exposition*, San Jose, CA, USA, 2009, pp. 459-464, doi: 10.1109/ECCE.2009.5316145.
- [13] E. Cilio, G. Mitchell, M. Schupbach and A. Lostetter, "SiC intelligent multi module DC/DC converter system for space applications," *2009 IEEE Aerospace conference*, Big Sky, MT, USA, 2009, pp. 1-19, doi: 10.1109/AERO.2009.4839520.
- [14] B. Reese, M. Schupbach, A. Lostetter, B. Rowden, R. Saunders and J. Balda, "High voltage, high power density bi-directional multi-level converters utilizing silicon and silicon carbide (SiC) switches," *2008 Twenty-Third Annual IEEE Applied Power Electronics Conference and Exposition*, Austin, TX, USA, 2008, pp. 252-258, doi: 10.1109/APEC.2008.4522730.
- [15] K. Siri, M. Willhoff, K. A. Conner and D. Q. Tran, "High-voltage-input, low-voltage-output, series-connected converters with uniform voltage distribution," *2009 IEEE Aerospace conference*, Big Sky, MT, USA, 2009, pp. 1-9, doi: 10.1109/AERO.2009.4839519.
- [16] H. Tao, A. Kotsopoulos, J. L. Duarte and M. A. M. Hendrix, "Transformer-Coupled Multiport ZVS Bidirectional DC-DC Converter With Wide Input Range," in *IEEE Transactions on Power Electronics*, vol. 23, no. 2, pp. 771-781, March 2008, doi: 10.1109/TPEL.2007.915129.
- [17] S. Falcones, R. Ayyanar and X. Mao, "A DC-DC Multiport-Converter-Based Solid-State Transformer Integrating Distributed Generation and Storage," in *IEEE Transactions on Power Electronics*, vol. 28, no. 5, pp. 2192-2203, May 2013, doi: 10.1109/TPEL.2012.2215965.
- [18] A. M. Ari, L. Li and O. Wasynczuk, "Modeling and analysis of N-port DC-DC converters using the cyclic average current," *2012 Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*, Orlando, FL, USA, 2012, pp. 863-869, doi: 10.1109/APEC.2012.6165920.
- [19] A. M. Ari, L. Li and O. Wasynczuk, "Modeling and analysis of N-port DC-DC converters," *2012 IEEE Power and Energy Conference at Illinois*, Champaign, IL, USA, 2012, pp. 1-7, doi: 10.1109/PEEL.2012.6184611.
- [20] R. W. Erickson and D. Maksimovic, *Fundamentals of Power Electronics*. New York, NY, USA: Springer, 2007.
- [21] S. Shao, H. Chen, X. Wu, J. Zhang, and K. Sheng, "Circulating current and ZVS-on of a dual active bridge DC-DC converter: A review," *IEEE Access*, vol. 7, pp. 50 561–50 572, 2019.
- [22] A. R. Alonso, J. Sebastian, D. G. Lamar, M. M. Hernando, and A. Vazquez, "An overall study of a dual active bridge for bidirectional DC/DC conversion," in *Proc. IEEE Energy Convers. Congr. Expo.*, Sep. 2010, pp. 1129–1135.
- [23] H. Bai, C. Mi, C. Wang, and S. Gargies, "The dynamic model and hybrid phase-shift control of a dual-active-bridge converter," in *Proc. 34th Annu. Conf. IEEE Ind. Electron.*, Nov. 2008, pp. 2840–2845.
- [24] H. Bai, Z. Nie, and C. C. Mi, "Experimental comparison of traditional phase-shift, dual-phase-shift, and model-based control of isolated bidirectional DC-DC converters," *IEEE Trans. Power Electron.*, vol. 25, no. 6, pp. 1444–1449, Jun. 2010
- [25] R. W. A. A. D. Doncker, D. M. Divan, and M. H. Kheraluwala, "A three-phase soft-switched high-power-density DC/DC converter for high-power applications," *IEEE Trans. Ind. Appl.*, vol. 27, no. 1, pp. 63–73, Jan. 1991.
- [26] L. Chen, L. Lin, S. Shao, F. Gao, Z. Wang, P. W. Wheeler, and T. Dragičević, "Moving discretized control set model-predictive control for dual-active bridge with the triple-phase shift," *IEEE Trans. Power Electron.*, vol. 35, no. 8, pp. 8624–8637, Aug. 2020.
- [27] Zhang, Z. Shan, and J. Jatskevich, "Large- and small-signal average value modeling of dual-active-bridge DC-DC converter considering power losses," *IEEE Trans. Power Electron.*, vol. 32, no. 3, pp. 1964–1974, Mar. 2017.
- [28] M. N. Kheraluwala, R. W. Gascoigne, D. M. Divan, and E. D. Baumann, "Performance characterization of a high-power dual active bridge DC-to-DC converter," *IEEE Trans. Ind. Appl.*, vol. 28, no. 6, pp. 1294–1301, Nov. 1992.
- [29] H. Qin and J. W. Kimball, "Closed-loop control of DC-DC dual-active bridge converters driving single-phase inverters," *IEEE Trans. Power Electron.*, vol. 29, no. 2, pp. 1006–1017, Feb. 2014
- [30] H. Tao, A. Kotsopoulos, J. L. Duarte, and M. A. M. Hendrix, "Transformer-coupled multiport ZVS bidirectional DC-DC converter with wide input range," *IEEE Trans. Power Electron.*, vol. 23, no. 2, pp. 771–781, Mar. 2008.
- [31] A. M. Ari, L. Li, and O. Wasynczuk, "Control and optimization of N-port DC-DC converters," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 4, pp. 1521–1528, Jul. 2016.
- [32] S. Falcones and R. Ayyanar, "LQR control of a quad-active-bridge converter for renewable integration," in *Proc. IEEE Ecuador Tech. Chapters Meeting (ETCM)*, Oct. 2016, pp. 1–6.
- [33] C. Gu, Z. Zheng, L. Xu, K. Wang, and Y. Li, "Modeling and control of a multiport power electronic transformer (PET) for electric traction applications," *IEEE Trans. Power Electron.*, vol. 31, no. 2, pp. 915–927, Feb. 2016.
- [34] P. Wang and M. Chen, "Towards power FPGA: Architecture, modeling and control of multiport power converters," in *Proc. IEEE 19th Workshop Control Modeling for Power Electron. (COMPEL)*, Jun. 2018, pp. 1–8
- [35] M. Neubert, S. P. Engel, J. Gottschlich, and R. W. De Doncker, "Dynamic power control of three-phase multiport active bridge DC-DC converters for interconnection of future DC-grids," in *Proc. IEEE 12th Int. Conf. Power Electron. Drive Syst. (PEDS)*, Dec. 2017, pp. 639–646.
- [36] P. Zumel, C. Fernandez, A. Lazaro, M. Sanz, and A. Barrado, "Overall analysis of a modular multi active bridge converter," in *Proc. IEEE 15th Workshop Control Modeling for Power Electron. (COMPEL)*, Jun. 2014, pp. 1–9
- [37] S. Falcones, R. Ayyanar, and X. Mao, "A DC-DC multiport-converter-based solid-state transformer integrating distributed generation and storage," *IEEE Trans. Power Electron.*, vol. 28, no. 5, pp. 2192–2203, May 2013.
- [38] P. Wang, Y. Chen, Y. Elasser, and M. Chen, "Small signal model for very-large-scale multi-active-bridge differential power processing (MAB-DPP) architecture," in *Proc. 20th Workshop Control Modeling for Power Electron. (COMPEL)*, Jun. 2019, pp. 1–8
- [39] N. S. R. Jakka, A. Shukla, and G. D. Demetriades, "Dual-transformer-based asymmetrical triple-port active bridge (DT-ATAB) isolated DC-DC converter," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 4549–4560, Jun. 2017.
- [40] B. J. D. Vermulst, J. L. Duarte, E. A. Lomonova, and K. G. E. Wijnands, "Scalable multi-port active-bridge converters: Modelling and optimised control," *IET Power Electron.*, vol. 10, no. 1, pp. 80–91, Jan. 2017
- [41] X. Ruan, W. Chen, L. Cheng, C. K. Tse, H. Yan, and T. Zhang, "Control strategy for input-series-output-parallel converters," *IEEE Trans. Ind. Electron.*, vol. 56, no. 4, pp. 1174–1185, Apr. 2009.
- [42] P. Zumel *et al.*, "Modular dual-active bridge converter architecture," *IEEE Trans. Ind. Appl.*, vol. 52, no. 3, pp. 2444–2455, May 2016.
- [43] K. Filsoof and P. W. Lehn, "A bidirectional multiple-input multiple-output modular multilevel DC-DC converter and its control design," *IEEE Trans. Power Electron.*, vol. 31, no. 4, pp. 2767–2779, Apr. 2016.
- [44] A. Rodriguez, A. Vazquez, D. G. Lamar, M. M. Hernando, and J. Sebastian, "Different Purpose Design Strategies and Techniques to Improve the Performance of a Dual Active Bridge With Phase-Shift

- Control,” IEEE Transactions on Power Electronics, vol. 30, pp. 790–804, Feb. 2015.
- [45] P. Wang and M. Chen, “Towards power FPGA: Architecture, modeling and control of multiport power converters,” in Proc. IEEE 19th Workshop Control Modeling for Power Electron. (COMPEL), Jun. 2018, pp. 1–8.
- [46] M. Neubert, S. P. Engel, J. Gottschlich, and R. W. De Doncker, “Dynamic power control of three-phase multiport active bridge DC-DC converters for interconnection of future DC-grids,” in Proc. IEEE 12th Int. Conf. Power Electron. Drive Syst. (PEDS), Dec. 2017, pp. 639–646.
- [47] I. N. Bronstein, K. A. Semendyayev, G. Musiol, and H. Mühlig, Handbook of Mathematics, 4th ed. Springer, 2004.
- [48] M. Biroš, „Manažment energie v mobilných zariadeniach“, Dizertačná práca, KEM FEI TUKE, 2018, Dostupné online na webe CRZP.
- [49] J. D. Glover, M. S. Sarma, and T. J. Overbye, Power System Analysis and Design, 4th ed. Toronto, ON, Canada: Thomson, 2008
- [50] R. D. Zimmerman, C. E. Murillo-Sanchez and R. J. Thomas, “MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education,” IEEE Transactions on Power Systems, vol. 26, no. 1, pp. 12-19, Feb. 2011.
- [51] Cayley, A. XXVIII. “On the theory of the analytical forms called trees”, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 13(85), pp. 172-176, (1857).
- [52] Demiroğlu, Uğur & Şenol, Bilal. (2018). Analytical Design of PI Controllers for First Order plus Time Delay Systems. Pamukkale University Journal of Engineering Sciences. 26. 40-47. 10.5505/pajes.2020.19616.
- [53] ECE 486: Control Systems, Lecture 9A: PI Tuning for First-Order Systems, dostupné online na: https://courses.engr.illinois.edu/ece486/fa2021/documentation/lectures/slides/Lecture9A_PITuningForFirstOrderSystems.pdf, 03.06.2023

Bezsnímačové riadenie synchronného motora s permanentnými magnetmi pre nízke rýchlosti

Viktor Petro

Katedra elektrotechniky a mechatroniky
Fakulta elektrotechniky a informatiky,
Technická univerzita v Košiciach, Slovensko
e-mail: viktor.petro@tuke.sk

Karol Kyslan

Katedra elektrotechniky a mechatroniky
Fakulta elektrotechniky a informatiky,
Technická univerzita v Košiciach, Slovensko
e-mail: karol.kyslan@tuke.sk

Abstrakt— Tento článok prezentuje hlavné aspekty dizertačnej práce, ktorá je venovaná analýze a návrhu regulačných algoritmov pre bezsnímačové riadenie synchronného motora s permanentnými magnetmi. Zameraná je na algoritmy vhodné pre riadenia v oblasti nízkych rýchlostí. Najčastejšie sa využívajú metódy založené na injekcii vysokofrekvenčných impulzov. Tieto impulzy môžu byť superponované na základný napäťový vektor slúžiaci pre reguláciu momentu, alebo môžu byť injektované v čase prerušenia riadenia motora. V práci je uvedený teoretický rozbor vybraných existujúcich metód bezsnímačového riadenia ako aj návrh nových, doposiaľ nepublikovaných prístupov. Prínosom sú simulačné výsledky v prostredí MATLAB/Simulink a experimentálne výsledky bezsnímačového riadenia motora. V článku sú prezentované dve doposiaľ nepublikované metódy, ktoré vznikli vylepšením existujúcich metód.

Keywords—bezsnímačové riadenie, synchronný motor s permanentnými magnetmi, vysokofrekvenčná injekcia

Abstract—This article presents the main aspects of the dissertation thesis, which focuses on sensorless control of permanent magnet synchronous motor in low-speed region. It contains approaches, which are used for motor control at low and zero-speed regions. Commonly, high-frequency signal injection is utilized for sensorless control. These high-frequency impulses can be superimposed on the fundamental voltage vector used for motor control, or they can be injected while the motor control is interrupted. The thesis describes the theoretical background of existing methods for sensorless control as well as presents yet unpublished approaches. The main results are simulation verification using MATLAB/Simulink and experimental verification of the sensorless motor control methods. This article contains yet unpublished approaches, created by the improvement of already existing methods.

Keywords—sensorless control, permanent magnet synchronous motor, high-frequency injection

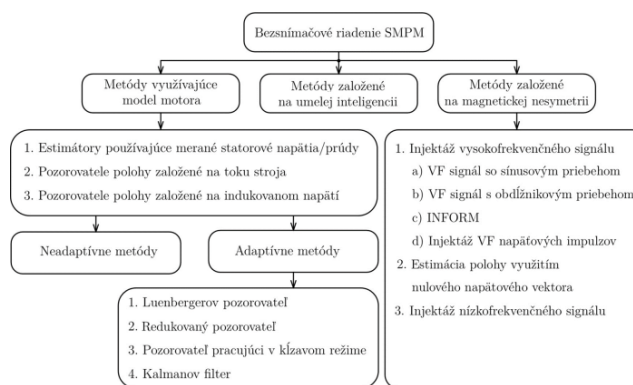
I. ÚVOD

Synchronný motor s permanentnými magnetmi (SMPM) je bežnou súčasťou náročných priemyselných aplikácií, pri ktorých sa od pohonu vyžadujú rýchle dynamické charakteristiky, vysoká účinnosť a bezúdržbovosť. Regulácia motora je zabezpečená vektorovým riadením, ktoré zabezpečí optimálny chod motora v celej pracovnej oblasti.

Tento spôsob riadenia využíva transformáciu statorovej súradnicovej sústavy do rotorovej, kde sa zvlášť reguluje momentovotvorná a tokotvorná zložka prúdu. Pre účel transformácie súradnicových sústav je nevyhnutná presná informácia o polohe rotora voči statorovému vinutiu stroja. Najčastejšie sa používa inkrementálny snímač, snímač typu resolver, respektíve snímače založené na Hallovom princípe. Snímač polohy rotora však znižuje celkovú spoľahlivosť systému, zvyšuje náklady na pohon a zaberá priestor. Preto eliminácia snímača mechanických veličín je aktuálna téma výskumu a metódy riadenia ktoré to umožňujú sú označované ako bezsnímačové.

Techniky bezsnímačového riadenia sa rozdeľujú v závislosti od rozsahu pracovných rýchlostí. Pre stredné a vysoké rýchlosti sú použité estimátory založené indukovanom napätí v motore. V prípade nízkych, resp. nulových otáčok je amplitúda indukovaného napätia nízka, a len ťažko rozoznateľná od šumu. Preto túto oblasť sa využívajú estimátory založené na magnetickej nesymetrii (nesymetrii) stroja. Základné rozdelenie bezsnímačových metód je uvedené na Obr. 1.

V článku prezentujeme prehľad metód bezsnímačového riadenia pre oblasť nízkych rýchlostí, a tiež návrhy na vylepšenie už existujúcich metód. Jedna z vylepšených metód bola aj experimentálne overená.



Obr. 1 Základné rozdelenie bezsnímačových metód

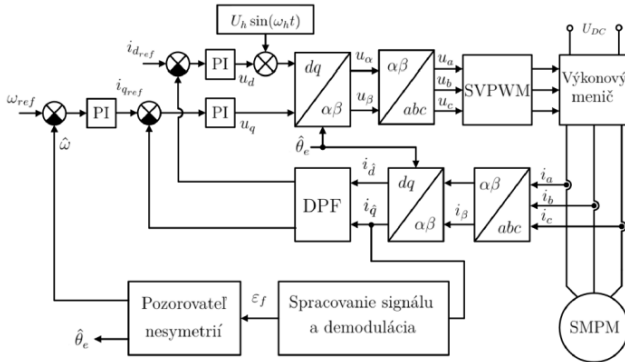
II. METÓDY ZALOŽENÉ NA MAGNETICKEJ NESYMETRII MOTORA

A. Injektáž VF signálu so sínusovým priebehom

Napäťový vektor rotujúci s vysokou frekvenciou (väčšinou v rozsahu 500 Hz až 2.5 kHz) je nasuperponovaný na základný napäťový vektor, ktorý je výstupom z regulačného procesu. Dôsledkom VF zložky napätia je vytvorenie VF zložky prúdu, ktorá je superponovaná na základný vektor prúdu. Táto VF zložka prúdu obsahuje informáciu o polohe rotora, ktorá môže byť získaná vhodným spracovaním a demoduláciou signálu. Matematický zápis injektovaného signálu vyzerať nasledovne [1] [2]:

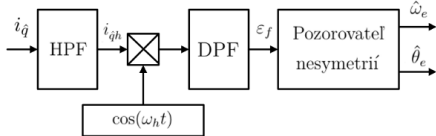
$$\begin{bmatrix} u_{dh} \\ u_{qh} \end{bmatrix} = \begin{bmatrix} U_h \sin(\omega_h t) \\ 0 \end{bmatrix}, \quad (1)$$

kde u_{dh} a u_{qh} predstavujú VF zložky injektované do pozorovanej \hat{d} a \hat{q} osi, U_h je amplitúda injektovaného napätia a ω_h predstavuje uhlovú frekvenciu VF napätia. Bloková schéma je zobrazená na Obr. 2.



Obr. 2 Bloková schéma bezsúťažného riadenia SMPM použitím injektáže sínusového VF signálu.

Injektované VF napätie vytvára vo fázach motora VF prúdové odozvy. Predtým, ako sa namerané prúdy použijú v regulačnej schéme, je nutné ich vyfiltrovať dolnopriepustným (DPF) filtrom, ktorý odstráni VF zložky. Nevyfiltrovaný prúd sa použije v bloku Spracovanie signálu a demodulácia. Tento proces je detailne uvedený na Obr. 3.



Obr. 3 Spracovanie a demodulácia signálu pre pozorovanie polohy a uhlovej rýchlosti.

V procese demodulácie sa pracuje len s VF zložkou prúdu i_{qh} , ktorú je možné získať hornopriepustným filtrom (HPF) alebo pásmovou priepustňou (PP). Následne je použitý DPF

a získa sa signál ε_f , ktorý je úmerný chybe pozorovanej polohy. Vyfiltrovaný signál ε_f vstupuje do pozorovateľa nesymetrií, ktorý má za úlohu regulovať tento signál na nulovú hodnotu, čím sa docielu nulový rozdiel medzi pozorovanou a skutočnou rotorovou súradnicovou sústavou. Pozorovateľ nesymetrií pozostáva z PI regulátora s veľkým pásmom priepustnosti.

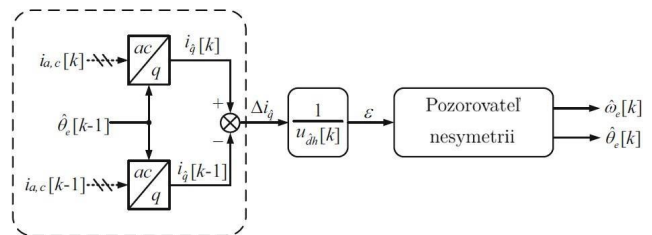
Nevýhoda metód založených na injektovaní signálu so sínusovým priebehom je obmedzená frekvencia injektovaného napätia, keďže musí byť zachovaný tvar sinusoidy. Zvýšením frekvencie injektovaného signálu by sa zlepšili dynamické vlastnosti pozorovania, avšak v praktických aplikáciách je frekvencia šírko-impulzovej modulácie (PWM - Phase Width Modulation) obmedzená.

B. Injektáž VF signálu s obdĺžnikovým priebehom

Pre vylepšenie dynamických vlastností pozorovania, ako aj celkového bezsúťažného riadenia, bola navrhnutá injektáž VF signálu s obdĺžnikovým priebehom. Podobne ako pri injektáži signálu so sínusovým priebehom, aj v tomto prípade sa injektáž vykonáva v rotorovom súradnicovom systéme. Výhoda spočíva v tom, že frekvencia injektovaného signálu môže byť rovnaká ako spínacia frekvencia PWM. Pretože frekvencia injektovaného signálu je oveľa väčšia ako šírka pásma priepustnosti regulátora prúdov, v prúdovej slučke nie je nutné použiť DPF. Injektované napätie s obdĺžnikovým priebehom je zadané ako [3]:

$$\begin{bmatrix} u_{dh} \\ u_{qh} \end{bmatrix} = \begin{bmatrix} U_h (-1)^k \\ 0 \end{bmatrix} \quad (2)$$

kde U_h je amplitúda injektovaného napätia v k -tom výpočtovom kroku. Bloková schéma je takmer identická ako na Obr. 2, ale injektovaný signál má podobu z rovnice (2). Proces získavania signálu ε_f , ktorý je úmerný chybe pozorovanej polohy je zobrazený na Obr. 4.



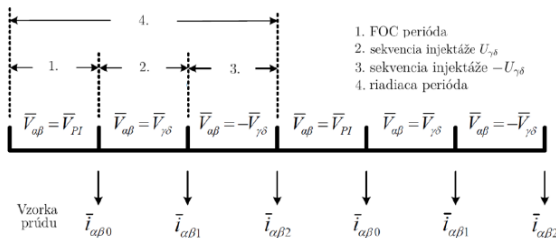
Obr. 4 Pozorovanie polohy rotora a uhlovej rýchlosti [4].

Doteraz spomenuté metódy sú založené na superponovaní vysokofrekvenčného signálu na signál so základnou frekvenciou. Pre získanie VF zložiek prúdu je väčšinou použitý HPF alebo PP. V procese extrakcie polohy z VF zložiek prúdu je často používaný dolnopriepustný filter. Používanie filtrov

zmenšuje pásmo priepustnosti regulátorov a zhoršuje dynamické vlastnosti pohonu s bezsnímačovým riadením. Okrem toho, na presnosť pozorovanej polohy môžu nepriaznivo vplyvať aj úbytky napätí, ktoré sú spôsobené nelinearitou meniča, úbytkami napätí na spínacích prvkoch a na odpore vinutí stroja, a pod. Spomenuté nedostatky metód so superponovaním VF signálu na základnú frekvenciu sa dajú do určitej miery odstrániť injektážou napät'ových impulzov, ktorá bude uvedená v nasledujúcej časti.

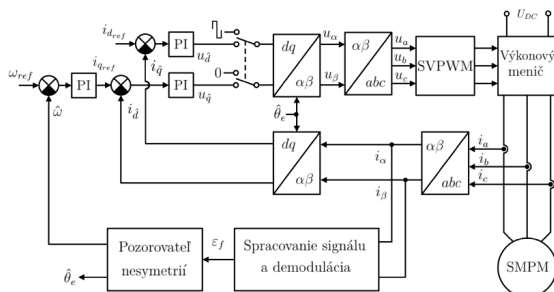
C. Injektáž VF napät'ových impulzov

Pri tejto metóde dochádza ku odpojeniu motora od riadiacej štruktúry počas injektáže VF signálu. Sekvencia injektáže je zobrazená na Obr. 5. Riadiaca perióda pozostáva z troch PWM spínacích periód. V prvej spínacej perióde sa vykonáva vektorové riadenie (FOC – Field Oriented Control). V druhej spínacej perióde sa vykonáva injektáž kladného VF napätia. V tretej spínacej perióde sa vykonáva injektáž záporného VF napätia. Ďalej platí, že prúdy sú snímané na začiatku každej spínacej (vzorkovacej) periódy. Vektorové riadenie zohľadňuje len prúd nameraný na začiatku periódy, v ktorej sa vykonáva samotné riadenie. To zabezpečí elimináciu digitálnych filtrov a v prúdovej slučke a zvýši odolnosť voči úbytkom napätia bez nutnosti použiť akékoľvek kompenzácie [5].



Obr. 5 Sekvencie injektáže napät'ových impulzov do pozorovanej rotorovej súradnicovej sústavy

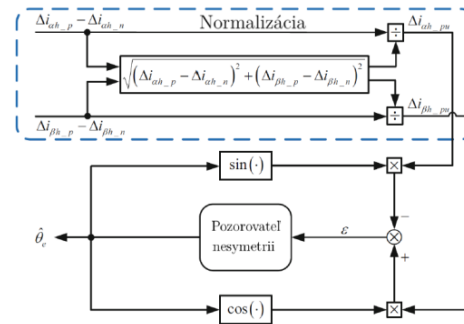
Bloková schéma bezsnímačového riadenia použitím injektáže VF napät'ových impulzov je zobrazená na Obr. 6.



Obr. 6 Bloková schéma bezsnímačového riadenia použitím injektáže VF napät'ových impulzov

V blokovej schéme na Obr. 6 je doplnený prepínač, ktorý zvyrazňuje, že motor je buď pripojený k regulačnej štruktúre (dolná poloha spínača), alebo je motor pripojený ku VF injektáži (horná poloha spínača). DPF v prúdovej slučke nie je potrebný, pretože prúd počas riadiacej periódy neobsahuje VF

zložky. To isté platí pre injektážne periódy, digitálne filtre sa nepoužívajú, pretože nameraný prúd je odozvou len VF napätia. Spracovanie VF prúdov a pozorovanie polohy a uhlovej rýchlosti je zobrazené na Obr. 7.



Obr. 7 Normalizácia a proces výpočtu pozorovanej polohy a uhlovej rýchlosti [6]

Pre prúdové derivácie v $\alpha\beta$ statorovej súradnicovej sústave platí:

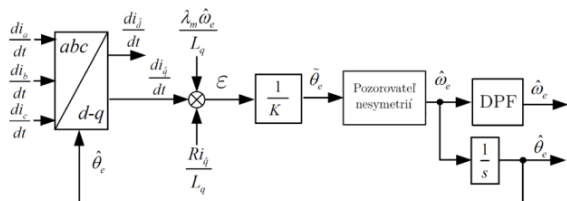
$$\begin{bmatrix} \Delta i_{\alpha h_p} - \Delta i_{\alpha h_n} \\ \Delta i_{\beta h_p} - \Delta i_{\beta h_n} \end{bmatrix} = \begin{bmatrix} 2i_{\alpha h1} - i_{\alpha h0} - i_{\alpha h2} \\ 2i_{\beta h1} - i_{\beta h0} - i_{\beta h2} \end{bmatrix} = 2K \begin{bmatrix} \Delta i_{\alpha h_{pu}} \\ \Delta i_{\beta h_{pu}} \end{bmatrix} \quad (3)$$

Tieto prúdové derivácie sú normalizované a následne transformované do rotorovej súradnicovej sústavy použitím pozorovanej polohy.

D. Estimácia polohy využitím nulového napät'ového vektora

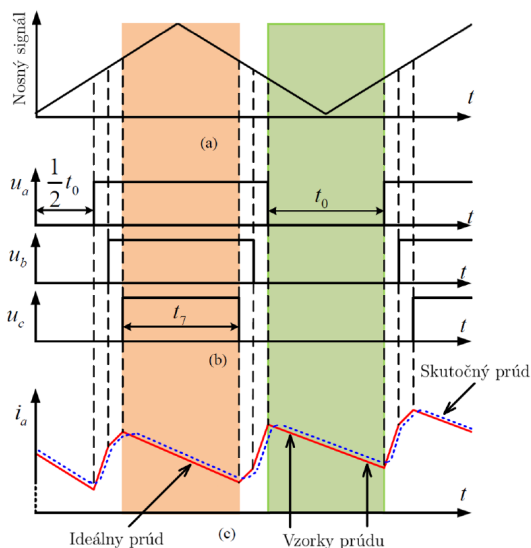
Všetky doposiaľ opísané metódy založené na injektáži rôznych typov VF signálov sa pri bezsnímačovom riadení SMPM využívajú veľmi často. Pre niektoré aplikácie však nežiaduce dva javy, ktoré pri injektáži vznikajú: akustické rušenie a momentové rázy. Pre potlačenie akustického šumu boli navrhnuté injektáže VF signálov so pseudonáhodnou frekvenciou. Injektovaním VF signálu s premenlivou frekvenciou sa docielu zníženie akustického rušenia na relatívne nízku hodnotu. Avšak neodstránia sa momentové rázy vytvorené injektážou VF napätia. Najmodernejšie metódy, ktoré sú predmetom aktuálneho výskumu, sa zameriavajú práve na odstránenie momentových rázov. Sú založené na meraní (resp. výpočte) prúdových derivácií bez použitia injektáže napätia. Pri týchto metódach je nevyhnutná vysoká presnosť hodnôt prúdových derivácií v procese vyhodnocovania polohy rotora. Pri bežnom vektorovom riadení je možné pre meranie prúdových derivácií využiť nulový napät'ový vektor (Zero Voltage Vector - ZVV). Princíp spočíva v nasledovnom. V oblasti nízkych rýchlostí motora je doba trvania nulového napät'ového vektora prirodzene predĺžená kvôli malej hodnote indukovaného napätia v motore. Preto v tejto oblasti prevádzky motora je možné vykonávať meranie prúdových derivácií s dostatočnou presnosťou. Pre meranie prúdových derivácií sa používa Rogovského cievka [7], alebo sa pre presné snímanie prúdu počas trvania nulového napät'ového vektora využijú A/D prevodníky so zvýšeným rozlíšením. Avšak spomenuté riešenia sú cenovo neefektívne a doplnenie systému o ďalší snímač pre snímanie prúdových derivácií je pre komerčné použitie väčšinou nežiaduce. Blokovaná schéma pre

pozorovanie polohy a uhlovej rýchlosti pre nulový napäťový vektor je znázornená na Obr. 8.



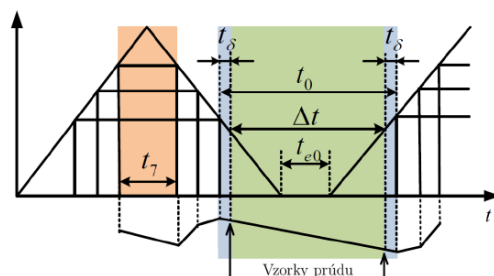
Obr. 8 Bloková schéma pozorovateľa pre metódu s nulovým napäťovým vektorom [8]

Priestorová vektorová modulácia (Space Vector Modulation - SVM) je jednou z najmodernejších a najpoužívanejších techník pre generovanie signálov pre ovládanie tranzistorov striedača. Pomocou SVM sa vytvára požadované modulované napätie pre napájanie SMPM. Želané napätie pre SVM sa dá opísať pomocou vektora, ktorého koncový bod v komplexnej rovine opisuje kružnicu [6]. Táto kružnica môže byť aproximovaná výkonovým striedačom, ktorý umožňuje kombináciou rôznych stavov spínačov vytvoriť dva nulové napäťové vektory: V_0 a V_7 a 6 aktívnych vektorov: V_1 až V_6 . Ak je aplikovaný vektor V_0 , všetky tri dolné tranzistory vo vetvách striedača sú zapnuté a na vinutia stroja sa pripojí záporný pól medziobvodu. Naopak, pri vektore V_7 sú zapnuté horné tranzistory striedača a na vinutia stroja sa pripojí kladný pól medziobvodu. Doba trvania vektora V_0 je t_0 a doba trvania vektora V_7 je t_7 , pričom platí, že $t_0 = t_7$. Nulové napäťové vektory sú farebne vyznačené na Obr. 9.



Obr. 9 Klasický prístup merania prúdových derivácií: (a) Nosný signál pre moduláciu, (b) Spínacie signály pre horné tranzistory vo vetvách striedača, (c) Ideálny a skutočný fázový prúd [6]

Pre zlepšenie presnosti snímania prúdových derivácií je možné proporcionálne zmeniť dĺžku intervalov t_0 a t_7 . Interval t_7 sa zmenší o dobu t_0 , pričom interval t_0 bude o túto dobu predĺžený, Obr. 10.



Obr. 10 Proporcionálna zmena intervalu t_0 a t_7 [6]

Takáto úprava podľa Obr. 10 umožní vytvorenie väčšieho časového okna pre meranie prúdových zmien, ktoré sú následne merateľné aj menej kvalitnými snímačmi prúdov.

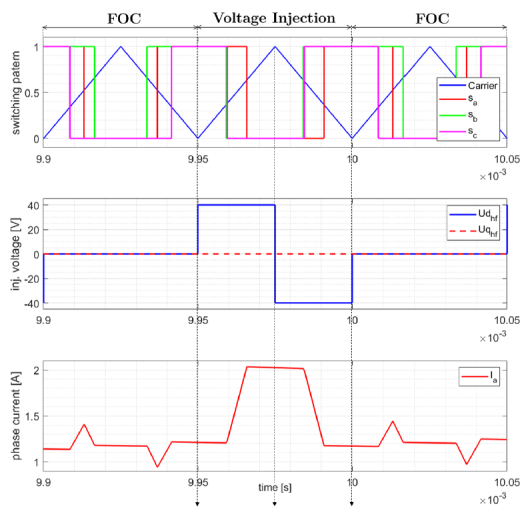
III. NOVONAVRHNUTÉ METÓDY BEZSNÍMAČOVÉHO RIADENIA

Ako už bolo spomenuté, pri injektáži VF napäťových impulzov jedna riadiaca perióda obsahuje tri spínacie periódy. Len v prvej perióde sa vykonáva vektorové riadenie. V nasledujúcich dvoch periódách je motor odpojený od riadiacej štruktúry, pričom sa injektuje kladné a záporné napätie. Následne z prúdových zmien počas injektáže sa vyhodnotí poloha rotora. Hlavnú nevýhodu tejto metódy predstavuje odpojenie motora od riadiacej štruktúry počas trvania dvoch injektážnych periód. Počas trvania injektáže nie je moment motora riadený, a preto skrátenie doby, alebo úplná eliminácia odpojenia motora od riadenia prispeje k potlačeniu momentových rázov a zlepší kvalitu regulácie uhlovej rýchlosti. Vylepšenie tejto metódy sme navrhli v dvoch krokoch. V prvom kroku sa odstráni jedna z dvoch injektážnych periód. Docieli sa to zavedením režimu dvojitej aktualizácie PWM, ktorý umožní v rámci jednej spínacej periódy injektovať kladné aj záporné VF napätie. V druhom kroku sa injektáž v režime dvojitej aktualizácie PWM ďalej upraví tak, že injektáž aj riadenie motora budú vykonávané v rámci jednej spínacej periódy. Takto vzniknutá metóda bude predstavovať druhý krok vylepšenia metódy injektáže napäťových impulzov. Navrhnuté metódy sú opísané v nasledujúcich podkapitolách. Tieto vylepšenia doposiaľ podľa najlepších vedomostí autora neboli v literatúre publikované.

A. Injektáž VF napäťových impulzov v režime dvojitej aktualizácie

Počas jednej spínacej periódy dochádza k zapnutiu aj vypnutiu tranzistorov. Ak by sme hodnotu modulovaného napätia aktualizovali aj v strede spínacej periódy, dokázali by sme preniesť kladné aj záporné injektované napätie v rámci jednej spínacej periódy. Spínanie tranzistorov, injektované napätia a odozva fázového prúdu sú zobrazené na Obr. 11. V hornej časti je zobrazený 20 kHz nosný trojuholníkový signál a spínacie signály pre horné tranzistory v jednotlivých vetvách. V prvej perióde sú spínacie signály s_a , s_b a s_c

vytvorené na základe modulovaného trojfázového napätia pre riadenie motora. V nasledujúcej spínacej perióde dochádza k injektáži kladného a záporného napätia do pozorovanej \hat{d} osi. Prúdové vzorky potrebné pre pozorovanie polohy sa odoberú na začiatku, v strede a na konci injektážnej periódy.



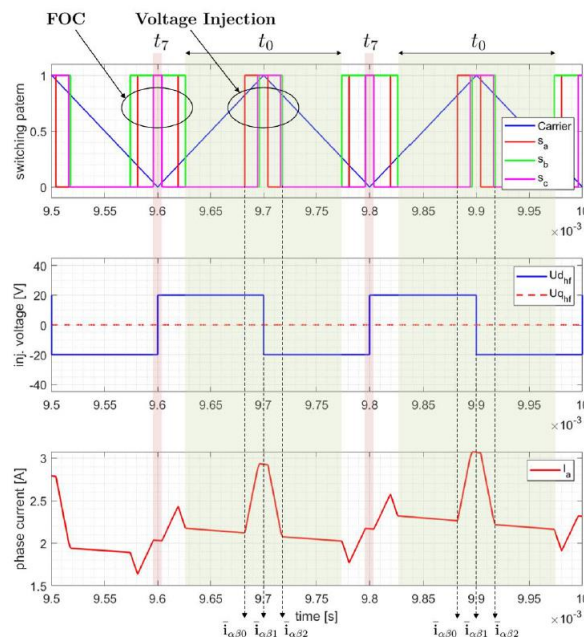
Obr. 11 Sekvencie injektáže VF napät'ových impulzov v režime dvojitej aktualizácie PWM s kompenzáciou nelinearit' striedača.

Z rozdielu $i_{\alpha\beta 0}$ a $i_{\alpha\beta 1}$ je vypočítaná prúdová zmena počas injektáže kladného napätia a z rozdielu $i_{\alpha\beta 1}$ a $i_{\alpha\beta 2}$ sa zistí zmena prúdu počas injektáže záporného napätia. Pre estimáciu polohy a uhlovej rýchlosti sa adaptuje identický pozorovateľ ako v pôvodnej metóde injektáže napät'ových impulzov, zobrazený na Obr. 7. Takýmto vylepšením sa zmenší aj počet spínaní tranzistorov pre potreby pozorovania polohy, čím sa zmenšia spínacie straty na výkonových súčiastkach, čo by sa malo prejaviť na zvýšení účinnosti bezsnímačového riadenia. Zároveň sa vektorové riadenie motora prepočítava o 33 % častejšie oproti pôvodnej metóde, preto sa zlepši kvalita riadenia uhlovej rýchlosti a dôjde k potlačeniu momentových rázov motora.

B. Injektáž VF napät'ových impulzov v nulovom napät'ovom vektore

Pre ďalšie utlmenie momentových rázov a zlepšenie kvality regulácie uhlovej rýchlosti motora by bolo vhodné, ak by sa vektorové riadenie motora vykonávalo v každej spínacej perióde. Preto je v tejto časti uvedený ďalší, druhý krok vylepšenia injektáže VF napät'ových impulzov, pri ktorom neodchádza k odpájaniu motora od riadenia a zároveň sa pri ňom využije režim dvojitej aktualizácie PWM. Pri klasickej modulácii sú nulové napät'ové vektory t_0 a t_7 rovnako dlhé. Ak by sme skrátili nulový napät'ový vektor t_7 a úmerne predĺžili napät'ový vektor t_0 , dostali by sme dostatočne veľké časové okno, t.j. časový úsek, v ktorom sa vykonáva injektáž. Navrhnutý algoritmus preto na začiatku každej spínacej periódy najprv vyhodnotí to, ktoré z troch fázových modulovaných napätí má najmenšiu hodnotu. Túto hodnotu napevno nastaví na 10 % z hodnoty nosného trojuholníkového

signálu. Ďalšie dve napätia sa upravujú tak, aby sa zachoval pôvodný napät'ový rozdiel od najmenej hodnoty. Pre injektovaný signál sme zvolili podobnú úpravu. Na začiatku každej spínacej periódy sa zistí, ktorá z troch modulovaných VF napätí v statorových súradniciach má najväčšiu hodnotu. Táto hodnota sa napevno nastaví na 90 % z hodnoty nosného trojuholníkového signálu. Ďalšie dve napätia sa upravujú tak, aby sa zachoval pôvodný napät'ový rozdiel od najväčšieho napätia. Riadiace signály, ktoré vznikli pre potreby riadenia motora sa spočítajú cez logický súčet s riadiacimi signálmi, ktoré vznikli pre VF injektáž. Výsledný riadiaci signál horných tranzistorov je zobrazený v hornej časti na Obr. 12.



Obr. 12 Injektáž VF napät'ových impulzov v modifikovanom nulovom napät'ovom vektore

Pre výpočet prúdových zmien sa odmeria prúd na začiatku a na konci injektáže kladného napätia $i_{\alpha\beta 0}$ a $i_{\alpha\beta 1}$, ako aj na konci injektáže záporného napätia $i_{\alpha\beta 2}$. Pre určenie polohy a uhlovej rýchlosti motora sa použije pozorovateľ z pôvodnej metódy injektáže napät'ových impulzov zobrazený na Obr. 7.

IV. EXPERIMENTÁLNE VÝSLEDKY

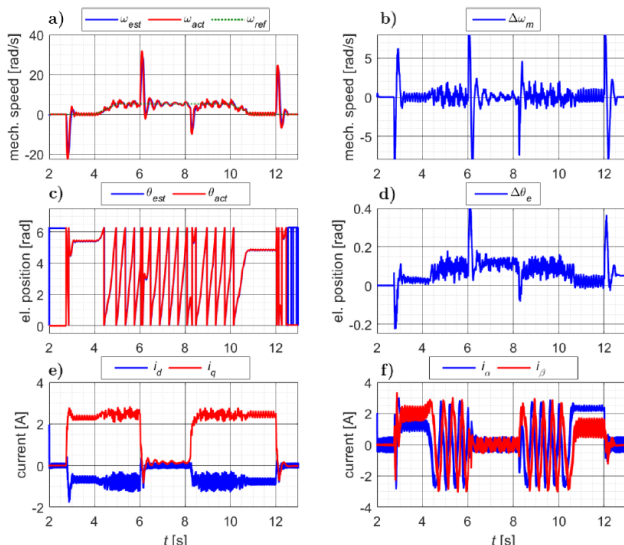
Pre experimentálne overenie metód bezsnímačových metód s VF injektážou bol použitý synchronný motor s permanentnými magnetmi, ktorého parametre sú uvedené v Tabuľke 1. Pre napájanie motora bol použitý prototyp výkonového striedača. Snímanie prúdov bolo zabezpečené prúdovými snímačmi CKSR 15-NP od firmy LEM, s citlivosťou 42 mV/A. Ako riadiaci člen bol použitý hardware-in-the loop (HIL) simulátor OP 5600 od spoločnosti Opal RT, umožňujúci rýchle prototypovanie z prostredia MATLAB/Simulink. HIL simulátor má 16 bit analógové vstupy so vzorkovaním 500 KS/s a digitálne výstupy so vzorkovaním 40 MS/s.

Tab. 1 Parametre SMPM od výrobcu Merkes

Napätie medziobvodu	$U_{DC} = 565 \text{ V}$
Nominálny výkon	$P_n = 2.53 \text{ kW}$
Nominálny moment	$M_n = 10.5 \text{ Nm}$
Nominálne napätie	$U_n = 380 \text{ V}$
Nominálny prúd	$I_n = 4.2 \text{ A}$

A. Injektáž VF napätia so sínusovým priebehom

Experimentálne výsledky sú zobrazené na Obr. 13. V čase $t = 2 \text{ s}$ je želaná rýchlosť natavená na nulovú hodnotu: $n_{ref} = 0 \text{ ot/min}$. V čase $t = 2.5 \text{ s}$ bol motor zaťažený 50% momentom nominálnej hodnoty. V čase $t = 4 \text{ s}$ bol motor rozbehnutý po rampe na $n_{ref} = 50 \text{ ot/min}$ (2% z nominálnych otáčok). V čase $t = 6 \text{ s}$ došlo k odpojeniu záťaže od motora a v $t = 8 \text{ s}$ bol motor opäť zaťažený. V čase $t = 10 \text{ s}$ bol zaťažený motor brzdený po rampe až na nulovú rýchlosť. Pri $t = 12 \text{ s}$ bola záťaž od motora opäť odpojená. Táto metóda vykazuje nárast chyby v pozorovanej polohe s narastajúcimi otáčkami motora. Najväčšie hodnoty chyby nastali pri prechodných dejoch po zaťažení motora. Maximálne hodnoty presahovali hodnotu 0.25 rad (14 el. stupňov). V ustálenom stave bola chyba v pozorovanej polohe do 0.2 rad (11.5 el. stupňov). Maximálne hodnoty v chybe pozorovanej rýchlosti nepresahovali hodnotu 10 rad/s (95.5 ot/min).

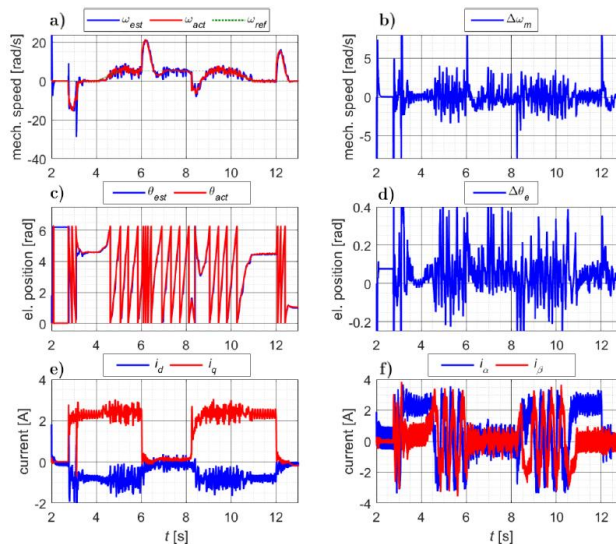


Obr. 13 Experimentálne výsledky injektáže VF napätia so sínusovým priebehom

B. Injektáž VF napätových impulzov

Na Obr. 14 sú zobrazené experimentálne výsledky pri zaťažení motora pri nulových otáčkach. Záťažový moment predstavoval 50% nominálnej hodnoty. Následne bol motor rozbehnutý na 50 ot/min (2% z n_n). Chyba v pozorovanej rýchlosti presahuje 5 rad/s (50 ot/min) a priebeh obsahuje väčšie množstvo šumu. Priebeh chyby v pozorovanej polohe tiež indikuje kmitanie pozorovanej polohy, pričom maximálne

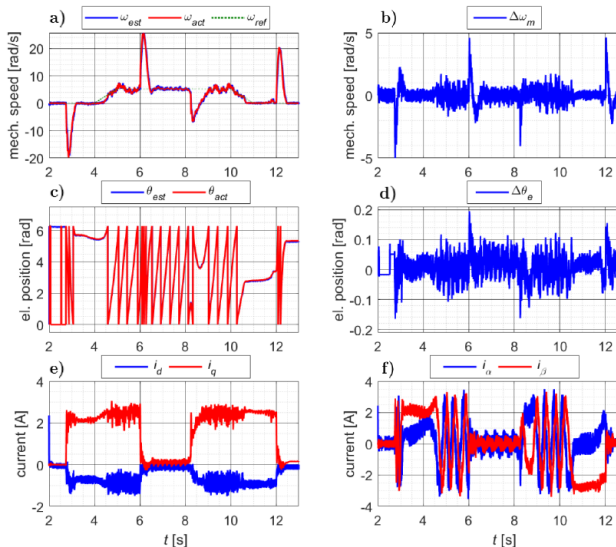
hodnoty sú okolo 0.4 rad (23 el. stupňov). Aj v tomto prípade platí, že počas celej doby bezsnímačového riadenia motora bola použitá len pozorovaná poloha a uhlová rýchlosť, pričom inkrementálny snímač sa využil len pre vyhodnotenie presnosti pozorovateľa.



Obr. 14 Experimentálne výsledky injektáže VF napätových impulzov

C. Injektáž VF napätových impulzov v režime dvojitej aktualizácie PWM

Experimentálne priebehy na Obr. 15 boli zhotovené pri 50% záťaži, pričom pohon bol rozbehnutý na $n_{ref} = 50 \text{ ot/min}$, čo predstavuje 2% z nominálnej uhlovej rýchlosti.



Obr. 15 Experimentálne výsledky injektáže VF napätových impulzov v režime dvojitej aktualizácie PWM

V prechodných dejoch bola maximálna chyba v pozorovanej uhlovej rýchlosti do 5 rad/s (48 ot/min) a v ustálenom stave bola blízka nulovej hodnote. Chyba v pozorovanej polohe

bola počas dynamických dejoch do 0.2 rad (12 el. stupňov) a v ustálenom stave sa držala pod hodnotou 0.1 rad (6 el. stupňov).

D. Zhodnotenie experimentálnych výsledkov

Pre štatistické porovnanie experimentálnych výsledkov sme použili strednú kvadratickú hodnotu chyby (root-mean-square error - RMSE) v pozorovaní polohy a uhlovej rýchlosti. Celkovo boli vybraté štyri úseky:

- T_1 zaťaženie pri nulových otáčkach - dynamický dej,
- T_2 zaťaženie pri nulových otáčkach - ustálený stav,
- T_3 rozbeh zaťaženého motora na želané otáčky,
- T_4 udržiavanie zaťaženého motora na želaných otáčkach.

Výsledky sú uvedené v Tabuľke 2. Pre prehľadnosť zavedieme zjednodušené označenia jednotlivých metód bezsnímačového riadenia:

- **Metóda 1:** Injektáž VF signálu so sínusovým priebehom,
- **Metóda 2:** Injektáž VF napätových impulzov,
- **Metóda 3:** Injektáž VF napätových impulzov v režime dvojitej aktualizácie.

V časovom úseku T_1 bola najmenšia hodnota strednej kvadratickej chyby v pozorovanej rýchlosti pri Metóde 3. Napriek skokovému zaťaženiu motora chyba v pozorovanej rýchlosti $RMSE_{\omega}$ nepresiahla hodnotu 13 ot/min. Pri ostatných metódach bola hodnota tejto chyby výrazne väčšia. Ani nominálne zaťaženie motora výrazne neovplyvnilo presnosť pozorovania rýchlosti použitím Metódy 3. Z pohľadu strednej kvadratickej chyby v pozorovanej polohe $RMSE_{\theta}$ mala opäť najlepšie výsledky Metóda 3. Maximálna hodnota bola 3.1 el. stupňov. Ani táto chyba sa nezväčšovala nárastom záťažného momentu. V časovom úseku T_2 a T_3 boli opäť najmenšie hodnoty $RMSE_{\omega}$ pri Metóde 3, ale pri nižších otáčkach bola presnosť pozorovania polohy bola presnosť pozorovania polohy lepšia s použitím pôvodnej metódy injektáže napätových pulzov. V úseku T_4 mala najväčšiu presnosť Metóda 3 v oboch kategóriách hodnotenia presnosti.

Tab. 2 Hodnoty strednej kvadratickej chyby pre pozorovanú polohu a rýchlost pre všetky 3 metódy.

	RMSE _ω [ot/min]	RMSE _θ [el. stupne]	Metóda 1		Metóda 2		Metóda 3	
			RMSE _ω	RMSE _θ	RMSE _ω	RMSE _θ	RMSE _ω	RMSE _θ
T_1	Mz = 0.5Mn	$n_{ref} = 15$ rpm	21.5	3.7	27	10.4	10.1	3.1
		$n_{ref} = 50$ rpm	23.9	7.6	33.7	14.6	10.8	2.8
		$n_{ref} = 50$ rpm	n/a	n/a	n/a	n/a	12.5	2.7
T_2	Mz = 0.5Mn	$n_{ref} = 15$ rpm	16.4	2.9	3.4	1.5	1.7	3.4
		$n_{ref} = 50$ rpm	11.8	5.9	3.8	2.4	1.8	1.9
		$n_{ref} = 50$ rpm	n/a	n/a	n/a	n/a	6.2	2.9
T_3	Mz = 0.5Mn	$n_{ref} = 15$ rpm	5.6	1.8	3.9	1	1.7	2.3
		$n_{ref} = 50$ rpm	6.7	5.4	10.2	4.3	4.3	2.2
		$n_{ref} = 50$ rpm	n/a	n/a	n/a	n/a	8.6	2.4
T_4	Mz = 0.5Mn	$n_{ref} = 15$ rpm	4.4	2.2	10	6.3	2.9	1.8
		$n_{ref} = 50$ rpm	6.2	3.9	13.7	7.2	5.7	2.4
		$n_{ref} = 50$ rpm	n/a	n/a	n/a	n/a	8.9	2.7

V. ZÁVER

V tomto článku boli opísané a experimentálne porovnané vybrané metódy bezsnímačového riadenia SMPM pre oblasť nízkych rýchlostí. Najprv bola experimentálne overená metóda injektáže VF signálu so sínusovým priebehom. Táto metóda spôsobuje kmitanie momentu, keďže VF signál je superponovaný na základnú frekvenciu riadiaceho napätia. Ďalšia nevýhoda sú digitálne filtre, ktoré sú nevyhnutné v procese pozorovania polohy. Filtre znižujú pásmo priepustnosti regulátorov, čo nepriaznivo vplyva na kvalitu regulácie. Odlišným prístupom je metóda, pri ktorej injektáž prebieha oddelene od vektorového riadenia motora. Táto metóda je označovaná ako injektáž VF napätových impulzov. Táto metóda je jedna z najnovších publikovaných metód a preto sme sa zamerali na preskúmanie možností jej vylepšenia. Ako hlavnú nevýhodu sme identifikovali nutnosť odpojenia motora od riadiacej štruktúry po dobu dvoch spínacích periód, ktoré nasleduje jedna riadiaca perióda. Injektáž napätových impulzov v režime dvojitej aktualizácie PWM umožňuje riadenie motora v každej druhej spínacej perióde čo potláča nežiaduce kmitanie otáčok ako aj momentu stroja. Pri tejto novonavrhnutej metóde sa dosiahli najlepšie experimentálne výsledky bezsnímačového riadenia motora z pohľadu presnosti v pozorovanej polohe a uhlovej rýchlosti motora.

REFERENCIE

- [1] W. Qian, X. Zhang, F. Jin, H. Bai, D. Lu, and B. Cheng, "Using high-control bandwidth fpga and sic inverters to enhance high-frequency injection sensorless control in interior permanent magnet synchronous machine," IEEE Access, vol. 6, pp. 42 454–42 466, 2018.
- [2] Y. Zhao, Z. Zhang, C. Ma, W. Qiao, and L. Qu, "Sensorless control of surface mounted permanent-magnet synchronous machines for low-speed operation based on high-frequency square-wave voltage injection," in 2013 IEEE Industry Applications Society Annual Meeting, 2013, pp. 1–8.
- [3] D. Kim, Y.-C. Kwon, S.-K. Sul, J.-H. Kim, and R.-S. Yu, "Suppression of injection voltage disturbance for high-frequency magnet square-wave injection sensorless drive with regulation of induced high-frequency current ripple," IEEE Transactions on Industry Applications, vol. 52, no. 1, pp. 302–312, 2016.
- [4] R. Ni, D. Xu, F. Blaabjerg, K. Lu, G. Wang, and G. Zhang, "Square-wave voltage injection algorithm for pmsm position sensorless control with high robustness to voltage errors," IEEE Transactions on Power Electronics, vol. 32, no. 7, pp. 5425–5437, 2017.
- [5] G. Xie, K. Lu, S. K. Dwivedi, J. R. Rosholm, and F. Blaabjerg, "Minimum voltage vector injection method for sensorless control of pmsm for low-speed operations," IEEE Transactions on Power Electronics, vol. 31, no. 2, pp. 1785–1794, 2016.
- [6] W. Gaolin, Z. Guoqiang, and X. Dianguo, Position Sensorless Control Techniques for Permanent Magnet Synchronous Machine Drives. Singapore: Springer Verlag, 2020.
- [7] C. Wu, Y. Zhao, and M. Sun, "Enhancing low-speed sensorless control of pmsm using phase voltage measurements and online multiple parameter identification," IEEE Transactions on Power Electronics, vol. 35, no. 10, pp. 10 700–10 710, 2020.
- [8] M. S. Mubarak, T.-H. Liu, C.-Y. Tsai, and Z.-Y. Wei, "A wide-adjustable sensorless ipmsm speed drive based on current deviation detection under space-vector modulation," Energies, vol. 13, no. 17, 2020. [Online]. Available: <https://www.mdpi.com/1996-1073/13/17/4431>

Design of Smart Network in an Industrial Zone

Rikin Tailor

Katedra elektroenergetiky
FEI TUKE
Košice, Slovenská republika
rikin.jitendrakumar.tailor@tuke.sk

Lubomír Beňa

Katedra elektroenergetiky
FEI TUKE
Košice, Slovenská republika
lubomir.bena@tuke.sk

Abstract— This paper draws its inspiration from the Industrial revolution, increasing electricity demand, rising electricity prices, and the necessity of an electricity management system. We are on the verge to change the whole concept of electricity generation, distribution, and utilization. In this paper, we have covered the topics such as how the industrial revolution happened, according to that global energy increase at industrial zone, present electricity management system at industry side, etc. As a solution to all these topics, we have designed the smart industrial network. This network is formed with the integration of all new technology in it. It also encourages to improvise in reliability, efficiency, price reduction in electricity bill, consistency, and stability of the system. To gain a comprehensive understanding of algorithm behavior in both the industry and at a zonal level, I have analyzed three different scenarios. This analysis aims to determine the most suitable approach for fostering industrial growth at the zone level.

Keywords— *Smart Network, Big Data, Industry 4.0, Internet of Electricity*

I. INTRODUCTION

Starting with the discussion on information that how industry evolves with respect to time,

Industry 1.0: Introduction of Machine Power:

Mankind is developing different objects for hundreds of years. The product or object developed by them was with very basic tools, which we can clearly say are handmade. The revolution had begun with England in 1760 and at the end of the decade, it reached the United States [1]. This Industrial revolution was labeled as a shift from a handicraft economy to dominated machinery.

Industry 2.0: The phase of Technological Change

The second revolution started with the first-ever commercial electricity generation source, and it is the present time one of the most sustainable electrical energy generation plants, Hydropower [2]. In between 1880 to 1900, many engineering marvels were noticed in terms of power generation. Coal power generation was started for driving DC dynamo with the steam engine. Sir Charles Parsons was the first person who initiated steam turbine generation with just 1.6% thermal efficiency [3].

Industry 3.0: The Digital Revolution

The third phase of revolution began in the 1950s and it's known as the Digital Revolution, the digital revolution just because it transformed the world from analog form to digital.

Industry 4.0: The Automation and advance Digital Revolution

Here we are at the latest industrial revolution which started at the beginning of the last decade in 2011 at the Hannover Trade fair in Germany. Industry 4.0 is characterized by the level of

automation; we can subject it as a union between physical assets and advanced digital technologies, such as Internet of Things (IoT), Artificial Intelligence (AI), Robotics, Drones, autonomous vehicles, 3d printing, cloud computing, big data, and others, that are coordinated, having the possibility to share the information, assessment and act on it.

There are four design principles identified as integral to Industry 4.0, and companies are foster to implement to fully utilize the capabilities of current technology [1].

- Interconnection
- Information transparency
- Technical assistance
- Decentralized decisions

II. ELECTRICITY LOAD

A. Traditional load.

It is very important to clear out what are the traditional load so that it would be easier to understand how their electricity consumption was and how it will be different with the new load. Traditional load is the load that we have been using for the past 2-3 decades in residence, commercial place, and industry. We can include residence load as air-conditioning, television, ceiling fan, coffee maker, cloth dryer, dishwasher, microwave oven, CFL lamps, refrigerator, vacuum cleaner, washing machine, etc. Some of the load from these we used as per the requirement. However, other loads are always connected to electricity. Similarly commercial load, we can count as the lighting of shops, offices, advertisements, heating, computer assembly, air-condition, etc. Heavy machinery, motors, PLC & SCADA systems, camera systems, etc., are comes in the category of Industrial load. All these loads fall in the category of capacitive, inductive, and resistive load.

B. Novel loads:

• Here are some electrical loads which rapidly took place into the market in recent time.

1. Electric Vehicle and charging infrastructure
2. Artificial Intelligence
3. Robotic Process Automation
4. Edge Computing
5. Quantum Computing
6. Virtual Reality
7. Blockchain
8. Internet of Things (IoT)
9. 5G
10. Cyber Security

C. Top Electricity consuming Industry

Chemical Industry:

The chemical industry is highly diverse, manufacturing a wide range of products falling within the chemical and fertilizer categories. However, the production of these chemical products requires a significant amount of energy. Examples include ammonia, chlorine, and various petrochemicals. Recent investigations have revealed that the fertilizer industry alone accounts for 1.2% of global energy consumption, with more than 90% of electricity being used in the production of ammonia

Metal Industry:

Iron and steel are the most prominent products in the metal industry, with steel being widely used in various aspects of everyday life. It serves as a crucial component in machines, tools, appliances, weapons, automobiles, buildings, infrastructure, and many other applications. The method of continuous casting plays a significant role in the production of steel, accounting for approximately 93% of the world's steel manufacturing

Cement Industry:

The cement industry is a major energy consumer and plays a vital role in infrastructure development and economic growth of countries worldwide. Cement production is prevalent across the globe, with developing countries accounting for 93% of the global cement output. Notably, China emerged as the largest energy consumer in the industry, with approximately 50% of its building materials energy consumption attributed to cement production. In India, small-scale cement factories consume an average of 104 kWh per ton of cement

Paper and Pulp Industry:

The European Union (EU), United States, China, and Japan are the leading regions in paper production, accounting for 80% of the total energy consumption in the industry. Roughly half of this energy is utilized in the pulping process, while the remaining half is dedicated to papermaking. The primary form of energy used in pulping and papermaking mills is heat and steam, which powers the integrated operations of these facilities.

The most crucial information is that, despite the fact that we previously discussed unique loads of electrical energy for the future, we never calculated these loads in order to forecast the deployment of renewable energy in the present and the future

III. SMART GRID CONCEPT

The Smart Grid represents a digital revolution in the power industry, incorporating advanced technologies like bidirectional communication and transmission line sensing. It functions as an independent system, swiftly resolving issues and providing accessible, reliable, secure, and high-quality electricity to all customers.

The development of the smart grid is the result of prior attempts to use electronic control, measurement, and analysis. First utilized in the 1980s to track large customers' loads, automatic meter reading later evolved into the Advanced Metering Infrastructure of the 1990s, whose meters could record how much power was used throughout the day [4]. The smart meter can be used as a gateway to demand response-aware devices and Smart Sockets in the home and has the ability to maintain communication so that monitoring can be done in real-time. Devices with higher electricity consumption, such as commercial and residential air conditioners, refrigerators,

heaters, dishwashers, clothes washers, and dryers, changed when they were supposed to run in order to avoid activating when the grid was experiencing peak demand.

Various technologies are employed in the Smart Grid, including broadband over power lines (BPL) and wireless mesh networking. These technologies ensure a reliable connection between different devices within homes and support utility metering for gas and water [5].

A. Features of the smart grid

1. Reliability
2. Efficiency
 - Load adjustment/Load balancing
 - Peak curtailment/leveling and time of use pricing
3. Flexibility in network topology
4. Sustainability
5. Market-enabling
6. Demand response support
7. Platform for advanced services
8. Provision megabits, control power with kilobits, sell the rest.

B. Technology [6]

1. Integrated communications
2. Sensing and measurement
3. Smart meters
4. Phasor measurement units.
5. Distributed power flow control
6. Smart power generation using advanced components
7. Power system automation.

The smart storage system plays a vital role in integrating renewable energy sources into the market and grid applications. It addresses the issue of resource unreliability and enables the delivery of scheduled power. Additionally, it helps to balance the synchronization between energy consumption and generation. Both stationary energy storage systems and mobile (vehicle) systems are expected to have a significant impact in this energy transition. Currently, only around 5% of the installed generation capacity in the European electricity system includes storage systems, primarily pumped hydroelectric energy storage. The required storage capacity will vary depending on the mix of renewable energy sources in the overall production capacity. It is projected that European scenarios by 2050 will introduce storage capacity ranging from 43 GW to 90 GW, with an estimated investment of approximately \$80 billion to \$130 billion [7].

IV. OBJECTIVE OF PAPER

The most amazing and cutting-edge technology that has made the grid smarter, autonomous, and sustainable is the concept of the smart grid and microgrid. The entire world has started investing and researching to improve their system's dependability and environmental friendliness. The primary goal of both smart grids and microgrids is to give consumers constant access to electricity. The conditions for effective use of the smart grid are bidirectional communication flow, bidirectional power flow, resilience in self-healing technology employing enhanced sensor parameters, and distributed generation. The goal of the

thesis is to create an advanced, semi-autonomous smart industrial management system with the ability to support diverse batteries effectively, install smart EV charging stations in parking lots, use self-healing technology, and exchange energy with other industries.

V. ADVANCEMENTS IN INDUSTRIAL GRID RESEARCH: A FOCUS ON RECENT DEVELOPMENTS

A starting point for grid research is understanding the feasibility of different renewables connected to grid, I have reviewed a research paper that focuses on the feasibility of grid-connected photovoltaic systems in small-scale industries. The study assessed photovoltaic technology based on three key criteria: energy, environment, and economics. The researcher conducted a sensitivity analysis considering various devices connected to the grid. A 53.66 kW photovoltaic array was installed, enabling the industry to generate 79 MWh/year of energy. Not only did this help meet the industry's electricity demand, but it also allowed them to sell 18 MWh of surplus electricity back to the utility company [8]. I came across another intriguing research paper that discusses the off-grid performance of industrial grids, with a specific focus on integrating renewable energy sources into industrial clusters. The paper addresses three key areas of investigation. Firstly, it explores wind turbine generation by utilizing the Weibull equation to accurately determine the probability density of wind speeds. Secondly, it examines the solar generation capacity by estimating the required surface area for solar arrays. Lastly, the study investigates the feasibility of transmitting electricity at different voltage levels over varying distances. Notably, the paper emphasizes the importance of a comprehensive approach that includes geographical and technological zoning of territories, as well as cumulative cost analysis. This approach proves effective in evaluating initial solutions and determining the most suitable power supply option, whether centralized or standalone [9].

The next step involves gaining a comprehensive understanding of the design aspects of industrial grids. This research paper focuses on exploring the various considerations and factors involved in designing an efficient industrial grid system. The author of [10] the paper has conducted a detailed analysis of the challenges encountered in developing a robust and efficient computational industrial grid. The study specifically examines optimization problems and proposes a solution using the Multi-objective Genetic Algorithm theorem. To enhance the development of the grid system, a Decision Support System (DSS) is introduced to ensure autonomous performance, reliability, and cost effectiveness. Throughout the design process, practical issues were identified and successfully resolved using Artificial Neural Networks. Wei Luo presents a novel approach to incrementally plan the distribution network in response to the growing number of businesses and industries in an Industrial zone. The incremental development is categorized into two classes: one focuses on power distribution planning, asset ownership, while the second class concentrates on distribution network investment, construction, and operation by enterprises. This approach introduces mixed ownership in power delivery, alleviating the burden on the distribution company. The revenue model is based on service fees. The paper proposes the implementation of a double ring/loop network with a transmission capacity of 110kV to ensure smooth and reliable

operation [11]. In the context of smart grid communication, an alternative approach to traditional electrical sensors is proposed using pulse position modulation-based energy detection. This method offers a viable solution for the House Area Network and Neighborhood Area Network applications. By leveraging this technology, more efficient and reliable communication can be achieved within the smart grid infrastructure [12].

There is a big part power grid in terms of supplying power by taking from power plant and delivering to industrial grid. The authors investigate an approach to optimize power supply in the power grid, focusing on maximizing the revenue of both power plants and the power grid. They propose the implementation of a direct power purchase from power plants along with an Incentive based demand response strategy (IBDRM) [13]. This strategy is designed for smart grid technology and requires real-time information about power requirements from industrial users.

Two different scenarios are examined: one without implementing IBDRM and the other with IBDRM. The results show that direct power purchase without IBDRM leads to high electricity demand and pricing, with minimal involvement of the power grid. In this scenario, the power plant takes a leading role, while industrial consumers act as followers.

However, when IBDRM is implemented, the power grid assumes the role of a manager, supporting both the power plant and industrial consumers to adhere to the IBDRM policy. In this case, the power grid becomes the leader, the power plant becomes a supporter, and industrial consumers remain as followers. The extent of benefits for the power grid depends on the level of participation from industries.

It is important to note that this system may not be highly beneficial for industries, as it requires compromising their production by reducing power consumption during peak demand. This can lead to higher labor and operational costs. Additionally, adjusting production schedules becomes challenging, especially considering the increasing global population and demand.

The research suggests that despite exploring various possibilities, the power grid may not reap significant benefits from this approach. A similar strategy involving time-of-use pricing is also presented by Gang Wang [14].

This research paper proposes the utilization of an enhanced version of the state network task (STN) model as a means to improve power management in industrial settings, going beyond the scope of traditional demand response strategies. The model incorporates two key elements: the state node and the task node, which are used to represent the industrial user's operational states and tasks, respectively.

The main objective of this study is to present optimal power management tactics that can effectively minimize costs and address operational constraints. By leveraging the STN model, industrial users can not only reduce their electricity expenses but also gain more control over their equipment usage, enabling them to optimize their operations based on specific time intervals.

To demonstrate the effectiveness of this approach, a comprehensive case study is conducted. The results highlight that the proposed system not only leads to significant cost reduction but also empowers users to better utilize their

equipment according to their specific needs and operational schedules.

Overall, this research emphasizes the advantages of adopting an improvised version of the STN model for industrial power management, offering an innovative approach to optimize costs and improve operational efficiency [15].

In the realm of distribution networks, it is crucial to address the impact of climate conditions and natural disasters on their operation. Events such as typhoons, ice disasters, and thunderstorms can cause significant damage to the network infrastructure, leading to prolonged power outages depending on the severity of these events. While it is impossible to prevent such occurrences, being prepared for them is essential. Early detection and warning systems play a vital role in this preparation.

In a research study conducted by Yan Zhang and colleagues, the focus is on providing a comprehensive understanding of early intelligent warning systems for distribution network operation failures. The study utilizes data mining techniques, correlation analysis, and causality analysis to assess a wide range of variables. By collecting and analyzing both electrical and non-electrical data, the researchers aim to generate successful early warnings for distribution grid failures [16].

The research paper presents a case study that incorporates actual data, enabling the simultaneous assessment of numerous variables. This comprehensive approach contributes to the development of effective warning systems for distribution networks, helping to mitigate the impact of climate-related events and reduce downtime.

Overall, the study underscores the importance of early intelligent warning systems in ensuring the resilience of distribution networks. By leveraging data analysis techniques and considering various variables, researchers aim to enhance the preparedness and response mechanisms for distribution grid failures caused by natural calamities:

VI. DESIGN OF SMART NETWORK IN AN INDUSTRIAL ZONE

The idea behind the smart network for the Industrial zone started with the true ambition of uninterrupted, efficient, sustainable, and autonomous electricity supply to all forms of consumers. I am using term all forms of consumers just because the residential load is highly dependent on industrial and commercial load. Taking little dive into it to understand how it is dependent, as we know due to higher demand compared to electricity generation, we are using different techniques on the residential side to balance electricity such as demand response and time of use. These techniques, we are using in peak demand time and peak demand time is what when industries are functioning. Most appropriately, we can say in between 9:00 AM to 6:00 PM. It means if Industries can generate even half of their demand, then it will reduce the burden from the utility, and then they can transfer this power to the residential side. As this system, it is also possible at residential side smart management system so it would be more effective to balance the electricity.

Many research papers, journals, and articles are being studied to see the real challenges in electricity. Here I am proposing an idea to facilitate industry on self-sustained in most of the cases. By acknowledging industry 4.0 (Industry revolution 4.0), we are designing autonomous systems in an industrial zone which will have solar generation, micro-wind generation, Smart

control panels, peer-peer electricity transfer, proper security, distribution management, EVs charging infrastructure, etc.

Digitalization has been growing rapidly in all sectors of technology. There is no doubt to say we live in a virtual universe. Indeed, to stay side by side or for moving forward, it would be obvious to acknowledge measures in electricity digitalization. As IoT (Internet of things) term in the computer universe, in electricity, it is known as the Internet of Electricity (IOE). It demonstrates the feasibility of real-time communication and remote control by digitalizing sensing of the delivery of electricity. Digitalize perspective also provides a platform for self-healing and instant response framework

In the paper, I have presented two different frameworks. The first framework is for the cement industry. Which has an electricity supply via the utility grid and other traditional approaches of diesel generators. This traditional approach is replaced with new renewable energy resources such as PV generation and Wind generation. A one-line diagram of the cement industry has been used. We will see detailed information on it in the following section. Another framework is for the Industrial zone, where I have considered 5 different industries, their load profile, energy generation, zone-wise energy storage system, and EV charging infrastructure. We will observe exact working conditions in the algorithm section.

A. Cement Industry Infrastructure

As shown in Fig.1, this is quite a big cement industry, the different parameters of this industry are as follows:

- Total Employees: 450
- Production capacity per day of cement plant : 1200 Tonnes
- Electric Power consumption per tonne : 110 kWh/tonne
- Considered variable load: 2-5%

consumption is 110 kWh. Based on this data daily consumption of the cement industry is,

Daily Energy Consumption (kWh) = Daily Production (tonnes) × Electricity Consumption per tonne (kWh/tonne)
 Daily Energy Consumption (kWh) = 1200 tonnes × 110 kWh/tonne = 132,000 kWh

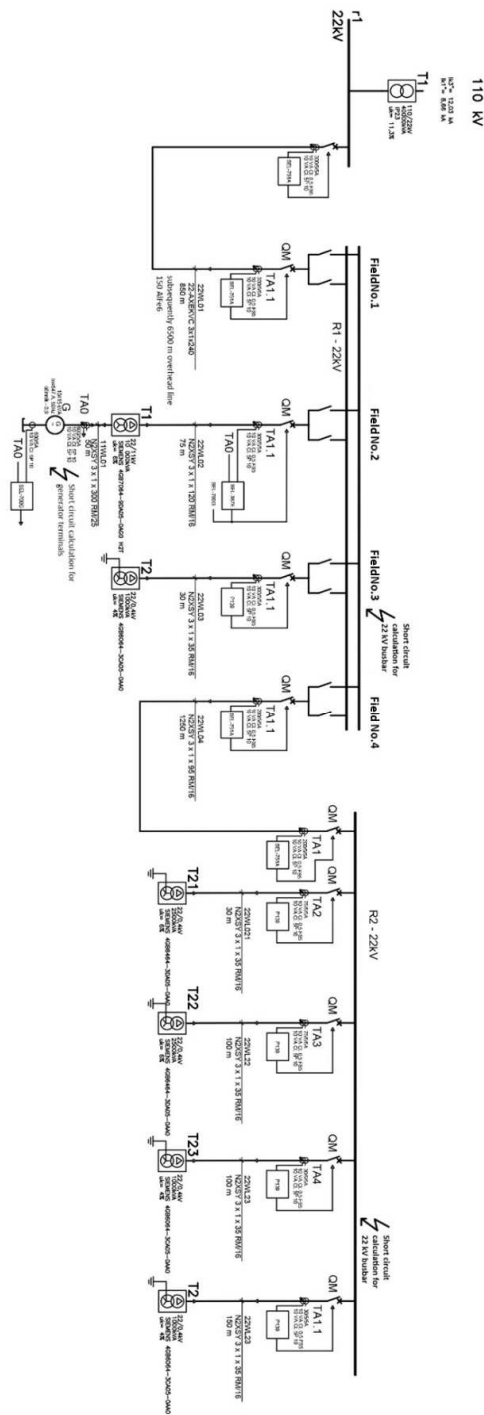


Fig. 1. One line Diagram of Industry without PV, Wind and Battery

Fig. 2 consists of load profile of cement industry. This is one of the industries which functions 24 hours' time. An industry has a daily production capacity of 1200 Tonnes and per tonne power

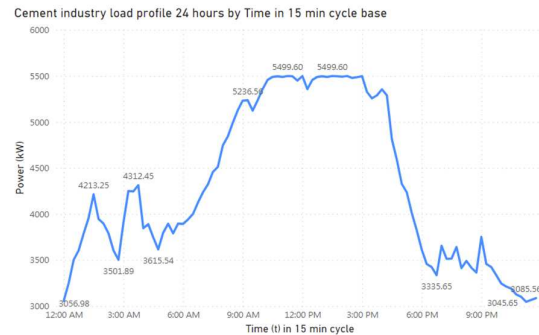


Fig. 2. Cement Industry Load Profile

A pick load of the cement plant is almost 5500 kWh, and we can see that it is during the pick hours. At the start of the day and end of the day the load profile is quite low compared to pick time. This load can create quite a big burden on the utility power supply during the pick time. Also due to pick hours, the per unit price is also quite high. This could be solved by accepting renewable energy and storage systems in industry.

Now let's jump to the implementation of PV generation, wind generation and storage systems of the industry. In accordance with the visual depiction of figure 3. Renewable energies are connected at the starting of the plant. The main reason behind the idea is that there is a heavy load at Field no. 1, 2, 3 and 4.

A new facility at the industry is as follows,

- Wind Generation capacity: 1 MW
- Battery Storage: 14.5 MWh for 2 hours back up (following is the calculation for it) with power dissipation is 3250 KW max.

The calculation process for estimating the battery capacity is as described below,

Determine the total energy requirement for 2 hours of battery backup: Backup Energy Requirement (kWh) = Daily Energy Consumption (kWh) × Backup Duration (hours)
 Backup Energy Requirement (kWh) = 132,000 kWh × 2 hours / 24 hours = 11,000 kWh

Adjust for battery efficiency: Usable Storage Capacity (kWh) = Backup Energy Requirement (kWh) / (Battery Efficiency)
 Usable Storage Capacity (kWh) = 11,000 kWh / 0.8 = 13,750 kWh

Adjust for depth of discharge (DoD): Required Battery Storage Capacity (kWh) = Usable Storage Capacity (kWh) / (1 - DoD)
 Required Battery Storage Capacity (kWh) = 13,750 kWh / (1 - 0.05) = 14,473 kWh

Therefore, the ideal battery storage capacity for 2 hours of backup for this cement plant is considered approximately 14,473 kWh or 14.5 MWh. Required Battery Power per hour (MW) = 14.5 MWh / 2 hours = 7.25 MW

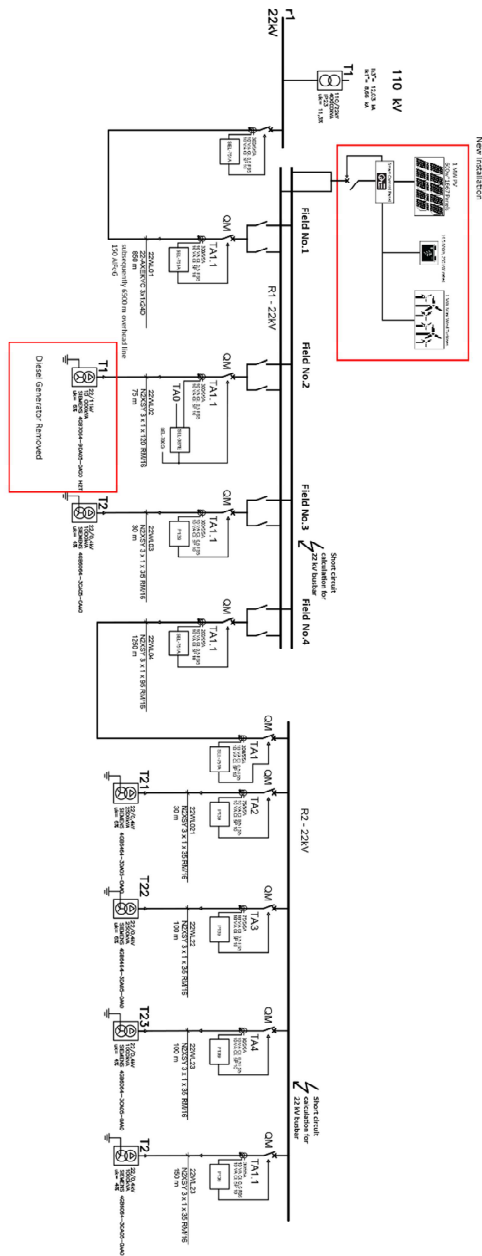


Fig. 3. One line Diagram of Industry with PV, Wind and Battery

After knowing the size of the battery, now let's discuss the PV and wind power generation curve. It is easy to analyze the speed of wind during the daytime in Fig. 4 and based on that wind turbine generation in Fig. 16. In the graph, we can see that the wind speed varies between 3.9 to 6.4 m/s. The maximum generation from the wind turbine is 885 kW at 2:45 AM, 12:15 PM and 5:45 PM. In addition to this, we could observe that wind is

available almost all the time in a day. It is obviously an intermittent source but, in this case, it is considered to be available the whole day. Minimum generation from the Wind Turbine is 530 kW when the speed is 3.90 m/s.

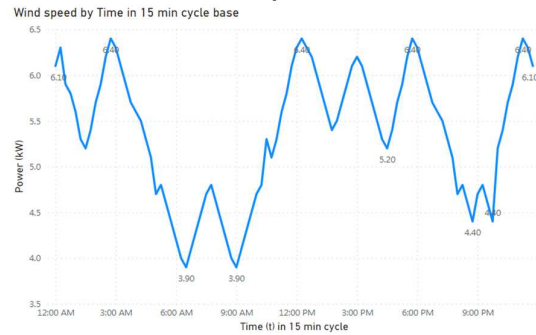


Fig. 4. Wind Speed

Furthermore, we can observe the PV power generation curve in Fig. 5. The period of PV generation spans from 6:15 AM until 6:45 PM, coinciding with daylight time. The maximum power generation by the PV plant is during the 12:15 PM to 2:45 PM. The peak power generation value is recorded as 981.20 kW.

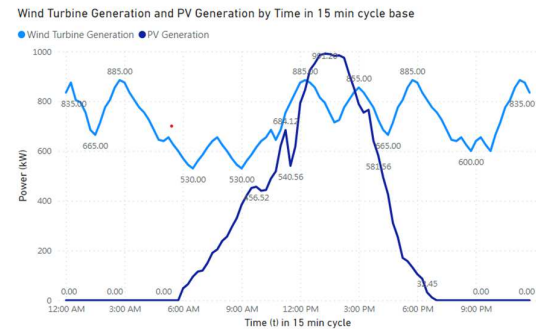


Fig. 5. Wind and PV generation Cement Industry

Battery storage system also performs major support during the pick time. The working principle of electricity flow is given in the algorithm. The impact of PV generation, wind generation, and the battery storage system on the industrial load curve is presented in the results section.

With a basic understanding of the load profile of the cement industry and the integration of renewable energy sources, the next step is to shift our attention towards designing a smart network at the industrial zone level.

B. Smart Network in industrial zone

Industries account for the largest share of electricity consumption, and research and development efforts have shown that utilities are actively engaging with industries to adopt time-of-use and demand response measures. However, there are two approaches to consider in these scenarios. The first option involves shifting the load curve from peak hours to normal hours, with the industry making certain compromises and benefiting from other utility incentives. However, many

industries are constrained by specific operating conditions and workforce availability, limiting their ability to operate outside of daytime hours. In such cases, a viable solution emerges: the implementation of a Smart network in the industrial zone. This approach goes beyond merely making individual industries or distribution grids intelligent; it aims to create a comprehensive smart system by integrating various components.

In order to gain a deeper understanding of smart networks and their operational principles, we have examined five distinct industries located within an industrial zone. These industries exhibit variations in terms of operating hours, electricity load patterns, load variability, energy generation resources, and storage systems.

To begin our exploration, let's analyze the layout of the Smart Network, as depicted in Figure 6. The network comprises multiple components, including industries, grid storage systems, electrical vehicle infrastructure, and crucial elements of the grid such as the Central grid connection equipped with monitoring and management controllers. The grid is structured in a ring configuration, optimizing the utilization of infrastructure by enabling peer-to-peer electricity exchange. Moreover, this configuration offers the advantage of alternate power supply routes in the event of a fault or disruption. Each industry is equipped with a smart network panel at the point of connection with the grid, incorporating sensors that enable connection and disconnection based on power supply conditions.

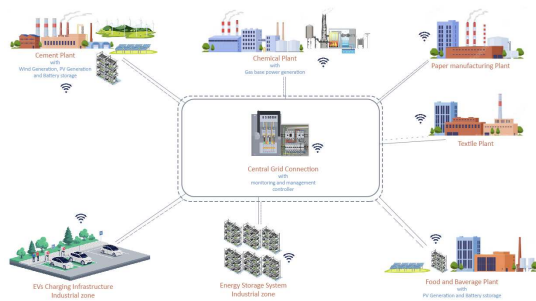


Fig. 6. Smart Network Layout

Central Grid connection with controller has three power supply possibilities:

1. Direct grid supply
2. Distribution storage supply
3. V2G energy supply

The management controller also offers the capability to feed distribution-generated energy back into the utility grid. Power Electronics plays a crucial role in converting energy at various levels and during different transitions. It is also interconnected with the zonal level Energy Storage System (ESS). Automated switches controlled by microprocessors drive transformer protection and various measuring sensors throughout the street. Additionally, multiple sensors are incorporated, such as an automatic reconfiguration algorithm sensor for detecting multiple energy sources, an automatic disconnection sensor to address electricity overuse or faults in any branch, a voltage regulation sensor, an automatic alert sensor for transmitting electricity to the grid from the central grid connection, a missing

sensor identification, and a self-healing device sensor. For this considered that investment could be possible by together with industry and utility for energy storage system at zone and EVs infrastructure. The government can also contribute with providing subsidies.

As mentioned earlier, the management controller provides access to three different energy sources. These sensors dynamically detect the availability of each source and establish the appropriate connections accordingly.

- a) Some of the characteristics of Smart Network is as follows:
- b) Sensor for isolation of industry in case of fault
- c) Voltage regulation sensor:
It will compare the actual voltage magnitude required with receiving magnitude and accordance with it, it triggers a voltage regulator at a control panel or sends the command to a utility center for any changes. The voltage regulation sensor's aim is to provide steady voltage to the consumer irrespective of how much power is drawn from the lines.
- d) Automatic alert sensor for transmitting electricity to the grid from smart network controller and vice versa.
- e) Automatic streetlight turn ON – OFF sensor:
This sensor is time-dependent, it will turn ON-OFF the light according to day-nighttime. The utility can also give a command from the center to turn ON-OFF.
- f) Missing sensor identification:
- This sensor can link with all the sensor; work is just to notify whether all sensor is working properly or not.
- g) Self-healing device sensor:
- If a fault does not have a severe cause and does not detect continuously then this sensor will give the command to reestablish the connection.

Now, let's delve into the concept of peer-to-peer energy transfer. As mentioned previously, in addition to the central grid connection, each industry is equipped with a smart network panel. These panels have the ability to communicate with each other through cloud-based platforms, enabling the detection and exchange of energy among various loads.

When one industry's panel transmits information regarding the power exchange to another specific industry, only that particular industry's network panel can receive and consume the transmitted power. This peer-to-peer energy transfer mechanism facilitates direct energy sharing and consumption between industries within the smart network.

Before assessing the algorithm that explains the behavior of electricity flow at both the individual industry and zonal levels, let's first examine the distinct industries and their load patterns. Which would be useful for the case study.

Having already covered the cement industry in our previous discussion, let's now turn our attention to the paper industry.

1) Paper manufacturing Industry

- Total Employees: 100
- Production capacity per day of paper plant: 10 Tonnes per day
- Electric Power consumption per tonne: 800 kWh/tonne

We can calculate energy consumption per day as,

Total Energy Consumption per day = Production Capacity × Energy Consumption per tonne
 Total Energy Consumption per day = 10 tonnes × 800 kWh/tonne = 8,000 kWh

To convert to megawatt-hours (MWh):

Total Energy Consumption per day in MWh = Total Energy Consumption per day / 1,000
 Total Energy Consumption per day in MWh = 8,000 kWh / 1,000 = 8 MWh

Therefore, the estimated electric power needed for a small-scale paper making industry with a production capacity of 10 tonnes per day would be approximately 8 MWh and rated power for the industry is 8 MWh / 24 = 0.333 MW

2) Chemical Industry

- Total Employees: 400
- Production capacity per day of chemical plant (Chlorine manufacturing): 400 Tonnes
- Electric Power consumption per tonne : 3000 kWh/tonne

A energy consumption per day,

Total Energy Consumption per day = Production Capacity × Energy Consumption per tonne
 Total Energy Consumption per day = 400 tonnes × 3000 kWh/tonne = 1,200,000 kWh

To convert to gigawatt-hours (GWh):

Total Energy Consumption per day in GWh = Total Energy Consumption per day / 1,000
 Total Energy Consumption per day in GWh = 1,200,000 kWh / 1,000,000 = 1.2 GWh

Rated Power in GW = 1,200,000 kWh / 24 hours / 1,000,000 = 0.05 GW

Therefore, it is projected that a chemical factory with a daily output capacity of 400 tonnes would require 1.2 GWh energy and 0.05 GW of rated electric power.

The chemical industry has its own 42 MW gas base power plant

3) Food and Beverage Industry

- The rating of frozen pizza making industry and its electricity consumption,
- Total Employees: 90
- Production capacity per day of pizza plant : 70,000 Frozen pizza per day
- Electric Power consumption per Pizza : 1.2 kWh/pizza
- PV Generation capacity : 1 MW

The energy demand per unit can be determined by computing,

Electricity Consumption per Pizza: 70,000 pizzas × 1.2 kWh = 84,000 kWh

Additional Equipment and Facilities: 20% of the pizza's electricity consumption: 84,000 kWh × 0.2 = 16,800 kWh

Total Electricity Consumption for Pizzas: Electricity consumption per pizza + Additional equipment and facilities: 84,000 kWh + 16,800 kWh = 100,800 kWh

Electricity Consumption per Day: Total electricity consumption per day: 100,800 kWh

To convert the electricity consumption from kilowatt-hours (kWh) to megawatt-hours (MWh), divide by 1,000: 100,800 kWh / 1,000 = 100.8 MWh

To determine the average electricity consumption in megawatts (MW), divide by the production duration: 100.8 MWh / 24 hours = 4.2 MW

Therefore, in this example, a food and beverage industrial plant with a capacity to make 70,000 frozen pizzas per day would require approximately 100.8 megawatt-hours (MWh) of electricity per day or have an average power demand of 4.2 megawatts (MW) during continuous production.

4) Textile Industry

Estimated rating for the textile manufacturing plant,

- Total Employees: 100,
- Production capacity per day of material: 16,666 meter
- Yearly Production capacity: 5 million meter
- Electric Power consumption per tonne : 0.45 kWh/meter

To calculate the electricity requirement for manufacturing 5 million meters of textile fabric,

Calculate the total electricity requirement for yearly production: Total electricity requirement = Yearly production * Electricity requirement per meter
 Total electricity requirement = 5,000,000 meters * 0.45 kWh
 Total electricity requirement = 2,250,000 kWh

Convert the yearly electricity requirement to MWh: Total electricity requirement (MWh) = Total electricity requirement (kWh) / 1,000
 Total electricity requirement (MWh) = 2,250,000 kWh / 1,000
 Total electricity requirement (MWh) = 2,250 MWh

Daily production = Yearly production / Number of working days
 Daily production = 5,000,000 meters / 300 days
 Daily production = 16,666.67 meters

Calculate the daily electricity requirement in MWh: Daily electricity requirement (MWh) = Total electricity requirement (MWh) / Number of working days
 Daily electricity requirement (MWh) = 2,250 MWh / 300 days
 Daily electricity requirement (MWh) = 7.5 MWh

Calculate the average daily electricity requirement in MW: Assuming 14 hours of operation per day: Average daily electricity requirement (MW) = Daily electricity requirement (MWh) / 14 hours
 Average daily electricity requirement (MW) = 7.5 MWh / 14 hours
 Average daily electricity requirement (MW) = 0.54 MW

VII. EQUATIONS

All equations related to mathematical modelling and case studies are given in this section.

A. Equation for the Energy Management in Cement Industry:

1. State of Charge of Battery (Storage system) at Cement Industry,

$$SOC_t^{ce} = SOC_{(t-1)}^{ce} + \left(p_t^{cce} * \eta_c^{cce} + \frac{p_t^{dce}}{\eta_d^{ce}} \right) * \Delta t \quad (1)$$

Where,

SOC_t^{ce} = State of charge of battery at time t, (ce- cement industry)

SOC_{t-1}^{ce} = State of charge of battery at time t-1, (ce- cement industry)

p_t^{cce} = Battery Charging power at time t at cement industry (kW)

η_c^{cce} = Battery Charging efficiency at time t at cement industry (%)

η_d^{ce} = Battery Discharging efficiency at time t at cement industry (%)

p_t^{dce} = Battery Discharging power at time t at cement industry (kW)

Δt = time constant

2. Battery charging equation of cement industry storage system,

$$p_t^{cce} * \eta_c^{ce} = p_{wt}^{ce} + U_t^{ce} + p_{pvt}^{ce} - (D_t^{ce}) \quad (2)$$

Conditions,

$$\text{Charging : } p_{wt}^{ce} \rightarrow D_t^{ce} < 3500 \text{ kW}$$

$$\begin{aligned} \text{Charging : } p_{wt}^{ce} + U_t^{ce} &\rightarrow D_t^{ce} \\ &< 3500 \text{ kW} \ \&\& \ t \\ &:: 11:00 \text{ PM to } 4:00 \text{ AM} \end{aligned}$$

$$SOC_{min}^{ce} \leq SOC_t^{ce} \leq SOC_{max}^{ce}$$

$$p_{min}^{cce} \leq p_t^{cce} \leq p_{max}^{cce}$$

Where,

p_{wt}^{ce} = Wind power generation at time t at the cement industry (kW)

U_t^{ce} = Utility power at time t at the cement industry (kW)

p_{pvt}^{ce} = Solar power generation at time t at the cement industry (kW)

D_t^{ce} = Total load at time t at the cement industry (kW)

SOC_{min}^{ce} = minimum limit of state of charge of battery at cement industry

SOC_{max}^{ce} = maximum limit of state of charge of battery at cement industry

p_{min}^{cce} = minimum limit of the charging power at cement industry (kW)

p_{max}^{cce} = maximum limit of the charging power at cement industry (kW)

3. Battery Discharging equation of cement industry storage system,

$$\frac{p_t^{dce}}{\eta_d^{ce}} = SOC_{(t-1)}^{ce} + p_{wt}^{ce} + U_t^{ce} + p_{pvt}^{ce} - (D_t^{ce}) \quad (3)$$

Conditions,

$$\text{Dis - charging : } D_t^{ce} \geq 5000 \text{ kWh}$$

$$p_{min}^{dce} \leq p_t^{dce} \leq p_{max}^{dce}$$

Where,

p_{min}^{dce} = minimum limit of the Discharging power at cement industry (kW)

p_{max}^{dce} = maximum limit of the Discharging power at cement industry (kW)

4. Cement industry load balance equation,

$$\begin{aligned} p_{wt}^{ce} + U_t^{ce} + p_{pvt}^{ce} + \left(\frac{p_t^{dce}}{\eta_d^{ce}} \right) \\ \geq D_t^{ce} + (p_t^{cce} * \eta_c^{ce}) \end{aligned} \quad (4)$$

5. Industrial zone load balance equation

$$\begin{aligned} U_t^z + \left(\frac{p_t^{dz}}{\eta_d^z} \right) + \sum_{j=1}^{45} p_{jt}^{V2Gev} \\ \geq \left(\sum_{i=1}^5 D_{it}^z + \sum_{j=1}^{45} p_{jt}^{G2Vev} \right. \\ \left. + (p_t^{cz} * \eta_c^z) \right) * \Delta t * k \end{aligned} \quad (5)$$

Where,

U_t^z = Utility power at time t at Industrial Zone (kW)

p_t^{dz} = Battery Discharging power at time t at Industrial Zone (kW)

η_d^z = Battery Discharging efficiency at time t at Industrial zone (%)

p_{it}^{V2Gev} = Electrical vehicle V2G power at time t, j = number of electric vehicle chargers

D_{it}^z = Load at time t at industrial zone (kW), i = number of industries

p_{it}^{G2Vev} = Electrical vehicle G2V power at time t, j = number of electric vehicle chargers

η_c^z = Battery charging efficiency at Industrial zone (%)

p_t^{cz} = Battery charging power at time t at cement industry (kW)

Δt = time constant

k = variable constant

6. State of Charge of Battery (Storage system) at Industrial zone,

$$SOC_t^z = SOC_{(t-1)}^z + \left(p_t^{cz} * \eta_c^z + \frac{p_t^{dz}}{\eta_d^z} \right) * \Delta t \quad (6)$$

Where,

SOC_t^z = State of charge of battery at time t, (z - industrial zone)

SOC_{t-1}^z = State of charge of battery at time t-1, (z - industrial zone)

7. Battery charging equation of Industrial zone storage system

$$p_t^{cz} * \eta_c^z = U_t^z + \sum_{j=1}^{45} p_{jt}^{V2Gev} - \left(\left(\sum_{i=1}^5 D_{it}^z + \sum_{j=1}^{45} p_{jt}^{ev} \right) * \Delta t * k \right) \quad (7)$$

Conditions,

Charging with: $U_t^z \rightarrow t :: 11:00 PM to 4:00 AM$

$$SOC_{min}^z \leq SOC_t^z \leq SOC_{max}^z$$

$$p_{min}^{cz} \leq p_t^{cz} \leq p_{max}^{cz}$$

Where,

SOC_{min}^z = minimum limit of state of charge of battery at industrial zone

SOC_{max}^z = maximum limit of state of charge of battery at industrial zone

p_{min}^{cz} = minimum limit of the charging power at industrial zone (kW)

p_{max}^{cz} = maximum limit of the charging power at industrial zone (kW)

8. Battery Discharging equation of Industrial zone storage system,

$$\frac{p_t^{dz}}{\eta_d^z} = SOC_{(t-1)}^z + U_t^z + \sum_{j=1}^{45} p_{jt}^{V2Gev} - \left(\left(\sum_{i=1}^5 D_{it}^z + \sum_{j=1}^{45} p_{jt}^{ev} \right) * \Delta t * k \right) \quad (8)$$

Conditions,

Dis - charging: $t :: 11:00 AM to 6:00 AM$

$$p_{min}^{dz} \leq p_t^{dz} \leq p_{max}^{dz}$$

Where,

p_{min}^{cz} = minimum limit of the charging power at industrial zone (kW)

p_{max}^{cz} = maximum limit of the charging power at industrial zone (kW)

VIII. ALGORITHMS

This section incorporates all the algorithms pertaining to Smart Network Management. Four algorithms have been included: the Cement Industry Energy Flow Algorithm, the Energy Storage Algorithm of Cement Industry, the Energy Flow at the Zone, and the Energy Storage System at the Industrial Zone Algorithm. Each algorithm is accompanied by equations, which are provided in above Section of the document. All the algorithms in the Industrial zone include individual industry algorithms and connected to each other.

A. Cement Industry Energy flow

Figure 7 illustrates the energy flow within the cement industry. The industry's maximum load is approximately 5500 kW. Due to which the algorithm begins by checking the real-time load. If the load exceeds 5000 kW, the algorithm evaluates the availability of solar and wind generation to meet the demand. Alternatively, it assesses the battery status in conjunction with solar and wind power to fulfill the load requirements. If none of

these combinations are sufficient, the remaining power is sourced from the utility system. The algorithm verifies the entire process every 15 minutes, in alignment with the 15-minute load change interval specified by ENTSO-E.

Conversely, if the load is below 5000 kW, the algorithm rechecks the load condition to determine if it is lower than 3500 kW. If the load is below 3500 kW, the algorithm examines the state of charge of the battery. If the battery's state of charge exceeds the minimum threshold, the algorithm attempts to supply the load using battery storage. If the battery status is insufficient, the algorithm combines renewable energy resources to meet the load requirements. If combination is inadequate, then utility power supply become mandatory to fulfill the load. Additionally, if the battery status falls below the minimum threshold, the algorithm initiates the battery charging algorithm to commence charging.

When load is higher than the combination of all sources then alert will signal will be sent to the zone controller and that will send information to the utility increase the supply and provide information for the forecasting.

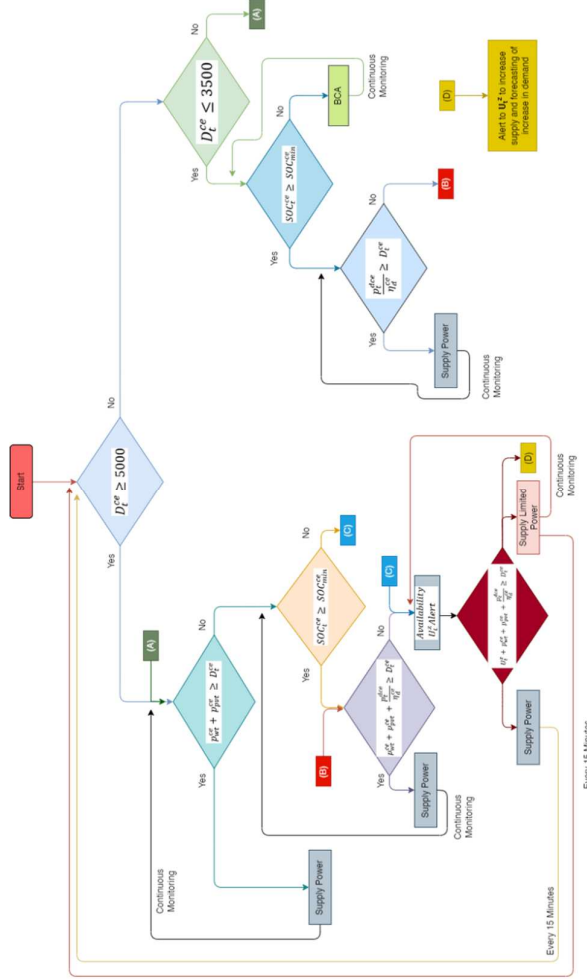


Fig. 7. Cement Industry Energy Flow Algorithm

B. Cement Industry Battery Storage system

Figure 8 illustrates the algorithm specifically designed to charge the battery within the cement industry. The input value for this algorithm is derived from the available signal provided by the energy flow algorithm. It commences by assessing the time dependency aspect. If the time falls between 11:00 PM to 4:00 AM, power will be supplied using a combination of wind, utility, and PV until the battery reaches the maximum state of charge. The algorithm consistently monitors the timing and evaluates the feasibility of power supply from various sources. If the time falls outside the specified range, the battery will only be charged using wind power generation. Otherwise, the battery will not be charged.

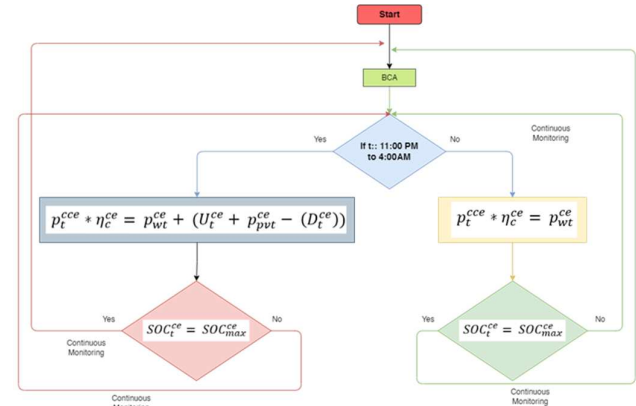


Fig. 8. Cement Industry Battery Charging Algorithm

C. Energy flow in Industrial Zone

Figure 9 illustrates the energy flow algorithm implemented in an industrial zone. The algorithm is time-dependent and takes into account the current time for decision-making.

When utility power availability is checked, it is given the highest priority. If utility supply is absent, the algorithm verifies whether the industrial storage has reached the minimum state of charge. Simultaneously, an alert is sent to all industries notifying them of the power absence. If the battery has sufficient capacity, power is supplied accordingly. Otherwise, the algorithm examines the possibility of utilizing electric vehicles (EVs) to charge the battery storage system. In such cases, no power is supplied to the industries directly. If the battery storage is adequate to fulfill the load, it is supplied up to the limit. Additionally, the algorithm considers the vehicle-to-grid (V2G) condition and, in conjunction with battery storage, supplies power. If none of these combinations are sufficient, the process is halted until a condition is fulfilled. Continuous monitoring is conducted during the absence of utility power to determine power readiness.

When utility power becomes available, the algorithm compares the timing with the predefined 9:00 PM to 4:00 AM limit to account for zone battery charging. If the time falls outside this range, the algorithm assesses the battery's state of charge. Following this, it checks the V2G condition and supplies power if the available power is sufficient. If the power is insufficient, a combination of utility power and battery supply is used to

requirements, power consumption patterns, and load fluctuations throughout different time periods are important factors to analyze.

By accurately assessing power demand and the capacity of the utility grid, industries can ensure a reliable power supply. Understanding the limitations of a normal grid scenario can help identify the need for additional energy storage solutions or EV infrastructure to enhance energy resilience and operational reliability.

Case 2: Grid with EV Infrastructure but Without Industrial Zone Storage - Adding Variable Load in Chemical and Cement Plants

A total of 45 charging stations and 151 car charging (G2V) and discharging (V2G) conditions were assessed in this analysis. The focus was on evaluating the performance and behavior of these charging stations and the associated G2V and V2G processes. In this case, we examine a scenario where EV infrastructure is available, but there is no on-site energy storage within the industrial zone. Additionally, there are variable load fluctuations in chemical and cement plants, resulting in a temporary 2-5% load increase for a few hours during daytime.

The objective is to determine the extra power that needs to be supplied by the utility to meet the demands of the EV infrastructure and the variable load. We also assess the feasibility and manageability of this additional power requirement. To evaluate the power requirements, we analyze the total power demand of the EV infrastructure and the variable load during the specified hours. By comparing this demand with the capacity of the utility grid, we can determine if the grid can effectively handle the extra power load. Considering factors such as charging patterns, load fluctuations, and grid capacity, we assess the manageability of the additional power requirements. This evaluation ensures that the grid can handle the increased load without compromising stability and reliability.

If the power demand exceeds the grid capacity, alternative solutions such as energy storage systems or grid upgrades may be necessary to effectively meet the requirements and maintain a stable power supply.

Case 3: Grid with EV Infrastructure and Industrial Zone Storage - Grid Management Scenario

Half hour power for all industry together is 32 MWh

Current battery percentage $82\% = 32 * 82\% = 26.24 \text{ MWh} = 26240 \text{ KWh}$

Charging efficiency: 85%

Discharging efficiency: 77%

8000-watt rating for charge and discharge

Calculation, one battery size 1 MWh and its rating is 250 kW with response time 500 milli second

Total rating of 32 battery = $250 * 32 = 8000 \text{ kW}$

In this case, we focus on a scenario where EV infrastructure and on-site energy storage are available within the industrial zone, enabling comprehensive grid management.

The objective is to develop an effective strategy for managing the grid, considering the power requirements of the EV infrastructure and the industrial zone. By integrating these components, we can optimize renewable energy utilization, ensure grid stability, and improve energy efficiency.

The algorithm assesses power demands from the EV infrastructure and the industrial zone, considering factors like charging patterns, load fluctuations, and energy storage levels. By analyzing these parameters, the algorithm determines the optimal allocation of power from renewable sources, storage, and the utility grid.

Real-time monitoring enables the algorithm to make adjustments based on grid conditions, ensuring stability and minimizing reliance on non-renewable sources. Additional considerations include energy storage capacity, renewable energy availability, and contractual agreements with the utility provider.

Implementing this grid management scenario allows industries to maximize renewable energy use, reduce grid dependency, and optimize energy consumption. This approach promotes sustainability, contributing to a greener and more efficient energy ecosystem

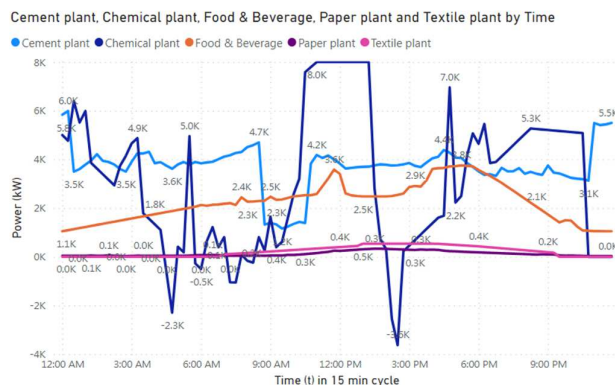


Fig. 11. Industrial load

The load pattern of five industries is presented in Fig. 11, highlighting the distribution of power consumption. The chemical and cement industries exhibit the highest power demands, with the chemical industry reaching a peak load of 8000 kW or 8 MW during periods of high demand. Notably, a fault occurred in the chemical industry, causing a drop in power supply, which was subsequently compensated by the gas power generation within the industry. The surplus power generated by the gas plant was then supplied to the local industrial grid, benefiting other industries.

Furthermore, the paper and textile industries demonstrate the lowest power loads among the five industries, indicating their

relatively lower energy requirements. Interestingly, during peak demand periods, the cement industry experiences a significant decrease in power consumption, suggesting possible load-shifting practices or adjustments in production processes. Lastly, the food and beverage industry maintains an average power consumption level compared to the other industries, reflecting its moderate energy demands.

Analyzing these load patterns allows for a comprehensive understanding of each industry's power usage dynamics. This information is crucial for implementing efficient energy management strategies tailored to the specific needs and characteristics of each industry.

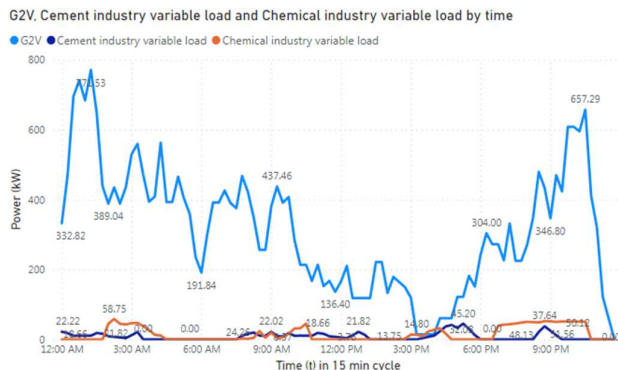


Fig. 11. G2V, Cement and Chemical industry variable load

Figure 11 illustrates the variable load profiles of the Cement and Chemical industries, along with the Grid-to-Vehicle (G2V) power supply. Throughout the day, G2V power is consistently available, effectively meeting the variable load requirements. The maximum power supplied through G2V reaches 771.53 kW, ensuring ample capacity to accommodate the fluctuating demands.

Upon evaluating the power availability for the entire day, the average G2V power hovers around 300 kW. This indicates a steady and reliable power supply that can effectively support the operational needs of the industries. Notably, the variable load requirements are relatively moderate, with the highest recorded variable load of 58.75 kW observed in the Chemical industry.

These findings demonstrate that the G2V power supply is well-suited to meet the dynamic energy demands of the Cement and Chemical industries. With sufficient power availability and manageable variable load profiles, industries can optimize their operations and effectively utilize the G2V power supply to support their processes and contribute to sustainable energy consumption practices.

X. RESULTS

I have thoroughly analyzed each case and incorporated the findings into the corresponding graphs, accompanied by comprehensive descriptions. The results obtained from the evaluation provide valuable insights into the energy dynamics and performance of the system under different scenarios.

Case 1 : Normal Grid: No industrial zone storage and No EV infrastructure

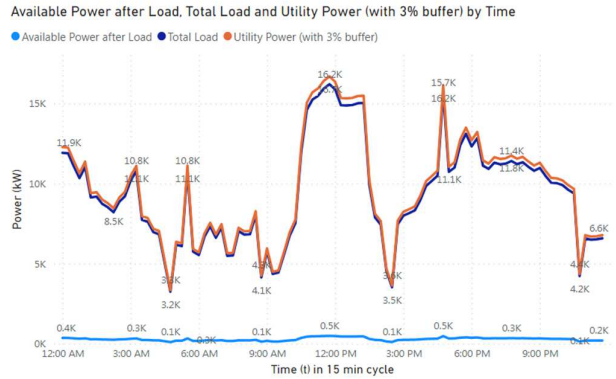


Fig. 12. Result: Normal Grid: EV, Utility Supply and available power

As depicted in Fig. 12, it is evident that under normal operating conditions, the utility power is sufficient to meet the load requirements in the absence of variable load and EV infrastructure. However, it is crucial to highlight that the maximum load observed at 11:45 AM reaches a substantial value of 16.2 MW, indicating a significant power demand during peak hours. Managing such high load levels poses challenges for the utility in maintaining a stable supply. Also, electricity prices are very high during peak hours which is not compatible for the industries.

The analysis of these observations provides valuable insights into the load dynamics and emphasizes the importance of effective load management strategies, considering both peak load conditions and the limitations of the utility's power supply. Such insights can inform decision-making processes and assist in the development of robust energy management plans to optimize load fulfillment and enhance overall system efficiency

Case 2 : Grid with EV Infrastructure but Without Industrial Zone Storage - Adding Variable Load in Chemical and Cement Plants

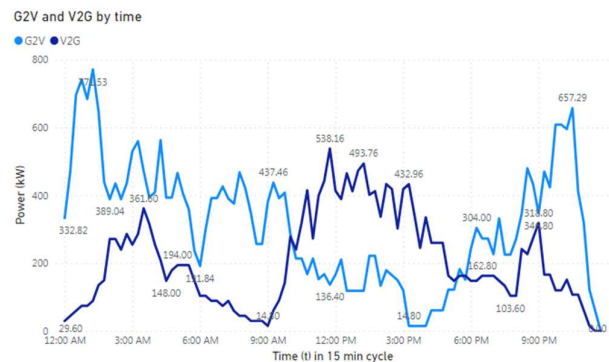


Fig. 13. Result: G2V and V2G

The figure 13. illustrates the operation of G2V (grid-to-vehicle) and V2G (vehicle-to-grid) processes. One noticeable trend is the significant number of V2G operations during peak

demand hours. This indicates that vehicles are supplying a substantial amount of power back to the grid, reaching a peak value of approximately 400 kW. On the other hand, the highest load observed in the G2V condition occurs around midnight, with a value of around 700 kW. This indicates a well-balanced interaction between vehicle charging and discharging, contributing to efficient energy management. This analysis provides valuable insights into the dynamics of G2V and V2G operations and helps in optimizing the utilization of vehicle-to-grid capabilities.

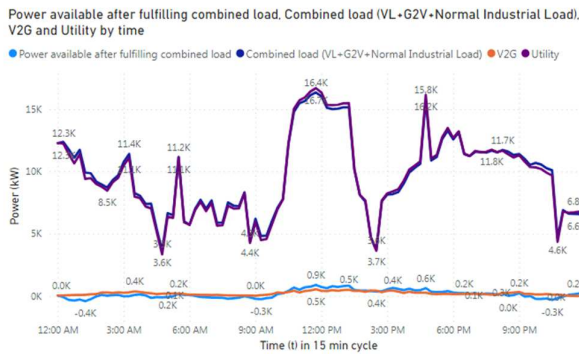


Fig. 14. Result: Normal Grid: Combined load (VL+G2V+Normal Industrial Load), V2G and Utility

The analysis of Fig. 14 reveals that the combination of utility grid power and V2G operations falls short in supplying adequate power to both industries and G2V processes. It is evident that there are multiple instances throughout the day where the available power is negative, indicating an insufficiency in power supply. Particularly, there are three prominent negative peaks observed at 1:30 AM, 9:00 AM, and 11:15 AM, reaching values of approximately -0.4 kW, -0.3 kW, and -0.3 kW, respectively. These fluctuations highlight the limitations of V2G technology in effectively balancing the grid as initially anticipated. Further analysis and improvements are required to address these challenges and enhance the reliability and efficiency of the power supply system.

Case 3 : Grid with EV Infrastructure and Industrial Zone Storage - Grid Management Scenario

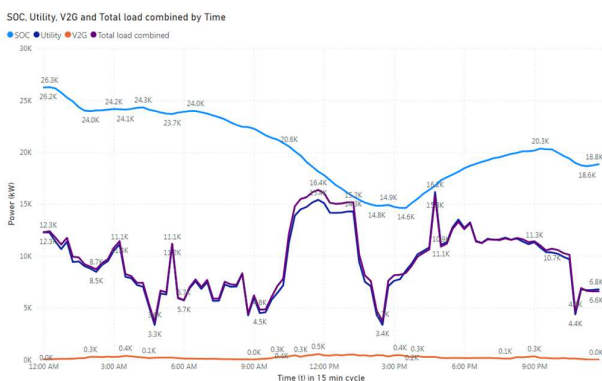


Fig. 14. Result: Normal Grid: Combined load (VL+G2V+Normal Industrial Load), V2G, Energy storage system at zone and Utility

In Case Study 3, an Energy storage system has been incorporated at the Industrial zone to ensure a reliable power supply during peak demand conditions. As depicted in Figure 14, the graph highlights the significance and capacity of the system in maintaining grid stability. At the beginning of the day, the battery State of Charge (SOC) was recorded at 26.3 kW, which gradually decreased to 14.6 kW during the peak demand time. However, by the end of the day, the battery SOC recovered to 18.8 kW. This clear trend showcases the effectiveness of the algorithm in efficiently managing the grid balance and optimizing the utilization of the energy storage system. The results affirm the value of incorporating such systems to mitigate power fluctuations and ensure a reliable power supply throughout the day. Pick demand with value of 16.4 k can be handle also with support of storage system and it will give less burden on the utility. Many other picks also noticeable in daytime but algorithm configured in such way that it will give priority during the pick demand time.

XI. CONCLUSION

All in all, I have tried to put the actual fact and status of our present grid. The global electricity load is rapidly rising by new innovative technologies, if we will have to meet the demand of this virtual world concept, digitalization in a transaction of money, mobile working, online data processing, and more importantly complete security solutions then we will need most reliable electricity. Not only this, to cope with the reduction in global carbon emissions, people will have to stand with government and utility by accepting green energy generation systems such as PV and micro-wind.

I have demonstrated how global industrialization changed with respect to time. The trading load is challenging the world to move forward and go autonomous. I have also touched on a very important topic that scientists and researchers will have to more focus on the demand graph, as this steep crypto growth and metaverse growth was not fully counted a few years back when actual prediction was taken.

After evaluating all the cases, including the normal grid, the normal grid with EV infrastructure and variable load, and the case with an energy storage system at the industrial zone, it is evident that the implementation of an energy storage system yields the most favorable results in terms of grid stability, power supply reliability, and load management.

In the first case, where a normal grid is considered without any additional infrastructure or energy storage, the power supply is solely dependent on the utility. While it can meet the regular industrial load, challenges arise during peak demand periods. The load analysis reveals instances where the available power is insufficient, leading to unfulfilled demand and potential grid instability.

In the second case, the introduction of EV infrastructure along with variable loads adds complexity to the grid management. Although EVs can provide additional power through V2G operations, the system still faces limitations in meeting the industrial load requirements during peak hours. The power balance remains a challenge, with negative power peaks and a reliance on utility supply to compensate for the shortfall.

However, in the third case, the incorporation of an energy storage system at the industrial zone proves to be the most effective solution. The energy storage system, combined with smart network management algorithms, addresses the challenges faced in the previous cases. It optimizes power supply by intelligently utilizing the stored energy, considering factors such as load patterns, variable loads, and EV charging/discharging operations.

The graph analysis demonstrates the successful performance of the energy storage system in maintaining a stable grid during peak demand periods. The algorithm effectively manages the battery state of charge, ensuring a consistent power supply and reducing reliance on the utility. The integration of EV infrastructure further enhances the system's flexibility and contributes to load balancing.

Overall, the case with an energy storage system at the industrial zone emerges as the best solution. It offers improved grid stability, reliable power supply, and efficient load management. The successful operation of the system, as depicted in the graphs and analysis, highlights the effectiveness of energy storage in enhancing industrial energy management.

By implementing an energy storage system and employing smart network management algorithms, industrial sectors can optimize their power consumption, reduce dependency on the utility grid, and contribute to a more sustainable energy ecosystem. The integration of EV infrastructure further enhances the system's capabilities, making it a comprehensive and recommended approach for industrial energy management.

ACKNOWLEDGMENT

This work was supported by the Slovak Research and Development Agency, under the contracts APVV-19-0576, APVV-21-0312 and SK-UA-21-0024 and the Slovak Academy of Sciences, under the contract VEGA 1/0757/21.

REFERENCES

- [1] Adrian Dima, Short history of manufacturing: from Industry 1.0 to Industry 4.0, blog-article, 25/01/2021, <https://kfactory.eu/short-history-of-manufacturing-from-industry-1-0-to-industry-4-0/>.
- [2] What is Industry 4.0 and 5.0? Your Guide to Understanding Industry 4.0 and 5.0, <https://gesrepair.com/industry-4-and-5/#industry>.
- [3] Patel, S. (October 1, 2017). History of Power: The Evolution of the Electric Generation Industry. From Power Mag.
- [4] Uses of Coal. Retrieved from world coal: <https://www.worldcoal.org/coal/uses-coal>
- [5] Assessment of Demand Response and Advanced Metering (Docket AD06-2-000)" (PDF). United States Department of Energy: 20. Retrieved 2008-11-27. Federal Energy Regulatory Commission staff report (August 2006).
- [6] Retrieved from: Wikimedia: <https://commons.wikimedia.org/wiki/Electricity>.
- [7] Berger, L. T., & Iniewski, K. e. (April 2012). Smart Grid - Applicacions, Communications and Security. ISBN 978-1-1180-0439-5. John Wiley and Sons.
- [8] European Energy Industry Investments. (2017., February). Retrieved from www.europarl.europa.eu: [http://www.europarl.europa.eu/RegData/etudes/STUD/2017/595356/IPO_L_STU\(2017\)595356_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2017/595356/IPO_L_STU(2017)595356_EN.pdf)
- [9] L.berkani, A. Hamidat, F.fodhil, O.nadjemi, Z. Alliche, Feasibility study of grid-connected photovoltaic systems in small industrial enterprises, 978-1-7281-2220-5/19/\$31.00 ©2019 IEEE
- [10] Eroshenko S.A., Samoylenko V.O., Pazderin A.V., Renewable Energy Sources for Perspective Industrial Clusters Development, 016 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), 978-1-5386-6410-0/18/\$31.00 ©2018 IEEE
- [11] Tynchenko V.V., Solovyov E.A., Tynchenko S.V., Design of Grid Systems to Solve Complex Industrial Engineering Problems, 2016 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), 978-1-5090-1322-7/16/\$31.00 ©2016 IEEE
- [12] Wei Luo, Youzhong Miao, Weimin Lei, Analysis of Distribution Network Planning of Industrial Park Under Incremental Distribution Businesses Opening, 2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC 2018), 978-1-5386-1803-5/18/\$31.00 ©2018 IEEE
- [13] Haroon Rasheed, Farah Haroon, Nandana Rajatheva, Pulse Position Modulation Based Energy Detection for Smart Grid Communication, 2018 the 6th IEEE International Conference on Smart Energy Grid Engineering, 978-1-5386-6410-0/18/\$31.00 ©2018 IEEE.
- [14] Chenwei Jiang, Fushuan Wen, Yusheng Xue, Fei Chen, Yikai Sun, Lijun Zhang, Optimal power management strategy for Industrial user based on the state task network considering user preferences, 978-1-6654-4119-3/21/\$31.00
- [15] Gang Wang, Xiaokui Zang, Chen Yang, Yichen Xu, Peng Nan, Leijiao Ge, Power Optimization Strategy for Typical Industrial Users under the Condition of Time-of-use Pricing, 2019 IEEE PES Innovative Smart Grid Technologies Asia, 978-1-7281-3520-5/19/\$31.00 ©2019 IEEE.
- [16] Haomiao Wang, Nan Hu, Shuran Jia, Bo Xu, and Dongliang Guo, An Analysis From the Perspective of Direct Power-Purchase for Industrial Users: Should the Power Grid Company Implement Incentive-Based Demand Response Management?, DOI 10.1109/ACCESS.2021.3081631©2021IEEE

Modifikácia elektrických a magnetických vlastností amorfných feromagnetických mikrodrôtov

Simeon Samuhel

Katedra fyziky

Fakulta elektrotechniky a informatiky, TUKE

Košice, Slovensko

simeon.samuhel@tuke.sk

Jozef Onufer

Katedra fyziky

Fakulta elektrotechniky a informatiky, TUKE

Košice, Slovensko

jozef.onufer@tuke.sk

Abstract— Tento príspevok sa zaoberá štúdiom sklom pokrytých amorfných mikrodrôtov. Je prezentovaná nová experimentálna metóda na štúdium dynamiky jednej doménovej steny v amorfných bistabilných mikrodrôtach v oblasti rýchlo sa meniaceho magnetického poľa. Keď je pohybujúca sa stena nútená zmeniť svoje parametre vo veľmi krátkom čase, pozorujeme zmeny jej rýchlosti analýzou indukovaných signálov v dlhej snímačej cievke. Tieto zmeny doménovej steny sú odlišné pre rýchly nárast a pokles intenzity magnetického poľa. Získané výsledky podporujú myšlienku, že doménová stena v bistabilných mikrodrôtach nie je pevný objekt, ale mení svoje parametre (napr. rýchlosť, tvar, rozmer steny a pod.) v závislosti od použitého magnetického poľa. V príspevku je tiež analyzovaný vplyv mechanického namáhania na vlastnosti mikrodrôtov. Účinky vplyvu aplikovaného axiálneho ťahového napätia sa prejavili ako dodatočné vrcholy v indukovaných signáloch z dlhej snímačej cievky. Tie predstavujú lokálne zvýšenie rýchlosti doménovej steny. Získané výsledky prispievajú k rozvoju poznatkov v oblasti sklom pokrytých amorfných mikrodrôtov a poskytujú podklady pre ďalší výskum a vývoj v tejto oblasti.

Abstract- This paper deals with a study of glass-covered amorphous microwires. A new experimental method is presented to study the dynamics of a single domain wall in amorphous bistable microwires in a region of rapidly changing magnetic field. When a moving domain wall is forced to change its parameters in a very short time, we observe changes in the speed of the wall on the induced signals in long pick-up coil. These domain wall changes are different for the rapid rise and fall of the magnetic field strength. The obtained results support the idea that the domain wall in bistable microwires is not a fixed object, but changes its parameters (eg wall, shape, dimension velocity etc.) depending on the applied magnetic field. The contribution also analyzes the influence of mechanical stress on the properties of microwires. The effects of the applied axial tensile stress were manifested as additional peaks in the induced signals from the long pickup coil. These represent a local increase in the velocity of the domain wall. The obtained results contribute to the development of knowledge in the field of glass-covered amorphous microwires and provides the basis for further research and development in this area.

Keywords—Amorphous ferromagnetic microwire, domain wall, domain wall velocity, bistable behaviour, anisotropy, magnetization reversal.

ÚVOD

V súčasnej dobe je vývoj mikro a nanotechnológií kľúčovým faktorom pri dosahovaní pokroku vo viacerých odvetviach, ako je napríklad elektronika, senzorika a medicína [1]-[14]. V rámci týchto odvetví je neustálym cieľom hľadať nové materiály a štruktúry s vylepšenými vlastnosťami, ktoré by umožnili vývoj nových a pokročilejších zariadení. V tejto súvislosti je jednou z možností štúdium sklom pokrytých amorfných mikrodrôtov ako perspektívneho materiálu s vysokým potenciálom.

Amorfné magnetické sklom pokryté mikrodrôty pritiahli značnú pozornosť kvôli svojim zaujímavým magnetickým vlastnostiam a ich potenciálu pri využití v senzorových zariadeniach a iných aplikáciách [1-4]. Tieto materiály pozostávajú z kovového jadra s priemerom 0,5–30 μm a skleneného obalu s hrúbkou 1–20 μm .

Amorfné sklom potiahnuté mikrodrôty sú zaujímavé materiály z hľadiska základného výskumu aj technologických aplikácií. Taylor vyvinul priamy výrobný proces na výrobu krátkych sklom potiahnutých vzoriek už v roku 1924 [7]. Túto metódu neskôr (v rokoch 1950 až 1964) vylepšil Ulitovský tak, že je možné realizovať výrobu súvislých sklom potiahnutých mikrodrôtov [8]. V súčasnosti je preto možné vyrobiť veľké množstvá mikrodrôtov pomocou modifikovanej Taylor-Ulitovského metódy [9]. Okrem toho moderné výrobné zariadenia na prípravu sklom pokrytých mikrodrôtov umožňujú pomocou presného ovládania výroby pomocou počítača definovať geometriu mikrodrôtu (priemer kovového jadra d a priemer mikrodrôtu D) [10]. Preto v posledných rokoch štúdium sklom potiahnutých mikrodrôtov pritiahuje značný záujem, a to najmä vďaka ich technickému využitiu, napríklad ako snímačov v rôznych systémoch [11]. Feromagnetické sklom potiahnuté mikrodrôty patria medzi pokrokové magnetické materiály, kvôli ich jedinečnej kombinácii fyzikálnych vlastností (mechanické, magnetické a biokompatibilita) a ich rýchlejšej, lacnej metóde prípravy [9],[12].

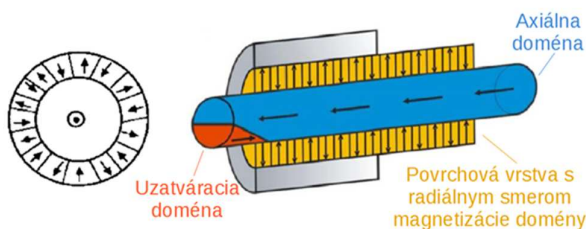
V tomto príspevku je prezentovaná nová experimentálna metóda na štúdium dynamiky doménovej steny (DS) v bistabilných mikrodrôtach v rýchlo sa meniacom magnetickom poli.

ZÁKLADNÉ ROZDELENIE MIKRODRÔTOV PODĽA
MAGNETOSTRIKcie

Významnou charakteristikou, podľa ktorej rozdeľujeme magnetické vlastnosti a správanie magnetických mikrodrôtov, je hodnota magnetostrikcie. Magnetostrikcia priamo súvisí so zložením zliatiny, z ktorej je mikrodrôt vyrobený. V závislosti od hodnoty magnetostrikcie možno amorfné sklom pokryté mikrodrôty rozdeliť do troch skupín:

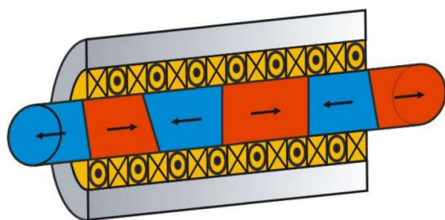
1. mikrodrôty s kladnou magnetostrikciou
2. mikrodrôty s magnetostrikciou blízkou nule
3. mikrodrôty so zápornou magnetostrikciou

Mikrodrôty s kladnou magnetostrikciou potiahnuté sklom na báze Fe patria medzi najvýznamnejšiu skupinu mikrodrôtov. Hysterézná slučka takého mikrodrôtu je perfektné pravouhlá (obr. 4a) a magnetizácia môže mať len dve hodnoty $\pm M_s$ (kde M_s je magnetizácia nasýtenia). Ich doménová štruktúra (obr. 1) pozostáva z jednej veľkej axiálnej domény v strede kovového jadra, ktorá je obklopená vonkajšími doménami s orientáciou radiálnej magnetizácie [13], [14]. Navyše sa na konci mikrodrôtu objavujú malé uzatváracie doménové štruktúry, aby sa zmenšili rozptyľové polia [13]. Premagnetovanie mikrodrôtu nastáva pri jednom veľkom Barkhausenovom skoku, kedy sa zmení smer magnetizácie vnútornej domény na opačný. Proces premagnetovania prebieha pohybom jednej DS pozdĺž celého mikrodrôtu a umožňuje študovať proces šírenia jednej DS.



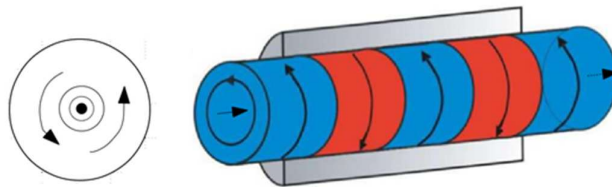
Obr. 1 Doménová štruktúra mikrodrôtu s kladnou magnetostrikciou [15].

Mikrodrôty s malou magnetostrikciou blízkou nule sú typické pre mikrodrôty potiahnuté sklom na báze Co-Fe. Doménová štruktúra (obr. 2) takýchto mikrodrôtov pozostáva z axiálnej doménovej štruktúry obklopenej cirkulárnymi doménami. Hysterézne slučky takýchto mikrodrôtov majú veľmi nízku koercitivitu a vysokú počiatočnú permeabilitu (obr. 4 c).

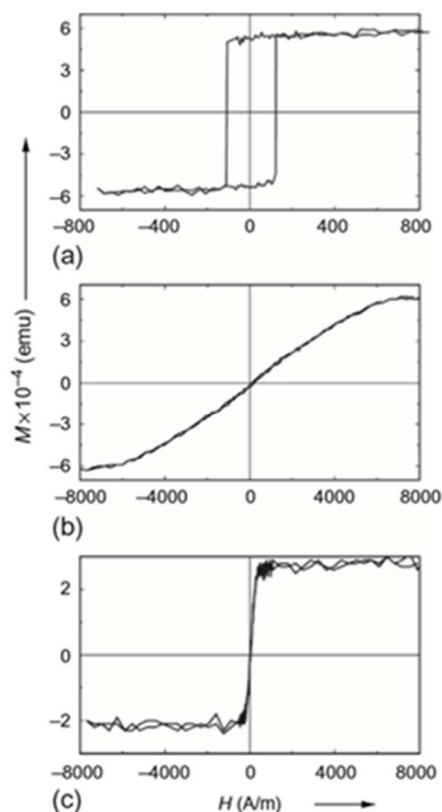


Obr. 2 Doménová štruktúra mikrodrôtov s malou magnetostrikciou blízkou nule [15].

Mikrodrôty so zápornou magnetostrikciou majú os ľahkej magnetizácie v cirkulárnom smere a ich doménová štruktúra sa preto skladá z cirkulárných domén (obr. 1). Proces premagnetovania v axiálnom smere prebieha reverzibilnou rotáciou magnetických momentov vo vnútri doménovej steny (DS). Preto pri pôsobení axiálneho magnetického poľa je pozorovaná takmer lineárna slučka s pomerne nízkou hysteréziou (obr. 4b).



Obr. 3 Doménová štruktúra mikrodrôtu so zápornou magnetostrikciou [15].



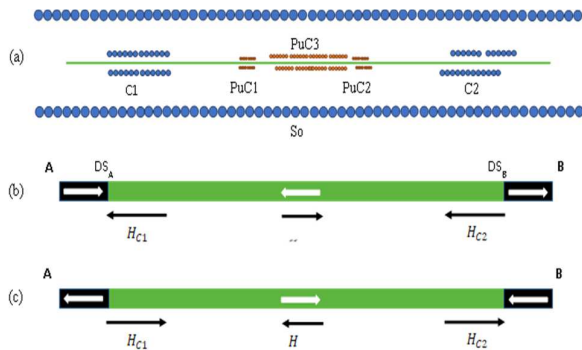
Obr. 4 Hysterézne slučky tenkých sklom potiahnutých mikrodrôtov s kovovým jadrom na báze a) železa ($\lambda_s > 0$) b) kobaltu ($\lambda_s < 0$) c) železo-kobaltu ($\lambda_s = 0$) [16].

EXPERIMENT

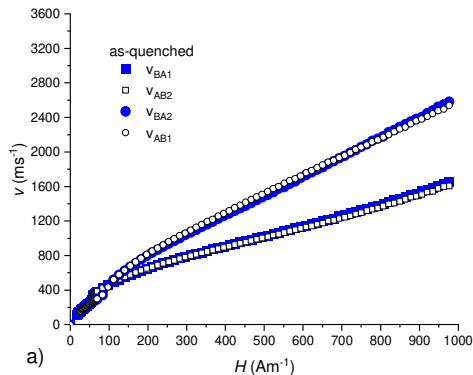
A. Experiment na skúmanie geometrie doménovej steny

Experimentálna zostava je schematicky znázornená na obr. 5(a), pozostáva z troch magnetizačných cievok So, C1 a C2 a z troch snímacích cievok PuC1, PuC2 a PuC3. Dlhý solenoid So

(dlhší ako dĺžka vzorky), ktorý vytvára axiálne pole H_{So} , sa používa v prvom kroku na premagnetovanie vzorky vo zvolenom smere (na definovanie počiatočného magnetického stavu mikrodrôtu). Ďalej sa aplikuje opačné magnetické pole H_{So} , ktoré je dostatočne vysoké na uvoľnenie DS z oboch koncov drôtu (na obr. 5 je A - ľavý koniec a B - pravý koniec). Uvoľnené DS sú zastavené magnetickým poľom blokovacích cievok C1 a C2 (H_{C1} zastaví DS zo smeru A a H_{C2} zastaví DS zo smeru B), ktoré sú aplikované spolu s H_{So} , ale v opačnom smere ($H_{C1} = H_{C2}$, s hodnotou mierne vyššou ale opačného smeru ako hodnota aplikovaného H_{So}). V tejto fáze nastáva situácia, keď sa dve DS nachádzajú v pozíciách, kde súčet poľí H_{So} , H_{C1} , H_{C2} je takmer nulový.



Obr. 5 Experimentálna zostava (a) počiatočný magnetický stav, keď je mikrodrôt zmagnetovaný poľom generovaným solenoidom H_{So} s vypnutými kľúčmi K1 a K2 (b); stav po aplikácii opačného H_{So} dostatočne vysokého na uvoľnenie DS z koncov drôtu a so zapnutými kľúčmi K1 a K2 (c).



Obr. 6 Štyri $v(H)$ závislosti pre vzorku s jednosmerovým efektom.

V záverečnom kroku sa po vypnutí poľa H_{C1} alebo H_{C2} môže šíriť po mikrodrôte jedna DS. Na základe toho vieme získať informácie o procese premagnetovania ako aj o výslednom stave magnetizácie v strednej časti mikrodrôtu. Priemernú rýchlosť DS je možné merať štandardnou Sixtusovou – Tonkovou metódou pomocou dvojice snímacích cievok PuC1 a PuC2. Táto metóda umožňuje merať všetky štyri rýchlosti DS bez manipulácie so vzorkou: pre obe možné orientácie počiatočného

magnetického stavu (obr. 5(b), a 5(c)) rýchlosti DS v_{AB} (blokovaná cievka C1 je vypnutá) a rýchlosti DS v_{BA} (blokovaná cievka C2 je vypnutá). Ak je pre použitú vzorku prítomný jednosmerný efekt, potom sa hodnoty páru rýchlostí DS pre jeden východiskový magnetický stav rovnajú, ale líšia sa od páru rýchlostí získaných pre opačný výsledný stav magnetizácie, ako je možné vidieť na obr. 6.

Jednosmerový efekt znamená, že štyri rýchlosti DS spĺňajú vzťah:

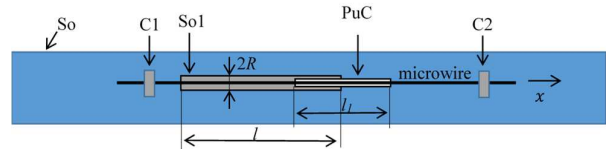
$$v_{AB1}(H) \approx v_{BA2}(H) \neq v_{AB2}(H) \approx v_{BA1}(H), \quad (1)$$

kde A, B (obr. 5) s indexmi označujú konce vzorky (smer šírenia DS) a indexy 1 alebo 2 zodpovedajú typu DS (head-to-head alebo tail-to-tail). DS šíriace sa vyššími rýchlosťami v_{AB1}, v_{BA2} (nižšie v_{AB2}, v_{BA1}) sú ďalej pre jednoduchosť označované ako rýchle DS (pomalé DS).

Informáciu o magnetickom stave centrálnej časti mikrodrôtu je možné získať pomocou signálu indukovaného v snímačej cievke PuC3 vplyvom pohybu DS.

Merania sa uskutočňovali na sklom potiahnutom mikrodrôte $Fe_{77.5}Si_{15}B_{7.5}$ pripravenom Taylor - Ulitovského metódou [8]. Priemer kovového jadra bol cca 15 μm , hrúbka skleneného obalu cca 7,5 μm a dĺžka vzorky 12,5 cm. Mikrodrôt bol s jednosmerovým efektom.

B. Experiment na skúmanie dynamiky DS prechádzajúcej cez oblasť rýchleho sa meniaceho magnetického poľa



Obr. 7 Experimentálna zostava upravená na meranie rýchlej zmeny poľa.

Hore popísaný experiment bol upravený tak aby sme mohli merať priebeh rýchlosti DS prechádzajúcej cez prudkú zmenu axiálneho magnetického poľa. Toto nové experimentálne usporiadanie je schematicky znázornené na obr. 7. Bistabilný mikrodrôt je umiestnený v dvoch koaxiálnych magnetizačných cievkach. Prvý je dlhý solenoid (So), ktorý vytvára pole H_{So} a druhý je tenký solenoid (So1) dĺžky $L = 3$ cm a polomeru $R = 0,258$ mm, ktorý vytvára pole H_1 . Oblasť blízko konca solenoidu So1 je zdrojom rýchlej zmeny magnetického poľa. Na sledovanie pohybu jednej DS po mikrodrôte sa používa tenká jednovrstvová snímačie cievka (PuC) dĺžky $l = 2$ cm a polomeru $r = 0,25$ mm s pravidelne navinutými závitmi, počet $z = 285$ závitov. Koniec solenoidu So1, ktorý je zdrojom rýchlej zmeny magnetického poľa, je umiestnený približne v strede snímačej cievky.

EXPERIMENTÁLNE VÝSLEDKYMERANÉ PRI RÝCHLEJ ZMENE POĽA

Experimentálne údaje boli spracované nasledujúcim spôsobom. Ak sa DS šíri vo vnútri snímačej cievky, indukované napätie môže byť vyjadrené ako

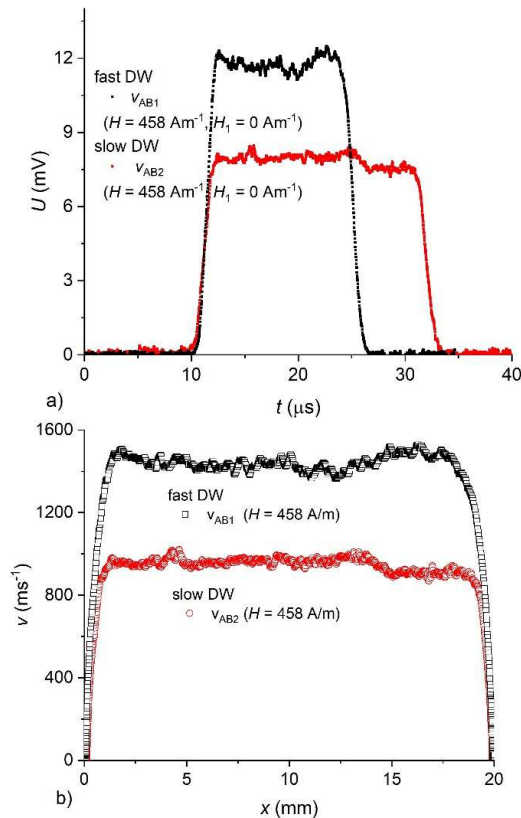
$$u(t) = -v \frac{z}{l} (\Phi_R - \Phi_L) = -v \frac{z}{l} \Delta\Phi \quad (2)$$

kde Φ_L a Φ_R sú magnetické toky cez ľavý a pravý koncový závit cievky.

Charakteristické signály indukovaného napätia na čase v snímačej cievke v dôsledku šírenia jednej DS pozdĺž mikrodrôtu v homogénnom magnetickom poli ($H_1=0$) sú znázornené na obr. 8a. V mikrodrôte je zdrojom rozptylového poľa samotná DS (head-to-head alebo tail-to-tail DS). Rozptylové pole neprispieva k zmene magnetického toku $\Delta\Phi$, vtedy ak je poloha DS dostatočne vzdialená od koncov dlhej snímačej cievky. Možno ju vyjadriť ako

$$\Delta\Phi = 2\mu_0 M_s A \quad (3)$$

kde M_s je saturačná magnetizácia a A je plocha prierezu axiálnej domény.



Obr. 8 Indukované napätiové signály merané vo snímačej cievke PuC (a); polohové závislosti lokálnych rýchlostí (b), pre rýchle (v_{AB1}) a pomalé (v_{AB2}) DS.

Ak je snímačia cievka PuC dostatočne dlhá, aby pohybujúca sa DS bola dostatočne vzdialená od koncov snímačej cievky, hodnotu $\Delta\Phi$ možno získať časovou integráciou signálu indukovaného v dlhej snímačej cievke.

$$\Delta\Phi_{tot} = \int_0^\infty u dt = \frac{z}{l} \Delta\Phi \int_0^l dx = -z\Delta\Phi \quad (4)$$

A potom

$$\Delta\Phi = \frac{-\Delta\Phi_{tot}}{z} \quad (5)$$

Nakoniec, lokálna rýchlosť pohybu DS vo vnútri snímačej cievky môže byť určená pomocou rov. (2) a (5)

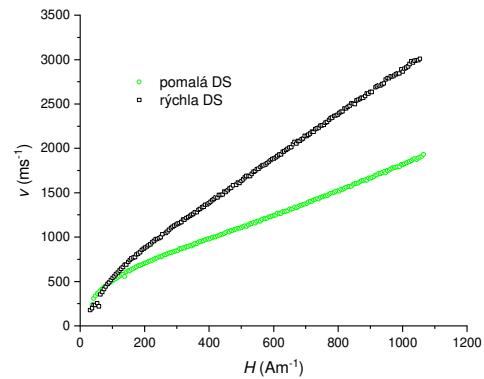
$$v = \frac{-l}{z} \frac{u}{\Delta\Phi} = l \frac{u}{\Delta\Phi_{tot}} \quad (6)$$

Z rov. 6 je možné vypočítať polohu DS ako

$$x = \frac{l}{\Delta\Phi_{tot}} \int_0^t u dt \quad (7)$$

Potom môže byť lokálna rýchlosť DS v polohe x v snímačej cievke PuC určená pomocou rovnice 6 a napätia indukovaného v dlhej snímačej cievke PuC. Vypočítané hodnoty v rýchlosti DS v závislosti na polohe x sú znázornené na obr. 8b. Pre tieto závislosti $v(x)$ boli lokálne rýchlosti v strede snímačej cievky ($x = 10$ mm) určené ako priemerná hodnota 50 bodov.

Spracovanie indukovaných napätiových signálov pri rôznych hodnotách magnetických polí H_{S0} umožnilo určiť závislosti $v(H)$ pre rýchle a pomalé DS (obr. 9).



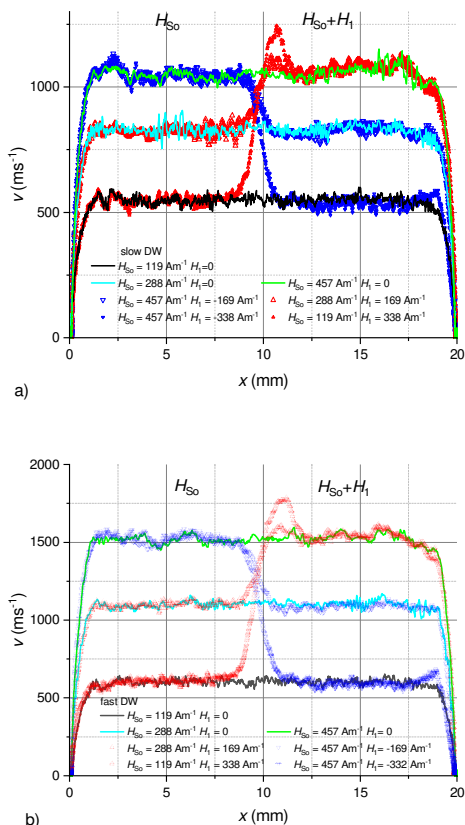
Obr. 9 Závislosť rýchlosti DS od aplikovaného magnetického poľa vypočítaná pomocou rovnice 5 ako priemerná hodnota 50 bodov pri strede závislosti $v(x)$ (obr. 8b). DS vychádza stále z toho istého konca.

V dôsledku magnetického poľa H_1 dochádza k rýchlej zmene magnetického poľa v oblasti na konci solenoidu So1, ktorý je zároveň v strede snímačej cievky PuC.

Ak je magnetické pole H_1 orientované v opačnom smere vzhľadom na magnetické pole H_{S0} , magnetické pole v oblasti vzorky umiestnenej v solenoidu So1 je nižšie ($H_{S0}-H_1$) a na konci solenoidu So1, ktorý sa nachádza v strede PuC, sa magnetické pole rýchlo zvyšuje pre DS, ktorá sa pohybuje v smere osi x (obr. 7). Ak sú magnetické polia H_{S0} a H_1 orientované rovnakým smerom ($H_{S0}+H_1$), potom pole na tomto mieste rýchlo klesá. V dôsledku toho je pozorované zvýšenie alebo zníženie rýchlosti v závislostiach $v(x)$ znázornené na obr. 10. Ako je možné vidieť na týchto obrázkoch, vrchol sa pozoruje, keď dôjde k zvýšeniu rýchlosti DS.

Pre danú polohu steny x možno vypočítať hodnotu ($H_{S0} + H_1$) a pomocou experimentálnych závislostí $v(H)$ na obr. 9 určiť zodpovedajúce rýchlosti. Závislosti $v(x)$ získané týmto spôsobom (nazývané ideálne a vynesené prerušovanou čiernou

čiarou na obr. 11) zodpovedajú prípadu, keď sa DS pohybuje bez zotrvačných účinkov, alebo prípadu, keď má DS dostatok času na prispôsobenie sa zmene magnetického poľa, v ktorom sa pohybuje. Ako je možné vidieť na obr. 10 reálne namerané krivky získané z experimentov majú rôzne tvary.



Obr. 10 Závislosť rýchlosti DS od jej polohy v PuC pre oblasť zmeny poľa medzi tromi úrovňami poľí. Tieto závislosti sú znázornené pre pomalé DS (a) a pre rýchle DS (b).

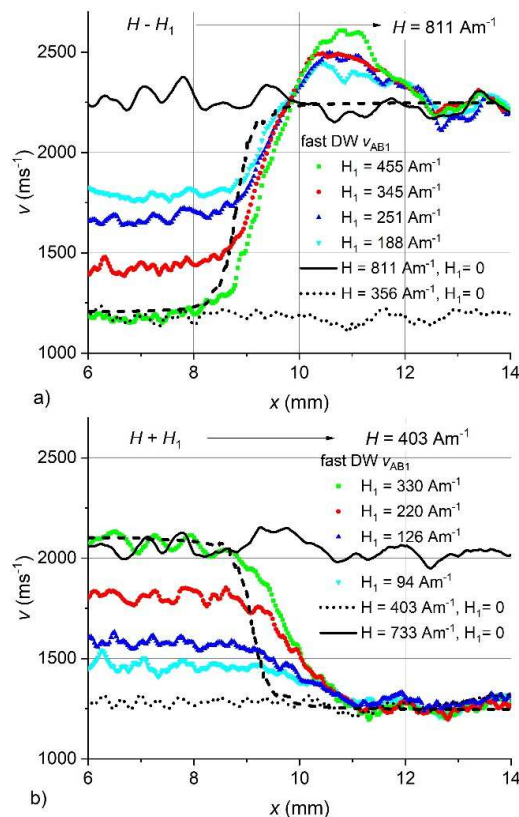
V tomto experimente sa magnetické pole mení zo svojich počiatočných na konečné hodnoty a tiež spôsobuje zmenu rýchlosti DS medzi dvoma hodnotami. Počas nárastu magnetického poľa (z nižšej na vyššiu hodnotu) môže byť rýchlosť DS dokonca vyššia ako jeho konečná (ustálená) hodnota (vrchol uvedený vyššie). Keď dôjde k poklesu magnetického poľa (z vyššej na nižšiu hodnotu), analogický účinok na rýchlosť DS nie je pozorovaný.

INDUKOVANÉ NAPÄTOVÉ VRCHOLY

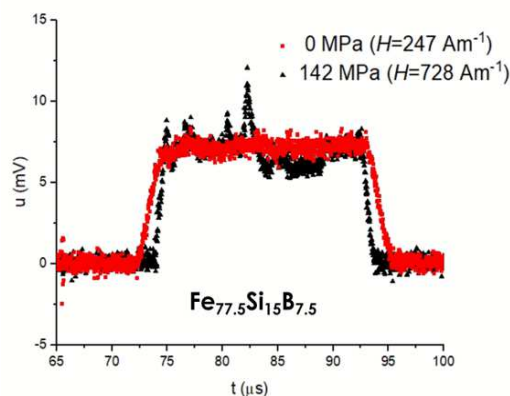
Tiež sme ďalej chceli zistiť ako aplikované ťahové napätie ovplyvňuje geometriu DS, preto sú ďalej prezentované výsledky meraní s experimentálnou zostavou na obr. 7 bez rýchlej zmeny poľa ale s ťahovým napätím aplikovaným na vzorku v axiálnom smere.

Porovnanie indukovaných napäťových signálov zo snímačej cievky môžeme vidieť na obr. 12. Porovnaním čiernej krivky meranej s aplikovaným ťahovým napätím a červenej krivky bez

ťahového napätia, ktoré pôsobí na vzorku ukazuje, že s aplikovaným ťahovým napätím sa objavujú nové vrcholy na indukovanom napäťovom signáli. Je zaujímavé, že amplitúdy niektorých vrcholov sú porovnateľné s priemernou hodnotou signálu. Tento jav je sprevádzaný výrazným rozptylom nameraných hodnôt v závislostiach $v(H)$ (merané pomocou Sixtus-Tonksovej metódy [17]) (obr. 13).

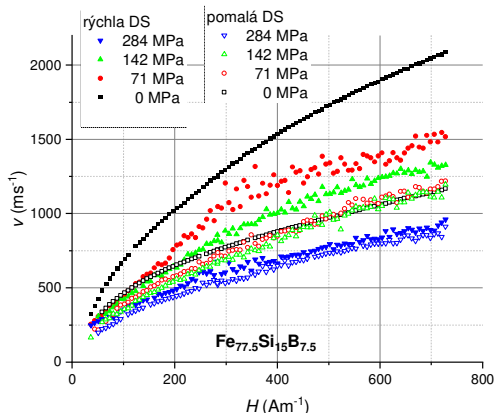


Obr. 11 Závislosť rýchlosti DS od jej polohy v PuC pre (a) rastúcu a (b) klesajúcu zmenu poľa pre rýchlu DS. Závislosti $v(x)$ znázornené prerušovanými čiernymi čiarami popisujú zmenu rýchlosti bez zotrvačnej hmotnosti.

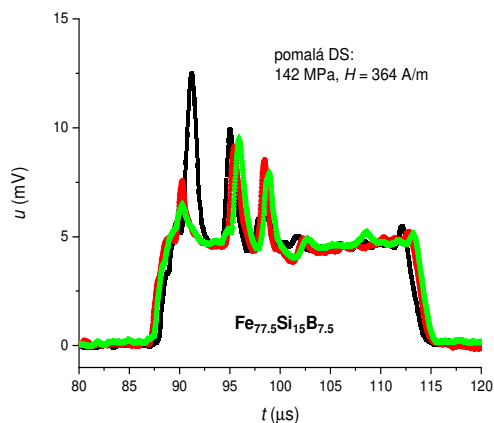


Obr. 12 Merané signály zo snímačích cievok PuC2 bez aplikovaného ťahového napätia (červená) a s aplikovaným ťahovým napätím (čierna).

Prezentované experimentálne výsledky namerané na mikrodrôtoch s jednosmerným efektom [18] ukazujú, že $v(H)$ závislosti rýchlych a pomalých DS sa stávajú menej strmými so zvyšujúcim sa aplikovaným ťahovým napätím (klesá pohyblivosť DS). Účinok ťahového napätia je však silnejší pre rýchle DS. $v(H)$ závislosti pre rýchle a pomalé DS sa k sebe približujú s rastúcim ťahovým napätím, až v určitom bode (aplikované napätie 283 MPa) sa dokonca prekrývajú a jednosmerný efekt už nie je pozorovaný. Toto je reverzibilný jav a pozoruje sa pri pomalých aj rýchlych DS.

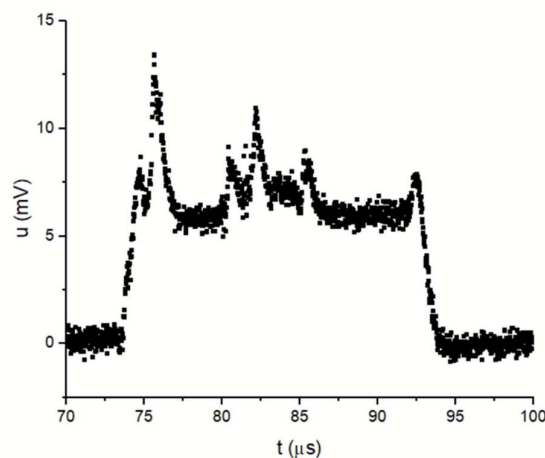


Obr. 13 Závislosti rýchlosti DS od axiálneho poľa získané vo vzorke pre rôzne hodnoty ťahového napätia, pre pomalú DS (otvorené symboly) a rýchlu DS (plné symboly).



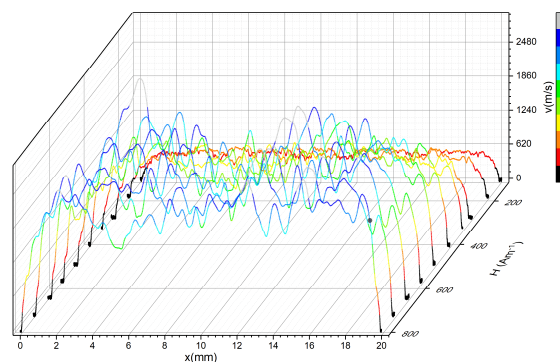
Obr. 14 Časová závislosť indukovaného signálu v dlhej snímačej cievke PuC2 meraná opakovane v rovnakom poli.

Keďže vyhodnotenie geometrie DS pomocou indukovaných signálov zo snímačej cievky je pomerne komplikované, zamerali sme sa na pôvod nových pozorovaných indukovaných vrcholov. Vynárajú sa otázky či tieto nové vrcholy na indukovaných signáloch majú fyzikálny význam a či sú reprodukovateľné alebo náhodné?



Obr. 15 Časová závislosť indukovaného signálu v dlhej cievke.

Záznamy signálov s rôznymi poľami (obr. 16) vyvolávajú dojem náhodnosti. Na druhej strane, ak sa tieto signály merajú opakovane v tom istom poli (obr. 14), výsledky naznačujú skôr systematický ako náhodný proces.

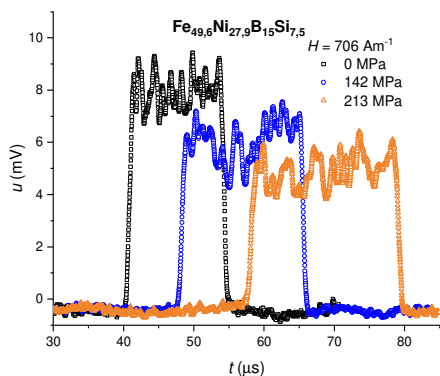


Obr. 16 Indukované signáli pri rôznych aplikovaných poliach.

Na zistenie, či sú tieto indukované vrcholy fyzikálne významné, boli vykonané merania aj na vzorke mikrodrôtu iného zloženia $Fe_{49,6}Ni_{27,9}B_{15}Si_{7,5}$ (obr. 17). Podobné vrcholy boli pozorované aj tu. Hoci ich amplitúdy neboli porovnateľné s priemernou hodnotou indukovaného napätia, ako to bolo možné pozorovať na vzorke bez obsahu niklu (obr. 14 a obr. 15), tak boli častejšie.

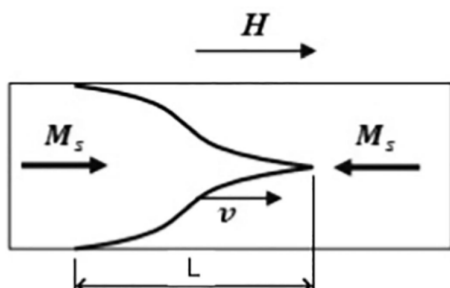
Pozorovaný pokles rýchlosti DS vplyvom axiálneho ťahového napätia (obr. 13) môže spôsobovať:

- a) skrátenie DS prostredníctvom zvýšenia energie DS,
- b) vratné zmeny v štruktúrach mikrodrôtov zodpovedných za jednosmerný efekt spôsobený aplikovaným ťahovým napätím na vzorke. Vyplýva to zo skutočnosti, že závislosti pre rýchlu a pomalú DS sa k sebe približujú so zvyšujúcim sa ťahovým napätím.



Obr. 17 Namerané signály indukované v snímačej cievke s aplikovaným ťahovým napätím vo vzorke $\text{Fe}_{49,6}\text{Ni}_{27,9}\text{B}_{15}\text{Si}_{7,5}$.

Dôležitou črtou javu (vysoké napät'ové špičky) vidného na obr. 15 je, že rýchlosť DS sa zvyšuje (v porovnaní s priemernou hodnotou), hoci by sa dal očakávať jej pokles od vplyvu lokálnych porúch a záchytných centier. Naše predchádzajúce výsledky [19], [20] tiež ukazujú, že DS v bistabilnom mikrodrôte nie je pevný objekt v tom zmysle, že by si zachovávala svoj tvar a rozmery bez ohľadu na rýchlosť (aplikované magnetické pole).



Obr. 18 Typ možného tvaru DS s valcovou symetriou [21].

Model kónického typu DS s vrcholom vpredu (obr. 18) kvalitatívne vysvetľuje skracovanie DS so zvyšujúcou sa rýchlosťou.

Uvažujme, že ťahové napätie aplikované na vzorku vytvárajú záchytné centrá na povrchu axiálne magnetizovaného jadra. Uvažujme ďalej, že v takýchto oblastiach pôsobí na DS brzdná sila. Táto lokálna sila spôsobí predĺženie DS. Tento proces nemusí nevyhnutne viesť k spomaleniu DS, pretože s predlžovaním DS sa zvyšuje jej pohyblivosť. V momente keď DS opustí záchytné centrum, DS je dlhá a jej rýchlosť sa začína zvyšovať a dosiahne maximum. Ďalej ako sa bude postupne skracovať rýchlosť DS sa vráti na pôvodnú hodnotu. Zaujímavý paradox vzniká, keď záchytné centrum môže spôsobiť lokálne zvýšenie rýchlosti DS. Takéto záchytné centrá môžu byť izolované (obr. 15) alebo rôzne prekryté (obr. 17). Práve prekrytie takýchto miest by mohlo spôsobiť zdanlivú náhodnosť pozorovaných vrcholov na indukovaných napät'ových signáloch s rôznymi magnetickými poľami.

Aj spomenutý experiment s rýchlou zmenou magnetického poľa naznačuje, že DS si nejaký čas zachováva svoje parametre (napr. rozmery). Ak sa napríklad pole zmení z nízkej hodnoty (dlhá DS, veľká pohyblivosť) na väčšiu hodnotu (krátka DS, malá pohyblivosť), takáto DS sa istú dobu (kým neprispôbi svojím rozmerom novým podmienkam) pohybuje vyššou rýchlosťou ako je „rovnovážna“ rýchlosť v danom poli teda pozorujeme „vrchol“ vo $v(x)$ závislosti. Vynára sa prirodzená otázka: môže byť mechanizmus tohto druhu zodpovedný za pozorované „vrcholy“ pozorované na meraných napät'ových signáloch v snímačích cievkach?

ZÁVER

Tento príspevok prezentuje experiment a spracovanie údajov, ktoré poskytuje informácie o zmenách rýchlosti DS šíriacej sa pozdĺž mikrodrôtu v oblasti rýchlej zmeny aplikovaného magnetického poľa. Merania sa uskutočňovali na mikrodrôte so silným jednosmerným efektom ako aj na vzorkách s obsahom Ni bez jednosmerového efektu. Bolo pozorované, že keď sa pohybuje DS cez oblasť rýchlej zmeny poľa je nútená zmeniť svoje parametre vo veľmi krátkom čase. Zodpovedajúca zmena rýchlosti DS je odlišná pre rýchle zvýšenie alebo zníženie hodnoty magnetického poľa. Experimentálne výsledky prezentované v tomto článku podporujú myšlienku, že DS v bistabilnom mikrodrôte nie je pevným objektom, ale mení svoje parametre (napr. rýchlosť DS) v závislosti od aplikovaného poľa. Tiež sú prezentované výsledky ktoré odhaľujú indukované signály merané na vzorkách vystavených vplyvu axiálneho ťahového napätia, ktoré nám zatiaľ neumožňujú vyhodnotiť zmeny v geometrii DS v dôsledku indukovaných napät'ových špičiek. Pokles rýchlosti DS bol pozorovaný vo vzorke s aplikovaným ťahovým napätím. Jednosmerný efekt nebol pozorovaný, keď aplikované napätie bolo dostatočne veľké. Ďalším zaujímavým pozorovaným javom bol výskyt lokálneho zvýšenia rýchlosti DS (významné vrcholy na $u(t)$, $v(x)$, závislosti).

POĎAKOVANIE

Tento výskum bol podporený Agentúrou na výskum a vývoj na základe zmluvy č. APVV-16-0079, grantom VEGA č. 1/0250/21 od Vedeckej grantovej agentúry Ministerstva školstva SR a grantom FEI v rámci zmluvy č. FEI-2023-92 udelená Technickou univerzitou v Košiciach.

LITERATÚRA

- [1] Herzer, G. Magnetic materials for electronic article surveillance. *J. Magn. Magn. Mater.* (2003), 254–255, 598–602.
- [2] Zhukov, A.; Ipatov, M.; Corte-Leon, P.; Blanco, J.M.; Zhukova, V. Advanced functional magnetic microwires for magnetic sensors suitable for biomedical applications. Ch. 18. In *Magnetic Materials and Technologies for Medical Applications*; Tishin, A.M., Ed.; Elsevier: Amsterdam, The Netherlands, (2022).
- [3] Fiorillo, F.; Bertotti, G.; Appino, C.; Pasquale, M. Soft Magnetic Materials. In *Wiley Encyclopedia of Electrical and Electronics Engineering*; Webster, J., Ed.; John Wiley & Sons, Inc.: Torino, Italy, (1999); p. 42.
- [4] Lenz, J.; Edelstein, A.S. Magnetic sensors and their applications. *IEEE Sens. J.* (2006), 6, 631–649.
- [5] Morón, C.; Cabrera, C.; Morón, A.; García, A.; González, M. Magnetic Sensors Based on Amorphous Ferromagnetic Materials: A Review. *Sensors* (2015), 15, 28340–28366.

- [6] Ripka, P.; Vertesy, G. Sensors based on soft magnetic materials Panel discussion. *J. Magn. Magn. Mater.* (2000), 215, 795–799.
- [7] Taylor, G.F. A Method of Drawing Metallic Filaments and a Discussion of Their Properties and Uses. *Phys. Rev.* 1924, 23, 655.
- [8] Ulitovsky, A.V.; Maianski, I.M. Avramenco A I 1960 Method of Continuous Casting of Glass Coated Microwire. USSR Patent 128427, 15 May 1960.
- [9] Zhukov, A.; Corte-Leon, P.; Gonzalez-Legarreta, L.; Ipatov, M.; Blanco, J.M.; Gonzalez, A.; Zhukova, V. Advanced Functional Magnetic Microwires for Technological Applications. *J. Phys. D Appl. Phys.* 2022, 55, 253003.
- [10] Jurc, R.; Frolova, L.; Kozejova, D.; Fecova, L.; Hennel, M.; Galdun, L.; Richter, K.; Gamcova, J.; Ibarra, P.; Hudak, R.; et al. Sensoric Application of Glass-Coated Magnetic Microwires, 2nd ed.; Elsevier Ltd.: Amsterdam, The Netherlands, 2020; ISBN 9780081028322.
- [11] Zhukova, V.; Corte-Leon, P.; Blanco, J.M.; Ipatov, M.; Gonzalez-Legarreta, L.; Gonzalez, A.; Zhukov, A. Development of Magnetically Soft Amorphous Microwires for Technological Applications. *Chemosensors* 2022, 10, 26.
- [12] Baranov, S.A.; Larin, V.S.; Torcunov, A.V. Technology, Preparation and Properties of the Cast Glass-Coated Magnetic Microwires. *Crystals* 2017, 7, 136.
- [13] Mohri, K.; Uchiyama, T.; Panina, L.V.; Yamamoto, M.; Bushida, K. Recent Advances of Amorphous Wire CMOS IC MagnetImpedance Sensors: Innovative High-Performance Micromagnetic Sensor Chip. *J. Sens.* 2015, (2015), 718069.
- [14] Zhukov, A.; Cobeño, A.F.; Gonzalez, J.; Blanco, J.M.; Aragonese, P.; Dominguez, L. Magnetoelastic sensor of level of the liquid based on magnetoelastic properties of Co-rich microwires. *Sens. Actuat. A Phys.* (2000), 81, 129–133.
- [15] R. Varga, Magnetization processes in glass-coated microwires with positive magnetostriction, *Acta Phys. Slovaca* 62(5) (2012) 411-518.
- [16] Zhukov A., Ipatov M., Zhukova V.: Processing magnetic microwires for magnetic bistability and magnetoimpedance. *Magnetic Nano- and Microwires, Design, Synthesis, Properties and Applications*, In *Electronic and Optical Materials*, 2015. pp 225-274
- [17] Hudak, J., et al. Improved Sixtus–Tonks method for sensing the domain wall propagation direction. *Sensors and Actuators A: Physical*, 2009, 156.2: 292-295.
- [18] J. Onufer, J. Ziman, M. Kládiová, “Unidirectional effect in domain wall propagation observed in bistable glass-coated microwire”, *J. Magn. Magn. Mater.* 396 (2015) 313–317.
- [19] J. Horniaková, J. Onufer, J. Ziman, P. Duranka, S. Samuhel, Changes in geometry of propagating domain wall in magnetic glass-coated bistable microwire, *Journal of Magnetism and Magnetic Materials*. 529 (2021), 1-5. ISSN 0304-8853, <https://doi.org/10.1016/j.jmmm.2021.167846>
- [20] J. Onufer, J. Ziman, P. Duranka, S. Samuhel, J. Horniaková, M. Kládiová, Dynamics of single domain wall propagating in bistable microwire in rapidly-changing magnetic field (2022), In: *IEEE Transactions on Magnetics: a Publication of the IEEE Magnetics Society*. - New York (USA): IEEE Magnetics Society Roč. 58, č. 11 (2022), s. 1-6. - ISSN 0018-9464
- [21] J. Onufer, J. Ziman, P. Duranka, M. Kládiová, The influence of annealing on domain wall propagation in bistable amorphous microwire with unidirectional effect, *Physica B: Condensed matter* 540 (2018) 58-64.

Metódy a algoritmy akcelerácie/paralelizácie výpočtov fyzikálnych modelov

Michal Solanik

Katedra počítačov a informatiky
Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
Košice, Slovensko
michal.solanik@tuke.sk

Ján Genči

Katedra počítačov a informatiky
Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
Košice, Slovensko
jan.genci@tuke.sk

Pavol Bobík

Ústav experimentálnej fyziky
Slovenská akadémia vied
Košice, Slovensko
bobik@saske.sk

Abstract—The development and verification of mathematical models in physics is a long time-consuming process. In this thesis, we focused on the process of development of mathematical models to propose and design methods and means, which can be used to accelerate selected parts of the process. We designed three concepts, Geliosphere - software utilizing a parallel approach on CPU and GPU level to simulate cosmic rays modulation in the heliosphere, Autosphere - software using C/C++-like domain-specific language to automate process parallelization, and optimization, and distributed system encapsulating both concepts in a distributed environment. The results of this thesis are verified concepts and published implementation of Geliosphere. Verification of key concepts of Autosphere, and distributed system proved, that the proposed design is usable, and can reduce the required time for the development of the models. On top of verification, we implemented a prototype of Autosphere, which can be further used, developed and improved.

Abstract—Vývoj a verifikácia matematických modelov v prostredí fyziky je časovo náročný proces. V tejto práci sme sa zamerali na proces vývoj matematických modelov, aby sme navrhli a načrtli metódy a prostriedky, ktoré môžu byť použité na akceleráciu vybraných častí procesu vývoja matematických modelov. Navrhli sme tri koncepty, Geliosphere - softvér využívajúci paralelný prístup na úrovni CPU a GPU na simulácie modulácie kozmického žiarenia v heliosfére, Autosphere - softvér využívajúci doméno-špecifický jazyk podobný C/C++ na automatizáciu procesu paralelizácie a optimalizácie distribuovaného systému, ktorý zapuzdruje tieto dva koncepty v distribuovanom prostredí. Výsledkom tejto práce sú verifikované koncepty a zverejnená implementácia softvéru Geliosphere. Verifikácia kľúčových konceptov softvéru Autosphere a distribuovaného systému potvrdila, že navrhovaný dizajn je použiteľný a môže redukovať potrebný čas pre vývoj matematických modelov. Navyše pri verifikácii sme implementovali funkčný prototyp softvéru Autosphere, ktorý môže byť ďalej používaný, vyvíjaný a vylepšovaný.

I. ÚVOD

Matematické modely nám umožňujú opísať rôzne javy v reálnom svete. Oblasti, ktoré sú bežným ľuďom najznámejšie sú finančníctvo, alebo napríklad aj meteorológia. Napríklad, GFS model [1] je model s nízkym rozlíšením, ktorého predikcia môže byť nepresná v horskom prostredí. Na druhú stranu, ICON model [2] môže poskytnúť dostatočnú presnosť v určitých regiónoch, ako napríklad v Európe.

Vo všeobecnosti, fyzika je jedna z oblastí, spolu s ekonómiou, biológiou a inžinierstvom, ktorá využíva vo

veľkom matematické modelovanie. Príkladom matematického modelovania s vysokým dopytom po výpočtovom výkone je fyzika kozmického žiarenia.

Kozmické žiarenie [3] tvoria vysoko energetické častice s pôvodom v Slnku (solárne energetické častice), alebo s pôvodom mimo Slnčnej sústavy (galaktické kozmické žiarenie). Širia sa naprieč oblasťou s prevahou magnetického poľa Slnka, ktorá sa nazýva heliosféra. Interagujú s nepravidlosťami tohto poľa a strácajú energiu, čím sa mení ich energetické spektrum. Tento proces sa nazýva modulácia kozmického žiarenia v heliosfére. Parker [4] sformuloval transportnú rovnicu, ktorá je základnou rovnicou pre moduláciu kozmického žiarenia v heliosfére. Kozmické žiarenia ovplyvňujú rôzne javy. Sú nevyhnutné na tvorbu C^{14} v atmosfére, čím umožňuje rádiokarbónové datovanie. Produkcia sekundárneho kozmického žiarenia umožňuje miónovú tomografiu. Kozmické žiarenie je dôležité v prípade takzvanej predpovede kozmického počasia. Simulácie modulácie kozmického žiarenia môžu odhadnúť, akú veľkú dávku radiácie môže kozmická loď alebo ľudská posádka dostať počas cestovania v heliosfére.

Paralelný prístup môže teoreticky výrazne skrátiť čas vykonávania na viacjadrových procesoroch. Na druhej strane prináša zvýšené náklady na komunikáciu medzi jednotlivými jadrami. Zrnitosť v paralelnom prístupe je jednou z najdôležitejších vlastností. Má významný vplyv na vykonávací čas. Použitie jemnozrného alebo hrubozrného paralelizmu môže viesť k predĺženiu času vykonávania v dôsledku vyšších nákladov na komunikáciu a synchronizáciu v neoptimalizovaných implementáciách. Komunikácia je previazaná s dekompozíciou problému. Efektívne paralelná implementácia vyžaduje správnu dekompozíciu problému.

Paralelizmus vieme dosiahnuť na jednom počítači, na štandardnom viacjadrovom procesore. Výpočtová sila jediného viacjadrového procesora nemusí byť dostatočná pre isté úlohy, aby sme boli schopní udržať vykonávací čas na úrovni minút, alebo hodín. Pre tieto typy úloh, paralelizmus môžeme aplikovať naprieč viacerými výpočtovými jednotkami alebo viacerými zariadeniami. V posledných rokoch získali na popularite grafické procesorové jednotky pre všeobecné použitie. Z hľadiska Flynnovej taxonómie GPU [5] predstavuje SIMD,

alebo viac špecificky SIMT model, ktorý je podmodelom SIMD prezentovaným v [6].

Modely modulácie kozmického žiarenia v heliosfére [7] sú zvyčajne založené na stochastickom prístupe. Väčšina autorov [8] [9] [10] využíva metódu Monte Carlo pre simulovanie modulácie kozmického žiarenia v heliosfére. Simulácia prebieha na úrovni nezávislých častíc. Nezávislosť častíc je veľmi dôležitá vlastnosť, čo umožňuje aby boli implementácie modelov efektívne na architektúrach SIMT, akou je napríklad GPU.

Vývoj modelov v oblasti fyziky je časovo náročný. Paralelizácia a automatizácia procesu paralelizácie môže byť veľmi užitočná pre výskumníkov, aby vyvíjali, verifikovali a ďalej vylepšovali svoje modely.

II. EXISTUJÚCE SYSTÉMY

Naším primárnym cieľom je akcelerácia simulácii modulácii kozmického žiarenia v heliosfére. V rámci tejto oblasti existuje niekoľko systémov, ktoré môžeme analyzovať:

- SOLARPROP - [9] otvorený softvér, ktorý obsahuje implementácie viacerých modelov modulácie kozmického žiarenia v heliosfére. Kappl vo svojom článku [9] taktiež publikoval svoj 2D model, ktorý je v softvéri SOLARPROP označený ako standard 2D. Softvér umožňuje vstup pomocou parametrov z príkazového riadku a pomocou vstupných súborov. Vstupné súbory umožňujú definovať parametre pre simulácie, vstupné spektrum a uhol sklonu.
- Program na simulovanie modulácii kozmického žiarenia v heliosfére - [10] je softvér, ktorý obsahuje implementácie dvoch 1D modelov, 1D backward-in-time (neskôr referovaného ako 1D B-p) a 1D forward-in-time (neskôr referovaného ako 1D F-p) modelu. Narozdiel od softvéru SOLARPROP poskytuje GPU implementáciu modelov modulácie kozmického žiarenia v heliosfére. Jedným z ďalších rozdielov je to, že výsledky jednotlivých simulovaných častíc sú zapisované na disk a následne vyhodnocované. To umožňuje používateľom podrobnejšiu analýzu výsledkov simulácii jednotlivých častíc. V porovnaní s pôvodnými implementáciami bolo dosiahnuté zrýchlenie na úrovni 8 v prípade 1D F-p modelu a minimálne 71 v prípade 1D B-p modelu.
- Distribuovaný systém - [19] softvér, ktorý sa zameriava na distribúciu simulácii modelov modulácie kozmického žiarenia v heliosfére naprieč viacerými počítačovými systémami. Architektúra v podobe jednotlivých kontajnerov je zobrazená na obrázku 1. Distribuovaný systém sa skladá z dvoch menších softvérových systémov - Hlavného uzla distribuovaného systému a uzla distribuovaného systému. Úloha hlavného uzla je rozdeľovať simulácia na menšie časti simulácie, ktoré sú následne na základe dostupnosti poslané dostupným uzlom distribuovaného systému. Uzol distribuovaného systému využíva Program na simulovanie modulácii kozmického žiarenia v heliosfére pre vykonávanie jednotlivých výpočtov. Komunikácia prebieha pomocou sprostredkovateľa správ (angl. message broker) pomocou protokolu AMPQ.

Paralelizácia na úrovni distribuovaného systému bola hrubozrná. Delenie simulácii sa uskutočňovalo na základe výkonnosti GPU v dostupných uzloch distribuovaného systému. Tabuľka I zobrazuje vykonávacie časy pre simulácie 1D F-p modelu bez použitia priepustnosti vo výpočte delenia simulácii. Rozdiel medzi použitými GPU je dost výrazný, v rozmedzí 90 minút. Preto sme sa rozhodli započítať priepustnosť pamäte do výpočtu, ktorá by mohla byť pôvodcom týchto rozdielov. Výsledky pre vykonávacie časy pre simulácie 1D F-p modelu s použitím priepustnosti sa nachádzajú v tabuľke II. V tomto prípade je rozdiel len 6 minút, čo potvrdzuje to, že táto korekcia prispela k zlepšeniu výkonu distribuovaného systému.

- HelMod - [21] 2D model pre solárnu moduláciu galaktického kozmického žiarenia. Podobne, ako ostatné modely v tejto oblasti je založený na Parkerovej transportnej rovnici. Je dostupný iba vo webovej aplikácii a je možné púšťať simulácie solárnej modulácie galaktického kozmického žiarenia v heliosfére.

TABLE I
VYKONÁVACIE ČASY PRE 1D F-P MODEL BEZ POUŽITIA PRIEPUSTNOSTI GPU PAMÄTI VO VÝPOČTE DISTRIBÚCIE

GPU	Počet simulácii v miliardách	Vykonávací čas [min]
GTX1080TI	57	238
RTX2060	43	328

TABLE II
VYKONÁVACIE ČASY PRE 1D F-P MODEL S POUŽITÍM PRIEPUSTNOSTI GPU PAMÄTI VO VÝPOČTE DISTRIBÚCIE

GPU	Počet simulácii v miliardách	Vykonávací čas [min]
GTX1080TI	64	251
RTX2060	36	257

Okrem systémov, ktoré sa priamo týkajú heliosféry existujú ešte systémy, ktoré sa zaoberajú veľmi príbuznými oblasťami ako je napríklad magnetosfére alebo vesmírne spršky:

- COR - [20] systém, ktorý umožňuje simulovať trajektórie kozmického žiarenia v magnetosfére. Celý systém tvorí webová aplikácia, systém na pridelovanie úloh a jednotlivé moduly na vizualizáciu, generovanie výstupu a podobne.
- CORSIKA - [22] systém, ktorý umožňuje simulovať vesmírne spršky. Je dostupný vo forme aplikácie s interakciou v príkazovom riadku. Jednou z jeho výhod je veľmi ľahká rozšíriteľnosť existujúcich modelov o nové modely.
- GALPROP - [23] systém používaný na simuláciu šírenia relativisticky nabitých častíc a difúzných emisií produkovaných počas ich šírenia v našej galaxii. Dostupný v dvoch formách, vo forme kódu a webovej aplikácie.

III. ŠPECIFIKÁCIA SYSTÉMOV

Nami vytvorené systémy, Program na simulovanie modulácii kozmického žiarenia v heliosfére a distribuovaný systém preukázali svoje limity, najmä čo sa týka flexibility

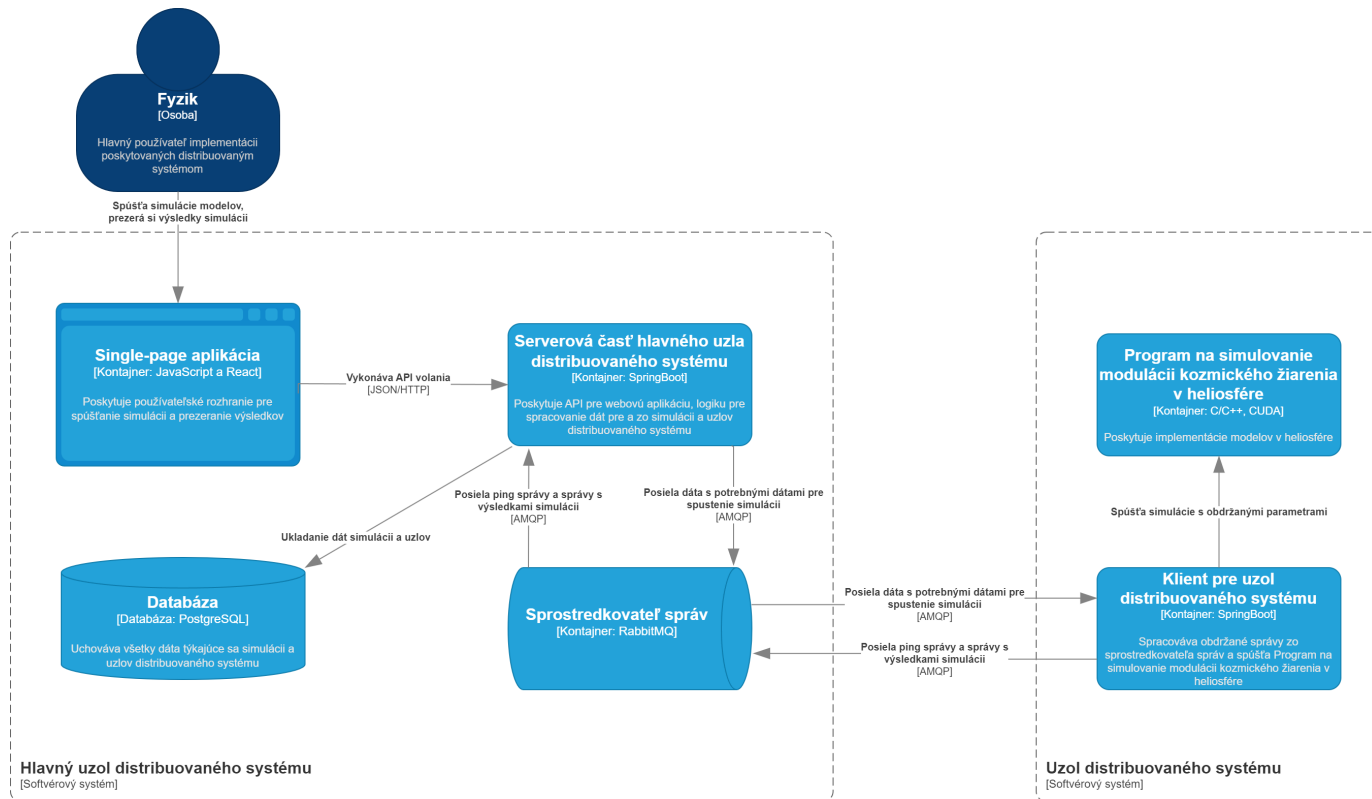


Fig. 1. Diagram kontajnerov starého distribuovaného systému.

a podpory vývoja nových modelov modulácie kozmického žiarenia v heliosfére.

Obrázok 2 zobrazuje proces vývoja modelu modulácie kozmického žiarenia v heliosfére. Fyzik je zodpovedný za odvodenie modelu, matematiky pre model, jej verifikáciu a základnú implementáciu modelu a jeho verifikáciu. Následne vývojár pripraví GPU verziu, ktorú verifikuje a integruje do Programu na simulovanie modulácii kozmického žiarenia v heliosfére. Z toho vyplýva, že by sme sa mali zamerať okrem samotnej implementácie aj na to, ako rýchlejšie testovať jednotlivé implementácie modelov. Jeden z najjednoduchších prístupov v tomto prípade by bolo mať pripravenú už paralelnú verziu bez zásahu vývojára. To môžeme dosiahnuť pomocou použitia rôznych knizníc pre paralelizáciu, akou je napríklad OpenMP opísaný v [11] alebo OpenACC opísaný v [12].

Ďalšou z možností je návrh systému, ktorý by dokázal automatizovať optimalizáciu a paralelizáciu modelu. Z hľadiska prístupov tieto nástroje môžeme deliť na doménovo-špecifické a nástroje všeobecného použitia. Ako jedným z príkladov doménovo-špecifických nástrojov je GeNN opísaný v [13]. Využíva preddefinované rozhrania na to, aby zabezpečili požadovanú formu, z ktorej generujú GPU implementácie modelu pre simulovanie činnosti neurónov v mozgu. Lindsey a spol. v [14] použili doménovo-špecifický jazyk založený na LaTeX rovniciach pre generáciu paralelného výstupného riešenia. Z kategórie nástrojov všeobecného použitia môžeme menovať napríklad nástroj OpenTuner opísaný v [15].

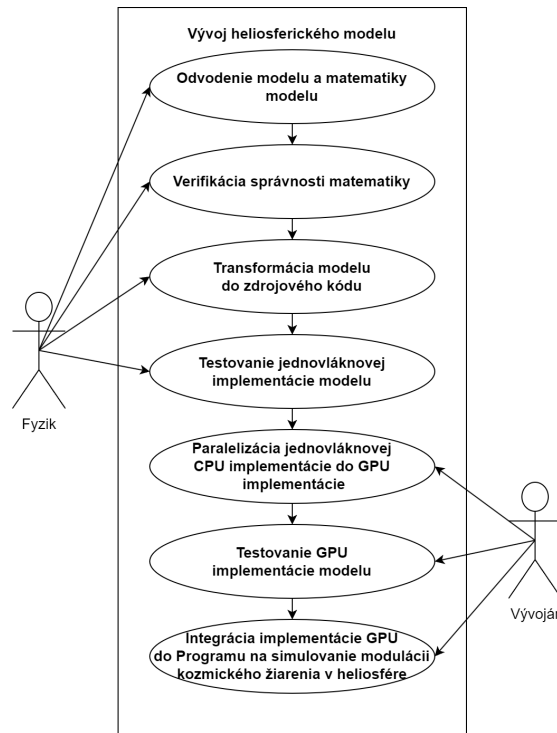


Fig. 2. Proces vývoja modelu modulácie kozmického žiarenia v heliosfére.

V konečnom dôsledku, môžeme navrhované systémy rozdeliť do troch kategórií:

- systém umožňujúci púšťať paralelne implementácie modelov pre simulovanie modulácii kozmického žiarenia v heliosfére
- systém umožňujúci na základe doménovo-špecifického jazyka definovať modelov a vygenerovať optimalizovanú paralelnú implementáciu podľa platformy, na ktorej sa nachádza
- systém ktorý zapuzdruje predchádzajúce dva do distribuovaného prostredia.

Tieto systémy by mali podporovať každý krok procesu vývoja modelu modulácie kozmického žiarenia v heliosfére okrem definície modelu a matematiky modelu. Z hľadiska ďalších špecifikácií, by mal systém umožňujúci púšťať paralelné implementácie modelov mať nasledujúce vlastnosti: možnosť interakcie cez príkazový riadok, podpora vstupu cez rok a mesiac a mať k dispozícii optimalizované paralelné implementácie.

Z hľadiska systému, ktorý umožňuje na základe doménovo-špecifického jazyka definovať modelov a vygenerovať optimalizovanú paralelnú implementáciu by sme mali zohľadniť nasledujúce vlastnosti: doménovo-špecifický jazyk by mal z hľadiska gramatiky pripomínať C/C++, vygenerované paralelné verzie by mali využívať GPU alebo CPU ako primárnu výpočtovú jednotku, životný cyklus by mal zabezpečiť optimalizáciu a paralelizáciu vstupného doménovo-špecifického kódu.

Nový distribuovaný systém by mal obsahovať nasledujúce vlastnosti: bežný používateľ by mal vedieť interagovať so systémom pomocou webového rozhrania, používatelia s administrátorskými právami by mali vedieť pridávať nové modely do systému, spravovať simulácie a uzly distribuovaného systému. Systém by mal využívať jemnozrný paralelizmus na distribúciu výpočtov, ktorý by zároveň mal poskytovať dobrú úroveň odolnosti voči zlyhaniam.

Pre lepšie rozoznanie medzi jednotlivými systémami budeme referovať na systém umožňujúci púšťať paralelne implementácie modelov pre simulovanie modulácii kozmického žiarenia v heliosfére ako na softvér Geliosphere. Na systém umožňujúci na základe doménovo-špecifického jazyka definovať modelov a vygenerovať optimalizovanú paralelnú implementáciu podľa platformy, na ktorej sa nachádza budeme referovať na softvér Autosphere. A na systém ktorý zapuzdruje predchádzajúce dva do distribuovaného prostredia budeme referovať ako nový distribuovaný systém.

IV. NÁVRH SYSTÉMOV PRE AKCELERÁCIU SIMULÁCII MODELOV MODULÁCIE KOZMICKÉHO ŽIARENIA V HELIOSFÉRE

V tejto sekcii sa budeme zaoberať návrhom návrhom jednotlivých systémom, ktoré sme navrhli v predchádzajúcej sekcii. Kladieme dôraz najmä na navrhnutie základnej architektúry a životného cyklu.

A. Geliosphere

Z hľadiska dekompozície môžeme implementáciu simulácie rozdeliť na tri časti: pred-simulačnú, simulačnú a post-simulačnú. Tieto tri časti nám uľahčia efektívnejšie využívanie pamäte, najmä kvôli GPU implementáciám, ktoré sú z hľadiska rýchlosti veľmi závislé na rýchlych pamätiach.

Životný cyklus softvéru Geliosphere by mal minimálne obsahovať nasledujúce kroky:

- spracovanie simulácii,
- spracovanie výstupných súborov zo simulácii,
- príprava výstupných súborov vo forme energetických spektier.

Na obrázku 3 je zobrazený životný cyklus softvéru Geliosphere. Vo všeobecnosti, môžeme ho rozdeliť na nasledujúce časti:

- spracovanie vstupu - vstupe je spracovaný z príkazového riadku a vstupných súborov. Jedným z typov vstupných súborov je súbor, ktorý obsahuje hodnoty difúzneho koeficientu, sklon uhla pre každý dostupný mesiac a rok založený na [16],
- spracovanie simulácii na CPU - pre vykonanie simulácii na CPU je potrebné inicializovať vlákna, ktorých sa budú simulácie vykonávať,
- alokácie pamäte pre GPU - GPU implementácie okrem štandardnej alokácie pamäte potrebujú inicializovať štruktúry pre generovanie pseudonáhodných čísel,
- Spracovanie simulácii na GPU - GPU implementácie okrem vykonávania jednotlivých funkcií zabezpečujú synchronizáciu vlákien a kopírovanie dát medzi GPU a pamäťou počítača, na ktorom sa simulovanie odohráva,
- Analýza výstupných dát - po dokončení simulácii sú výstupné súbory analyzované a na základe výstupných dát pre jednotlivé častice sú vypočítané energetické spektrá.

V rámci životného cyklus sme odseparovali CPU a GPU implementácie. To nám umožňuje neskôr definovať separátne moduly pre oba typy implementácie. Nutnosť kompilovať riešenie aj s GPU implementáciami by značne obmedzila použiteľnosť riešenia pre používateľov. Prezentované riešenie umožňuje kompiláciu aj bez GPU implementácii.

B. Autosphere

Proces optimalizácie a paralelizácie môžeme charakterizovať ako opakovaný proces, v ktorom porovnáваме medzi sebou viacero verzií na základe meraných charakteristík. Podobný prístup je charakteristikou genetických algoritmov, popísaných autorom Koza a spol. v [17]. Genetické algoritmy sú založené na princípe vygenerovania generácie programov a ich vyhodnotenia. Podobný prístup sme sa rozhodli aplikovať pri návrhu softvéru Autosphere. Životný cyklus softvéru Autosphere je zobrazený na obrázku 4.

Životný cyklus softvéru Autosphere pozostáva z nasledujúcich krokov:

- Spracovanie doménovo-špecifického jazyka do doménového modelu - vstupom pre softvér Autosphere bude doménovo-špecifický jazyk veľmi podobný C/C++ z

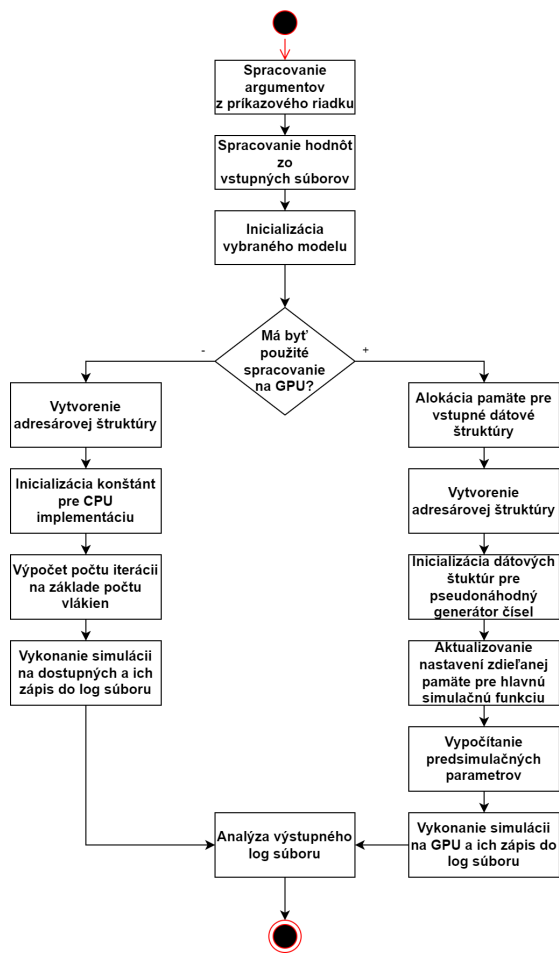


Fig. 3. Životný cyklus softvéru Geliosphere.

hľadiska gramatiky, reflektujúci dekompozíciu simulácii na pred-simulačnú časť, simulačnú časť a post-simulačnú časť.

- Analýza dostupných nástrojov a platforiem - pre ďalšie spracovávanie vstupného doménovo-špecifického kódu je potrebné zistiť, s akými nástrojmi a platformami môžeme pracovať počas automatizovanej paralelizácie a optimalizácie.
- Statická analýza - je to analýza, ktorá zodpovedá za generovanie odporúčaní na základe doménovo-špecifického modelu. Odporúčanie v tomto kontexte je doménový objekt, ktorý obsahuje postup, ako sa dopracovať k želanéj zmene, ktorá má za účel optimalizovať, alebo paralelizovať výstupný kód.
- Generovanie generácie programov - krok, ktorý slúži na vygenerovanie paralelných implementácií na základe dostupných platforiem a nástrojov.
- Dynamická analýza - je to analýza, ktorá generuje odporúčania na základe výstupu z iných programov, alebo meraní. Odporúčania v tomto prípade slúžia najmä ako pomôcka pre filtrovanie a vyhodnotenie, ale v prípade novej generácie sú aplikované na prvého potomka.

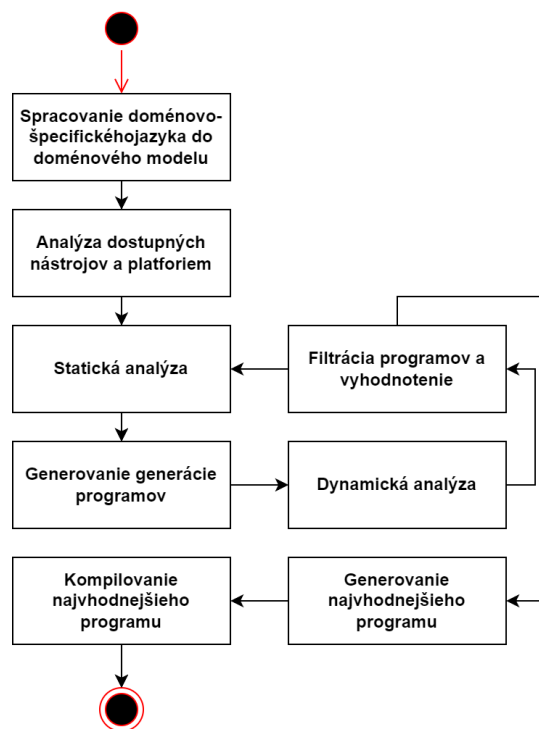


Fig. 4. Životný cyklus softvéru Autosphere.

- Filtrácia programov a vyhodnotenie - filtrácia programov slúži na odfiltrovanie riešení, ktoré sú pre nás nevyhovujúce. Napríklad v prípade GPU implementácií môžeme odfiltrovať kandidátske riešenia, ktoré obsahujú alokovanú lokálnu pamäť¹, ak sa v aktuálnej generácii nachádza aj riešenie, ktoré ju nepoužíva. Vyhodnotenie zodpovedá za to, že na základe dynamickej analýzy určí, či je potrebné pokračovať ďalej v optimalizácii.

Prezentované kroky zaisťujú vygenerovanie čo najoptimálnejšej paralelnej implementácie pre danú platformu.

C. Nový distribuovaný systém

Najväčšia zmena, ktorú realizujeme v novom distribuovanom systéme je zmena typu paralelizmu. Ako sme diskutovali v sekcii II, distribuovaný systém využívajúci hrubozrný paralelizmus bol problematický najmä z hľadiska efektívneho využívania výpočtových zdrojov. Pôvodné riešenie bolo problematické najmä z hľadiska využívania sprostredkovateľa správ, ktorého bezpečnostná stránka nebola celkom pod našou kontrolou. Z tohto hľadisa sme sa rozhodli zmeniť architektúru, aby sme ju prispôsobili novým podmienkam.

Na obrázku 5 sa nachádza diagram kontajnerov nového distribuovaného systému. Definovali sme niekoľko nových komponentov v systéme, ktoré môžeme zhrnúť ako:

- Služby pre hlavné doménové entity - pôvodnú aplikáciu sme rozdelili do mikroslužieb podľa doménového modelu. Simulačná služba zodpovedá za delenie simulácii na

¹Pamäť je alokovaná na globálnej pamäti, ktorá má pomalší čas prístupu oproti registrom, alebo zdieľanej pamäti.

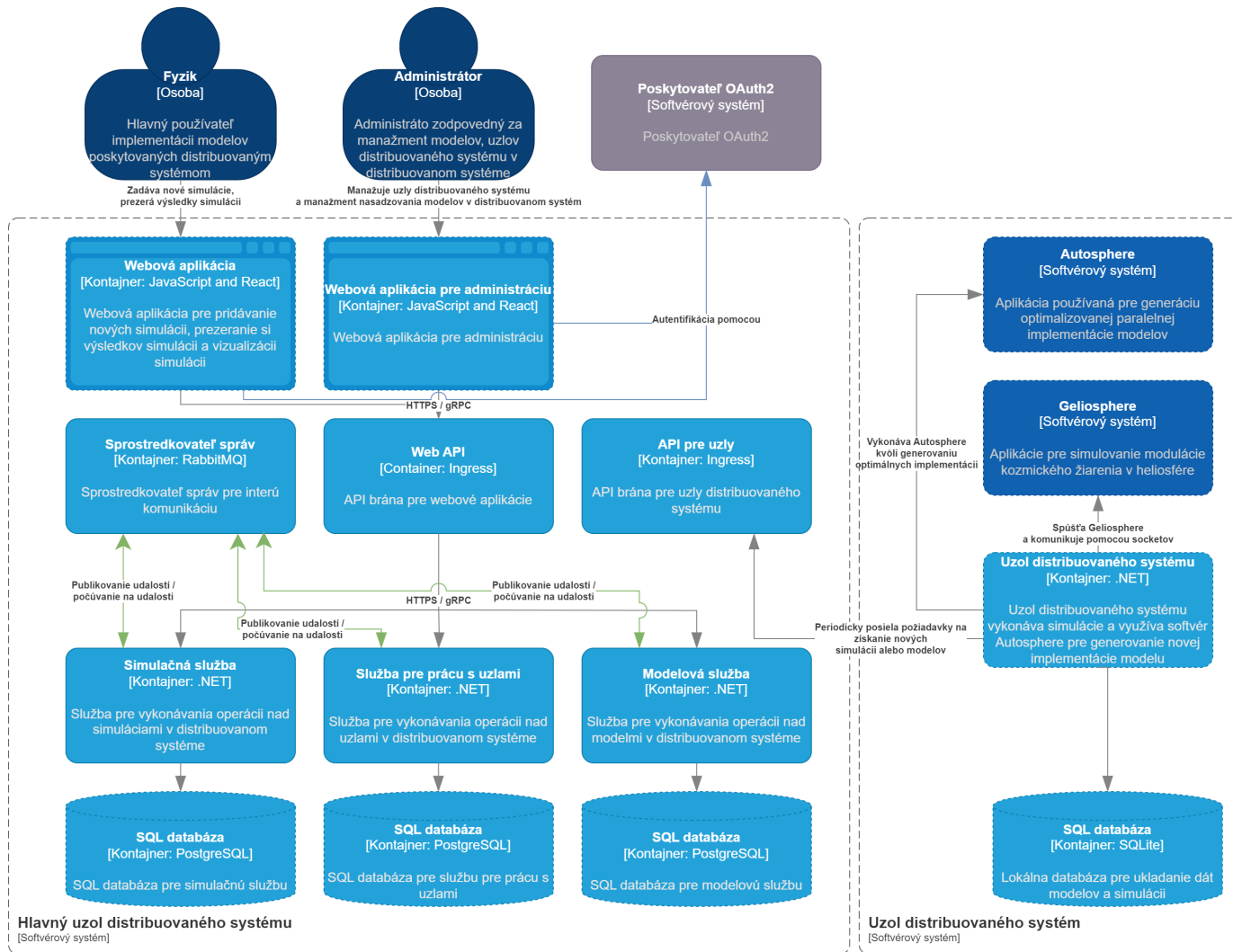


Fig. 5. Diagram kontajnerov nového distribuovaného systému.

menšie celky, ktoré sú neskôr spracované na uzloch distribuovaného systému. Služba pre prácu s uzlami registruje uzly a má na starosti to, ktoré uzly distribuovaného systému môžu obdržať ďalšie časti simulácie. Modelová služba zodpovedá za ukladanie modelov a ich jednotlivých verzií.

- Webové aplikácie - webové aplikácie slúžia na interakciu s bežnými používateľmi pre spúšťanie simulácií, prezeranie ich výsledkov alebo vizualizácií. Administrácia aplikácie je realizovaná pomocou samostatnej webovej aplikácie, ktorá umožňuje pridávať jednotlivé modely.
- Sprostredkovateľ správ - slúži internú komunikáciu medzi jednotlivými vnútornými službami. Pomocou počúvania na jednotlivé udalosti sme schopní v jednotlivých službách správne reagovať na zmenu.
- API brány - slúžia na komunikáciu s webovými aplikáciami a uzlami distribuovaného systému.
- Uzol distribuovaného systému - slúži na manažment výpočtov a model v rámci jedného počítača. Lokálna

databáza slúži na ukladanie základných dát o dostupných modeloch a simulácií v rámci uzla distribuovaného systému. So softvérom Geliosphere interaguje pomocou socketov a signálov. Sockety slúžia na textovú komunikáciu, kým signály indikujú pripravenosť softvéru Geliosphere a ukončenie prebiehajúcej simulácie. Uzol distribuovaného systému využíva softvér Autosphere na generovanie nových paralelných implementácií modelov.

V. VERIFIKÁCIA SYSTÉMOV

V rámci verifikácie systémov sme sa rozhodli pre verifikáciu najmä z hľadiska akcelerácie. V nasledujúcich podsekcích sa zameriavame na vyhodnotenie akcelerácia z hľadiska kľúčových konceptov jednotlivých nástrojov.

A. Geliosphere

GPU implementácie poskytujú značnú rýchlostnú výhodu oproti paralelných CPU implementáciám, ak sú správne optimalizované. V rámci jednotlivých implementácií sa pohybujeme na približne 50-100% využiteľnosti GPU.

Tabuľka III zobrazuje namerané časy pre SOLARPROPLIKE 2D model pre júl 1997. Prezentované zrýchlenie sa pohybuje na úrovni 3 až 34.

Tabuľka IV zobrazuje namerané časy pre Geliosphere 2D model pre júl 1997. Prezentované zrýchlenie sa pohybuje na úrovni 24 až 85.

Dosiahnuté zrýchlenie závisí na vstupných parametroch, ale zväčšuje sa s dlhšími výpočtami. V oboch prezentovaných prípadoch je ale výsledná GPU implementácia výrazne rýchlejšia s porovnateľnými výsledkami v presnosti.

TABLE III

NAMERANÉ VYKONÁVACIE ČASY PRE SIMULÁCIE 2D SOLARPROPLIKE MODELU PRE JÚL 1997.

Zariadenie	Dĺžka kroku [s]	Nameraný čas [s]
GTX1080TI	1000.0	281
Intel i5-11400F	1000.0	793
GTX1080TI	50.0	344
Intel i5-11400F	50.0	11 820

TABLE IV

NAMERANÉ VYKONÁVACIE ČASY PRE SIMULÁCIE GELIOSPHERE 2D MODELU PRE JÚL 1997.

Zariadenie	Dĺžka kroku [s]	Nameraný čas [s]
GTX1080TI	1000.0	376
Intel i5-11400F	1000.0	8 677
GTX1080TI	50.0	2 230
Intel i5-11400F	50.0	188 500

B. Atmosphere

Softvér Atmosphere sme sa rozhodli otestovať pomocou implementácie prvého prototypu softvéra. Na základe prezentovaného životného cyklu sme implementovali nasledujúce odporúčania generovaných zo statickej a dynamickej analýzy pre testovanie GPU implementácie:

- odporúčanie statickej analýzy pre aktualizáciu metadát vo vstupnom modeli
- odporúčanie statickej analýzy pre zníženie presnosti pomocou príznakov pri kompilácii,
- odporúčanie dynamickej analýzy pre určenie veľkosti vlákien a blokov pre GPU implementácie,
- odporúčanie dynamickej analýzy na základe množstva a typov používanej pamäti.

V prípade CPU implementácie sme navyše implementovali nasledujúce typy odporúčaní statickej a dynamickej analýzy:

- odporúčanie statickej analýzy pre aktualizáciu presnosti premenných,
- odporúčanie dynamickej analýzy založenej na vykonávacích časoch.

V prípade GPU implementácie sme boli schopní vygenerovať implementáciu, ktorej vykonávací čas bol 3.5 minúty. Vstupné parametre do simulácie bola rýchlosť slnečného vetra $V = 400 km/s$, veľkosť difúzneho koeficienta $K_0 = 5 \times 10^{23} cm^2/s$, veľkosť časového kroku $dt = 50.0s$ pre

TABLE V

VÝSLEDKY TESTOVANIA GRANULARITY V DISTRIBUOVANOM SYSTÉME [18]

Paralelizmus	Vykonávacie časy [hod.]
Hrubozrný paralelizmus	1.42
Hrubozrný paralelizmus po optimalizácii	1.17
Jemnozrný paralelizmus	0.942

117 miliónov simulačných trajektórií častíc. V porovnaní s vykonávacím časom zo softvéru Geliosphere bol vygenerovaný program rýchlejší, ale nezodpovedal za vygenerovanie energetického spektra.

V prípade CPU implementácie bol 58 minút pre vygenerovaný program a 161 minút pre softvér Geliosphere. To bolo ale zapríčinené najmä tým, že Geliosphere využíva dvojitú presnosť, kým optimalizovaný program používal odporúčanie statickej analýzy pre zmenu presnosti premenných.

C. Nový distribuovaný systém

Pre verifikáciu nového distribuovaného systému sme sa rozhodli použiť 1D B-p model. Vstupné parametre do simulácie bola rýchlosť slnečného vetra $V = 400 km/s$, veľkosť difúzneho koeficienta $K_0 = 5 \times 10^{23} cm^2/s$, veľkosť časového kroku $dt = 5.0s$ pre 1.6 miliardy simulačných trajektórií častíc.

V rámci verifikácie sme sa zamerali na použitý typ paralelizmu a či môže jemnozrný paralelizmus narušiť alebo zlepšiť prípadný výkon celého systému. Tabuľka V obsahuje výsledky testovania granularity v distribuovanom systéme. Jemnozrný paralelizmus nám zlepšil výkon o približne 22%. Znížil vykonávací čas zo 70 na 56 minút v distribuovanom systéme.

Tieto zmeny nám umožňujú zaviesť a implementovať oveľa efektívnejšie metódy na rýchlejšiu prácu pri vývoji a implementácii modelov modulácie kozmického žiarenia v heliosfére.

VI. ZÁVER

Hlavným cieľom tejto práce bolo navrhnúť systém, ktorý by umožňoval rýchlejší vývoj modelov vo fyzike. Prezentované návrhy riešia problémy s vývojom a testovaním modelov modulácie kozmického žiarenia v heliosfére. Softvér Geliosphere poskytuje akceleráciu vykonávacích časov simulácii, spolu s GPU implementáciami je schopný redukovat vykonávací čas z dní a hodín na hodiny a minúty. Koncept softvéru Atmosphere spolu s dostupnou možnosťou transformácie kódu v SymPy do doménovo-špecifického jazyka softvéru Atmosphere by mal odstrániť potrebu ručnej paralelizácie implementácie vývojárom. Koncept integrácie softvérov Geliosphere a Atmosphere do nového distribuovaného systému môže ešte ďalej znížiť vykonávací čas a čas na vývoj týchto modelov.

Hlavné prínosy práce v oblasti štúdia:

- **Publikácia softvéru Geliosphere ako open-source softvér**

V oblasti modelov modulácie kozmického žiarenia v heliosfére bola publikovaná iba jediná implementácia modelov, SOLARPROP, ktorý bol publikovaný v roku 2016. Ostatní autori zvyčajne zverejnia svoje modely, ale nezverejnia špecifické implementačné detaily, ktoré sa na nich viažu. Rozhodli sme sa navrhnuť, implementovať a zverejniť softvér Geliosphere s niekoľkými základnými modelmi - 1D f-p, 1D b-p, model založený na modele SOLARPROP standard 2D a Geliosphere 2D model. Každý model má svoju CPU a GPU implementáciu. V rámci heliosferickej oblasti, chceme zdôrazniť nasledujúce prínosy:

– **GPU implementácie**

GPU implementácie modelov modulácie kozmického žiarenia v heliosfére sa ukázali ako prínosné pre simulovanie trajektórií častíc kozmického žiarenia. Akcelerácie GPU implementácie pre SOLARPROP standard 2D modelu v porovnaní s ekvivalentnou implementáciou na CPU sa pohybovala na úrovni 3 až 34. V prípade Geliosphere 2D modelu, akcelerácia GPU implementácie sa pohybovala od 24 do 85 v porovnaní s paralelnou CPU verziou s identickými vstupnými parametrami.

– **Skripty pre automatizáciu porovnávaní s dátami z Ulysses misie**

Pripravili sme skripty, ktoré sme publikovali spolu so softvérom Geliosphere, ktoré zjednodušujú porovnávanie výsledkov z Geliosphere s meraniami detektora Ulysses KET P190. Skripty umožňujú spustenie Geliosphere v dávkovom behu, pripraví dáta na vizualizáciu a vygenerujú graf na porovnanie. Tieto skripty môžu byť využité napríklad na vyhodnocovanie nových modelov, ktoré môžu byť integrované do softvéru Geliosphere.

– **Verifikácia presnosti**

Verifikovali sme presnosť prezentovaného Geliosphere 2D modelu. Bol overený voči podobnému SOLARPROP standard 2D modelu, ale taktiež aj voči výsledkom z detektora Ulysses KET P190. Navyše, verifikovali sme aj to, či môže jednoduchá presnosť použitá v GPU implementácii znižovať presnosť výsledných spektier. Porovnanie výsledkov GPU a CPU implementácie Geliosphere 2D modelu ukázalo, že medzi výsledkami boli minimálne rozdiely, ktoré boli spôsobené najmä štatistickou chybou.

• **Návrh softvéru Autosphere**

Softvér Autosphere sa zameriava na zníženie vykonávacieho času potrebného na vývoj. Transformuje vstupný kód v doméno-špecifickom jazyku tak, aby výsledný vygenerovaný program bol paralelnou implementáciou týchto definovaných modelov. Návrh softvéru Autosphere využíva doméno-špecifický jazyk veľmi podobný jazykom C/C++, genetické algoritmy pre lepšiu optimalizáciu výstupných programov a

schopnosť generovať kód na viacero platforiem. Koncept transformácie vstupu založeného na knižnici SymPy do doméno-špecifického kódu umožňuje vývoj modelov modulácie kozmického žiarenia v heliosfére bez potreby písať implementáciu.

• **Návrh nového distribuovaného systému**

Koncept nového distribuovaného systému sa zameriava na viac efektívnu distribúciu častí simulácie, viac zabezpečenú architektúru a možné nasadenie v cloudovom prostredí. Koncept taktiež zahŕňa integráciu softvérov Geliosphere a Autosphere do nového distribuovaného systému. To by malo umožniť ešte rýchlejší vývoj a integráciu modelov modulácie kozmického žiarenia v heliosfére.

• **Verifikácia kľúčových konceptov softvéru Autosphere**

Verifikovali sme kľúčové koncepty softvéru Autosphere na prototyp. Verifikácia bola rozdelená do dvoch častí: verifikácia GPU implementácia a verifikácia CPU implementácia. Verifikácie oboch verzií implementácii ukázala, že softvér AUtosphere dokáže generovať optimalizované paralelné verzie, ktoré sú vhodné na ďalší vývoj. Verifikácia CPU implementácia a neskoršie porovnanie so softvérom Geliosphere odhalili isté časti kódu, ktoré môžu byť ďalej akcelerované.

• **Verifikácia kľúčových konceptov nového distribuovaného systému**

Čiastočne sme verifikovali kľúčové koncepty nového distribuovaného systému. Najvýraznejšia zmena v systéme bola zmena z hrubozrnej na jemnozrnú paralelizáciu. Počas testovania sme pozorovali zlepšenie výkonu o 22% oproti pôvodnému distribuovanému systému.

• **Všeobecná platforma pre vývoj nových heliosferických modelov**

Nový distribuovaný systém s integráciou softvéru Autosphere spolu s podporou vstupu zo SymPy a integráciou Jupyter notebookov obsahujúcich definície modelov v SymPy by mohol umožniť vývoj nových modelov priamo v distribuovanom systéme. Jupyter notebook môže obsahovať priamo definíciu, ale aj dokumentáciu k danému modelu. Tieto kombinované faktory môžu umožniť novému distribuovanému systému stať sa všeobecnou platformou pre vývoj nových modelov modulácie kozmického žiarenia v heliosfére.

V ďalšej práci by sme sa radi zamerali na integráciu prezentovaných konceptov do softvéru Autosphere a nového distribuovaného systému. Niektoré z požadovaných vlastností sme nezahrnuli v prezentovaných konceptoch. Môžu byť navrhnuté a implementované v rámci ďalšej práce a to menovite:

• **Navrhnuť a integrovať vyhodnocovanie simulácií do softvéru Autosphere**

Integrovanie tejto vlastnosti by umožnilo v novom integrovanom systéme ľahšie vyhodnocovanie výstupných súborov zo simulácií. Taktiež by to urýchlilo vývoj modelov.

- **Navrhnuť a integrovať generické spracovania energetického spektra do binov v softvère Geliosphere**

V tomto momente, delenie na biny je adaptované na základe vstupných energetických binov alebo výstupných energetických binov, ktoré sa zhodujú s experimentmi ako Ulysses KET P190 alebo AMS-02. Generický prístup by umožnil ľahšie rozšírenie softvéru Geliosphere a ďalšie porovnanie výsledkov voči ďalším experimentom.

- **Rozšíriť možnosti vstupov do programov generovaných v softvère Atmosphere**

Výstupné programy zo softvéru Atmosphere podporujú základné dátové typy, ktoré môžeme zadávať ako vstupy do simulácii. Mohli by sme rozšíriť tieto možnosti tak, aby výsledné programy mohli akceptovať vstup z komplexnejších dátových štruktúr, alebo súborov.

- **Podpora dávkového režimu v softvère Geliosphere, výstupného programu zo softvéru Atmosphere a novom distribuovanom systéme**

V aktuálnych implementáciach nepodporujeme možnosť dávkového spúšťania simulácii. Táto vlastnosť je v prípade softvéru Geliosphere súčasťou skriptov, ktoré slúžia na porovnanie výsledkov simulácii s výsledkami z Ulysses KET P190. Integrácia do softvéru Geliosphere, Atmosphere a nového distribuovaného systému by umožňovala ľahšiu verifikáciu modelov a porovnanie výsledkov voči experimentom, alebo ďalším modelom.

- **Poskytnúť ďalšie vizualizačné nástroje v novom distribuovanom systéme**

S podporou dávkového spúšťania simulácii, by sme mohli integrovať viac možností pre vizualizáciu simulácii. Jedným z príkladov je integrácia porovnávania simulácii z viacerých rokov na vybranej energii. Táto vlastnosť by bola integrovaná do webovej aplikácie a umožnila by používateľom vizualizáciu.

Vo všeobecnosti, v ďalšej práci by sme sa mali zamerať na propagáciu navrhnutých nástrojov, ich prototypov, čo by nám umožňovali získať väčšiu spätnú väzbu od reálnych používateľov.

REFERENCES

- [1] KLEIST, Daryl Timothy. An evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. University of Maryland, College Park, 2012.
- [2] REINERT, D., et al. DWD database reference for the global and regional ICON and ICON-EPS forecasting system. Technical report Version 2.1. 8, Deutscher Wetterdienst, 2020.
- [3] POTGIETER, Marius S. Solar modulation of cosmic rays. *Living Reviews in Solar Physics*, 2013, 10: 1-66.
- [4] PARKER, En N. The passage of energetic charged particles through interplanetary space. *Planetary and Space Science*, 1965, 13.1: 9-49.
- [5] FLYNN, Michael J. Very high-speed computing systems. *Proceedings of the IEEE*, 1966, 54.12: 1901-1909.
- [6] FLYNN, Michael J. Some computer organizations and their effectiveness. *IEEE transactions on computers*, 1972, 100.9: 948-960.
- [7] BOBIK, P., et al. On the forward-backward-in-time approach for Monte Carlo solution of Parker's transport equation: One-dimensional case. *Journal of Geophysical Research: Space Physics*, 2016, 121.5: 3920-3930.
- [8] DUNZLAFF, P.; STRAUSS, R. D.; POTGIETER, M. S. Solving Parker's transport equation with stochastic differential equations on GPUs. *Computer Physics Communications*, 2015, 192: 156-165.

- [9] KAPPL, Rolf. SOLARPROP: Charge-sign dependent solar modulation for everyone. *Computer Physics Communications*, 2016, 207: 386-399.
- [10] SOLANIK, Michal; BOBIK, Pavol; GENČI, Ján. Cosmic rays modulation in heliosphere models on GPU. 2021.
- [11] CHANDRA, Rohit, et al. Parallel programming in OpenMP. Morgan Kaufmann, 2001.
- [12] WIENKE, Sandra, et al. OpenACC—first experiences with real-world applications. In: Euro-Par 2012 Parallel Processing: 18th International Conference, Euro-Par 2012, Rhodes Island, Greece, August 27-31, 2012. *Proceedings 18*. Springer Berlin Heidelberg, 2012. p. 859-870.
- [13] YAVUZ, Esin; TURNER, James; NOWOTNY, Thomas. GeNN: a code generation framework for accelerated brain simulations. *Scientific reports*, 2016, 6.1: 1-14.
- [14] LINDSEY, Ben; LESLIE, Matthew; LUK, Wayne. A domain specific language for accelerated multilevel Monte Carlo simulations. In: 2016 IEEE 27th International Conference on Application-specific Systems, Architectures and Processors (ASAP). IEEE, 2016. p. 99-106.
- [15] ANSEL, Jason, et al. Opentuner: An extensible framework for program autotuning. In: *Proceedings of the 23rd international conference on Parallel architectures and compilation*. 2014. p. 303-316.
- [16] USOSKIN, Ilya G.; BAZILEVSKAYA, Galina A.; KOVALTSOV, Genady A. Solar modulation parameter for cosmic rays since 1936 reconstructed from ground-based neutron monitors and ionization chambers. *Journal of Geophysical Research: Space Physics*, 2011, 116.A2.
- [17] KOZA, John R. Genetic programming II: automatic discovery of reusable programs. MIT press, 1994.
- [18] SLANINA, Daniel. Jemnozrný paralelizmus v distribuovanom systéme. Bakalárska práca. 2022.
- [19] SOLANIK, MICHAL. Paralelizácia dvojrozmerných modelov kozmického žiarenia v heliosfére. Diplomová práca 2022.
- [20] GECÁŠEK, Daniel, et al. COR system: A tool to evaluate cosmic ray trajectories in the Earth's magnetosphere. *Advances in Space Research*, 2022, 70.4: 1153-1168.
- [21] BOSCHINI, M. J., et al. Deciphering the local interstellar spectra of primary cosmic-ray species with HelMod. *The Astrophysical Journal*, 2018, 858.1: 61.
- [22] HECK, Dieter, et al. CORSIKA: A Monte Carlo code to simulate extensive air showers. Report fzka, 1998, 6019.11.
- [23] VLADIMIROV, Andrey E., et al. GALPROP WebRun: an internet-based service for calculating galactic cosmic ray propagation and associated photon emissions. *Computer Physics Communications*, 2011, 182.5: 1156-1161.

System ochrany a manažment porúch pre mikrosiete a aktívne distribučné siete

Róbert Štefko

Katedra elektroenergetiky
Fakulta elektrotechniky a informatiky, TUKE
Košice, Slovenská republika
robert.stefko@tuke.sk

Zsolt Čonka

Katedra elektroenergetiky
Fakulta elektrotechniky a informatiky, TUKE
Košice, Slovenská republika
zsolt.conka@tuke.sk

Abstrakt— Predmetom tejto dizertačnej práce je vytvorenie modelu mikrosiete v programe Simulink a návrh nového systému chránenia pre aplikovateľnosť do praxe. Úvod práce obsahuje súčasný prehľad projektov mikrosietí a jednotlivé typy distribučných sietí. Podrobne sú popísané výsledky analýzy pre mikrosieť a distribučnú sieť, ktoré predstavujú základné faktory pre návrh samotného systému chránenia a ďalší výskum. Vytvorený model mikrosiete disponuje OZE a dynamickými záťažami, ktoré sú podrobené rôznym testom a analýzám. Model rovnako disponuje navrhnutými modulmi digitálnych ochrán ako pre smerovú nadprúdovú tak aj rozdielovú funkciu. Navrhnutý systém chránenia bol otestovaný pre rôzne poruchové stavy a následne aj prakticky reálne overený pričom práca poskytuje pozoruhodné výsledky pre jednotlivé testované ochrany.

Kľúčové slová— *Systém chránenia, manažment porúch, mikrosiete, aktívne distribučné siete, digitálne ochranné relé*

Abstract— The subject of this dissertation is the creation of a microgrid model in Simulink and the design of a new protection system for practical applicability. The introduction of the dissertation contains a current overview of microgrid projects and the various types of distribution networks. The analysis results for the microgrid and distribution network are described in detail, which are the fundamental factors for the design of the protection system itself and further research. The developed microgrid model has RES and dynamic loads, subjected to various tests and analyses. The model equally has designed digital protection modules for both directional overcurrent and differential elements. The proposed protection system has been tested for various fault conditions and then practically real verified while the work provides remarkable results for the various tested protections relays.

I. ÚVOD

Systém ochrany je jedným z najdôležitejších súčasti elektrizačnej sústavy a zabezpečuje ochranu zariadeniam a vedeniam a popri tom aj ochranu obslužnému personálu a koncovým užívateľom. Od samotného začiatku využívania elektrickej energie začali vznikať problémy v podobe porúch ako na strane výroby tak aj na strane prenosu a rozvodu. To podmienilo vznik novej oblasti výskumu, a to systémov chránenia. Ako prvé ochranné zariadenia boli používané poistky, ktoré, ale mali nevýhodu v samotnej možnosti

nastavenia, keďže samotná konštrukcia a aj funkčnosť poistky spočíva v tavnom drôte, ktorý je presne navrhnutý na určitú hodnotu elektrického prúdu, pri ktorej ak dôjde k prekročeniu tavný drôt sa roztaví čím preruší obvod. Z tohto dôvodu boli neskôr nahradzané ističmi a ochrannými relé a neskôr už prepojeným systémom niekoľkých ochranných relé až do podoby digitálnych ochrán, ktoré sa používajú dodnes. Hlavnou úlohou systému chránenia je zabezpečiť, aby chránený úsek (vedenie) alebo zariadenie neboli vystavované abnormálnym stavom, ako napríklad preťaženie, alebo skrat, ktoré pôsobia nepriaznivo na izoláciu a tým výrazne skracujú životnosť zariadenia, alebo chráneného úseku pričom môže dôjsť aj k ohrozeniu zdravia obsluhy a okolia. Pri návrhu systému chránenia je, preto potrebné poznať vzťah medzi pôsobením zariadenia, alebo chráneného úseku na okolie a okolia na zariadenie alebo chránený úsek. Preto je potrebné mať, čo najviac informácií o danom mieste pre návrh nového systému chránenia. V prípade neúplných alebo nejasných informácií o jednotlivých častiach, by sa ani nemalo začínať s návrhom systému chránenia, prípadne prísnejšie posudzovať jednotlivé časti.

Správne zareagovanie ochranného systému zabezpečí minimalizácia odpájania spotrebiteľov od elektrickej energie čím sa výrazne zvýši spoľahlivosť samotného systému. Z tohto dôvodu je možné predpokladať výrazný rozvoj v oblasti monitorovania pričom systémy chránenia pre ich aktívne využitie v praxi pri nasadzovaní miestnych sietí alebo mikrosietí definovaných, ako „Microgrids“ budú zohrávať dôležitú úlohu, čo povedie k výraznému skvalitneniu poskytovaných služieb a k neprerušovanému napájaniu spotrebiteľov, čo môže viesť v určitých prípadoch aj k záchrane života.

Neustále zvyšujúci sa dopyt po elektrickej energii podmieňuje rozvoj nového systému inteligentnej siete definovanej ako „Smart Grid“. Pre čoraz väčšie výkyvy v cene elektrickej energie získavajú na popularite obnoviteľné zdroje energie, ktoré začínajú výrazne prenikať do distribučných sietí pre napäťové hladiny NN a VN. Toto smerovanie naznačuje blížiace sa aplikovanie mikrosietí do prevádzky, čo povedie ku globálnej zmene konvenčného systému. Pre stále rozvíjajúcu sa problematiku je potrebné zdieľanie nových myšlienok a nápadov pre napredovanie vývoja takéhoto systému.

II. DEFINÍCIA MIKROSÍETÍ

Miestne siete alebo Mikrosiete môžeme definovať ako malé lokálne distribučné siete, ktoré napájajú spotrebiteľov elektrickou energiou, ktorá je vyrábaná pomocou distribuovaných zdrojov elektrickej energie. Takéto siete musia spĺňať predpoklad sebestačnosti vo výrobe elektrickej energie (prevádzka v ostrovnom režime), teda je potrebné, aby takéto siete mali podľa geografického umiestnenia vhodne zvolený energetický mix lokálnych zdrojov tak, aby zabezpečili dostatočnú výrobu pre miestnu spotrebu. Základnou výhodou mikrosietí v porovnaní s aktuálne využívanou elektrizačnou sústavou je v možnosti sa pripojiť a odpojiť od prenosovej a distribučnej sústavy, čiže s výhodou vedľa využívať výhodu režimu ostrovnej prevádzky [1].

A. Definícia Európskej únie mikrosiete

Mikrosieť zahŕňa nízkonapäťové distribučné systémy s DZE, akumuláčnymi zariadeniami, systémom skladovania elektrickej energie a flexibilnými záťažami. Pričom takéto systémy môžu fungovať buď pripojené alebo odpojené od hlavnej sústavy [1].

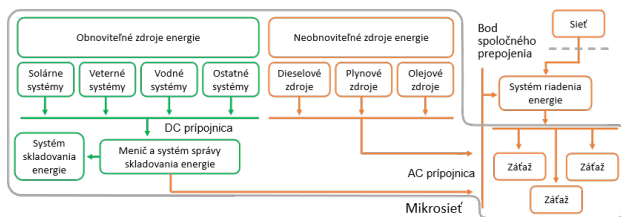
B. Definícia Spojených štátov amerických mikrosiete

Mikrosieť predstavuje skupinu vzájomne prepojených záťaží a DZE v rámci jasne definovaných hraníc, ktoré pôsobia ako jedna ovládateľná jednotka vzhľadom na sústavu. Mikrosieť sa môže pripojiť a odpojiť od sústavy, aby mohla využívať oba režimy (ako v pripojenom režime k sústave tak aj v ostrovnom režime) [1].

Mikrosieť možno jednoducho definovať, aj ako moderný autonómny systém distribúcie elektrickej energie, ktorý využíva primárne lokálne (miestne) zdroje energie. Vo všetkých definíciách sa používa pojem „distributed generation sources“ a vzťahuje sa na výrobu elektrickej energie umiestnenú v blízkosti miesta spotreby alebo záťaže [1].

C. Štruktúra mikrosietí

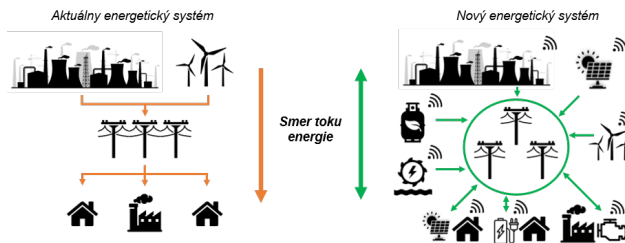
Mikrosieť je prepojenie rôznych obnoviteľných a neobnoviteľných DZE so zariadeniami schopnými akumulovať elektrickú energiu a rôznych typov záťaží. Štruktúra mikrosiete zahŕňa aj prepojenie na ostatné mikrosiete a sústavu spolu so systémom riadenia, stability a komunikácie medzi spotrebiteľom a výrobcom. Prepojenie medzi sústavou a mikrosieťou je známe, ako bod spoločného prepojenia tak, ako zobrazuje Obr. 1. Mikrosiete zahŕňajú rovnako technológie konvertorov, systémy riadenia výkonu, stability a efektívnej prevádzky [1].



Obr. 1 Základná štruktúra mikrosietí

D. Porovnanie súčasnej distribučnej sústavy s mikrosieťou

Zväčšujúci sa podiel OZE, ktoré prenikajú do distribučnej sústavy zásadne mení konvenčný koncept centrálnej výroby elektrickej energie k miestnej (lokálnej) výrobe a spotrebe. Integrácia OZE znižuje straty pri prenose a zvyšuje prevádzkovú spoľahlivosť, pričom vznikajú nové problémy spojené s touto zmenou hlavne v systémoch chránenia, riadenia a lokalizácie porúch. Z tohto dôvodu je vývoj stále potrebný pre uplatnenie takýchto mikrosietí do budúcnosti [2].

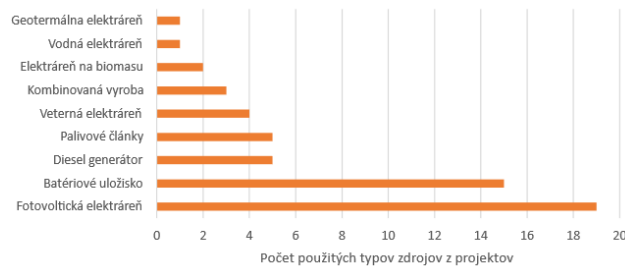


Obr. 2 Rozdiely v smeroch toku energie v elektrizačnej sústave

Pri porovnaní konvenčných (aktuálnych) energetických systémov s mikrosieťovými tak, ako to zobrazuje Obr. 2, je výrazný rozdiel v smere toku energie. V aktuálnom energetickom systéme je vidieť smer toku elektrickej energie od veľkých elektrární (výroby) cez prenos a rozvod elektrickej energie pomocou dlhých vedení až k spotrebiteľom (záťaž). Pripájaním DZE do sústavy sa však tento tradičný radiálny smer toku energie mení na obojsmerný tok elektrickej energie z dôvodu premiestnenia výroby priamo do miesta spotreby alebo do jej blízkosti cez distribučný systém krátkych vedení. Pri aplikácii mikrosietí však vzniká otázka: Ako vhodne zvoliť miestny (lokálny) energetický mix zdrojov tak, aby zabezpečil optimálnu, udržateľnú a hlavne dostatočnú flexibilitu výroby elektrickej energie za všetkých poveternostných podmienok? To prináša nové výzvy pri budovaní a vývoji mikrosietí. Nasledujúce podkapitoly budú zamerané na typy mikrosietí podľa spôsobu využitia.

E. Zhodnotenie energetického mixu zdrojov pre mikrosiete a nanosiete

Analýza opísaných svetových projektov v tejto kapitole naznačuje smerovanie aktuálne nasadzovaných energetických zdrojov do prevádzky pre mikrosiete a nanosiete v rôznych podnebných pásmach a nadmorských výškach.



Obr. 3 Analýza nasadzovaných OZE v mikrosietach

Z analýzy vychádza ako najčastejšie nasadzovaný zdroj práve fotovoltaická elektrárňa spolu s batériovým systémom ako záložným zdrojom. Následne, ako záložné zdroje boli použité Palivové články a diesel generátory, čo potvrdzuje aj Obr. 3.

III. ANALÝZA MIKROSÍETÍ A DISTRIBUČNÝCH SIETÍ

Pre lepšie pochopenie všetkých vzájomných vzťahov a pôsobení je potrebné mať, čo najviac informácií o danej problematike. Z tohto dôvodu predmetom tejto kapitoly bude, čo najpodrobnejšie opísanie problematiky OZE a prevádzky distribučnej sústavy a už realizovaných mikrosietí.

A. Stabilita a kvalita elektrickej energie

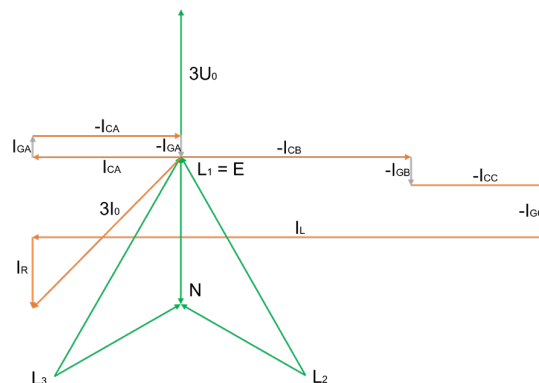
Problémy so stabilitou sú v mikrosieťach v porovnaní s aktuálne využívanou prenosovou a distribučnou sústavou častejšie, pretože mikrosiete majú oveľa menej výkonné energetické zdroje, ale na druhej strane tieto zdroje majú oveľa menej viazanú zotrvačnosť a akumuláciu elektrickej energie v kinetickom pohybe, čo sa prejaví hlavne na stabilite [3].

Reguláciou činného a jalového výkonu možno regulovať napätie a frekvenciu systému. V mikrosieťach je dôležitá regulácia frekvencie siete. Okrem otázky stability je hlavným problémom mikrosietí aj kvalita elektrickej energie. Na základe regulácie napätia a frekvencie sa stabilita rozdeľuje na dva typy: stabilitu napätia a frekvencie. K zmenám frekvencie systému dochádza v dôsledku prechodných javov a porúch, napríklad pri prechode mikrosiete z ostrovného režimu do sieťového režimu, pričom tieto zmeny je potrebné eliminovať, čo sa označuje ako stabilita frekvencie. Problém so stabilitou napätia v mikrosieťach môže byť spôsobený prechodnými javmi ako napr. spínacími úkonmi alebo v dôsledku zmeny v dynamickom zaťažení (jalového výkonu). To, ale vieme regulovať pomocou odľahčenia vedenia, prepínaním odbočiek distribučných TR a pripájaním tlmiviek a kondenzátorov [4]-[6].

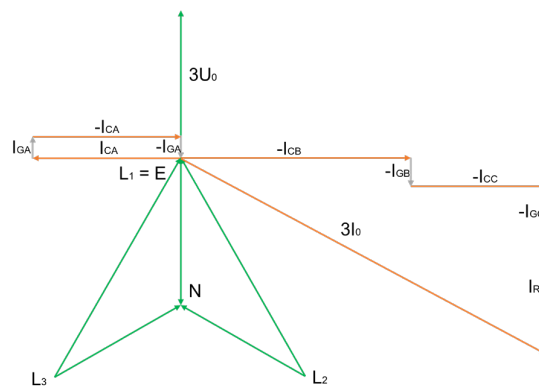
B. Zmeny napätia a prúdu kompenzovanej siete cez tlmivku

Pri vzniku zemného spojenia dôjde k zmene napätových pomerov pričom najvýznamnejšie sa prejaví netočivá zložka napätia U_0 , ktorá mala pred vznikom poruchy nulovú hodnotu a ktorá pri vzniku poruchy vzrástla na úroveň fázového napätia a má opačné smerovanie fázora ako porušená fáza pred poruchou. Pričom vzniká napätová nesymetria v sieti. Neporušenými fázami tečú poruchové prúdy s príspevkom admitančného prúdu. Zemná ochrana zdravého vývodu by v takomto prípade zaznamenala trojnásobok netočivej zložky napätia $3U_0$ a prúd $3I_0$, ktorý pozostáva z príspevku admitančného prúdu zdravých fáz. Fázový posun medzi napätím U_0 a prúdom I_0 sa pohybuje v blízkosti 90° , výsledky merania poukazujú, že tento uhol sa pohybuje v rozsahu 86° až 89° v závislosti od kvality izolácie siete a veľkosti príspevku prúdu konduktancie. Zemná ochrana poruchového vývodu by v takomto prípade zaznamenala trojnásobok netočivej zložky napätia $3U_0$ a prúd $3I_0$. Fázový posun medzi napätím U_0 a prúdom I_0 je mierne väčší ako 90° , výsledky merania poukazujú, že tento uhol sa pohybuje v rozsahu 91° až 97° v závislosti od kvality izolácie siete a

veľkosti príspevku prúdu konduktancie. Ak zemné spojenie v kompenzovanej sieti trvá dlhšie ako 1 s, automatika pripojí sekundárny odporník. V prípade kovového zemného spojenia sa napätové pomery takmer nezmenia. Nezmení sa ani tok prúdov zdravých vývodov. Pripnutím sekundárneho odporníka dôjde k zmene toku prúdu v porušenom vývode, ktorým začne tečť prúd vnútený odporníkom. Prúd vnútený odporníkom je v protifáze s napätím $3U_0$. Po pripojení odporníka zemná ochrana poruchového vývodu zaznamená fázový posun medzi $3U_0$ a $3I_0$, ktorý je výrazne väčší ako 90° . Veľkosť tohto uhla závisí na vyladení tlmivky a na veľkosti prúdu vnúteného odporníkom. Ak bude tlmivka podladená o veľkosť príspevku prúdu susceptancie jednej fázy môže tento uhol dosiahnuť až hodnotu 180° [7]. Pre porovnanie zmeny fázorov prúdu je zobrazené porovnanie na Obr. 4 pre sieť s uzemneným uzlom transformátora cez tlmivku a cez odporník.



Sieť s neúčinne uzemneným uzlom transformátora cez tlmivku



Sieť s neúčinne uzemneným uzlom transformátora cez odporník

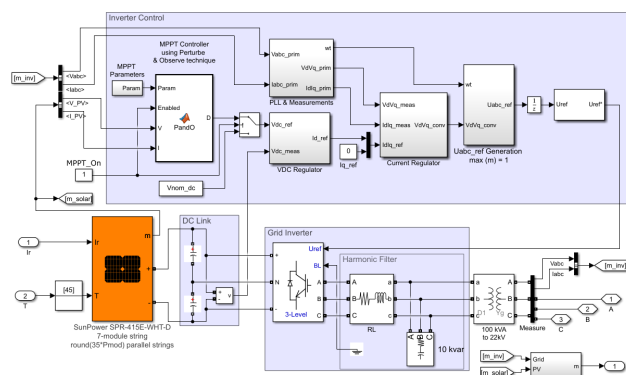
Obr. 4 Fázorový diagram prúdov a napätí pre poruchový vývod [7]

C. Vplyv obnoviteľných zdrojov energie na sieť

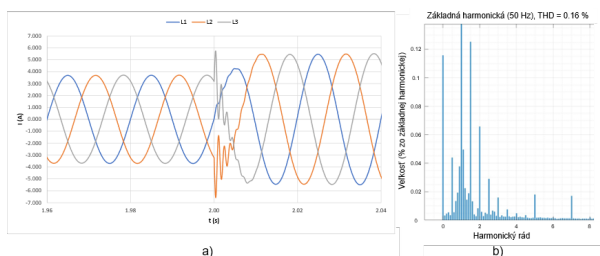
Podkapitola sa zaoberá hardvérovými limitmi striedačov a poukazuje na ich základné vlastnosti, modelovanie diesel generátora, batériového systému, tepelnej elektrárne a vodnej elektrárne v prostredí Matlab & Simulink, v ktorom je realizovaná aj odskúšanie pôsobenia nového systému chránenia na vzniknutý poruchový stav.

D. Model fotovoltaického systému

Aktuálne pre modelovanie energetických zdrojov existujú rôzne programy pre simulovanie zariadení. Práca bude využívať modelovacie prostredie Matlab & Simulink, v ktorom boli vytvorené modely najrozšírejších typov energetických zdrojov. Pri porovnaní modelov budeme uvažovať so fotovoltaickými systémami, dieselovými generátormi, batériovými systémami, tepelnými elektrárnami a vodnými elektrárnami. Model fotovoltaického systému bol otestovaný pri výkone 100 kWp. Fotovoltaické pole má sedem modulov na jeden reťazec zapojených do série, ktoré sú paralelne zapojené do tridsiatich piatich reťazcov. Elektrický potenciál jednosmerného napätia je 480 V a výstupné striedavé napätie z meniča je 250 V, ktoré sa nakoniec pomocou zvyšovacieho transformátora transformuje na 22 kV. Na Obr. 5 je znázornený schematický model fotovoltaického systému. Obr. 6 a) zobrazuje charakteristické prevádzkové parametre bežného stavu a poruchového stavu. Hodnota simulovaného menovitého prúdu bola menšia ako 3.75 A [8].



Obr. 5 Schéma modelu fotovoltaického systému [8]

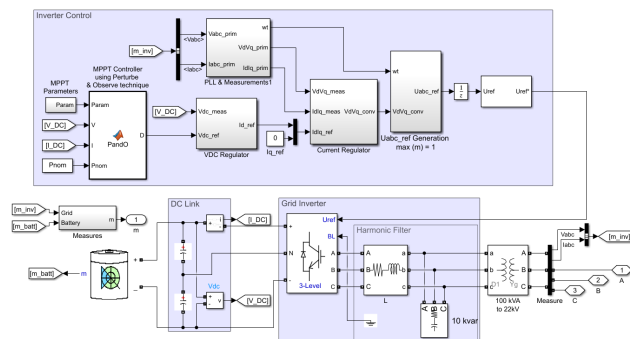


Obr. 6 a) Hodnoty prúdu fotovoltaického modelu počas poruchového stavu [8]; b) FFT analýza výstupného napätia modelu fotovoltaického systému [8]

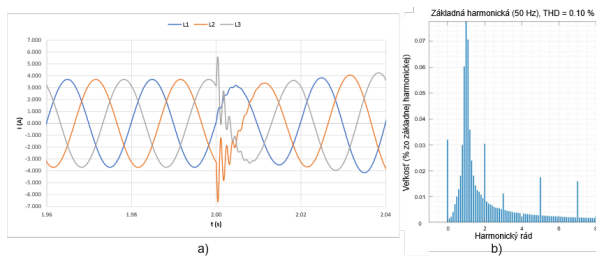
Simulovaná hodnota poruchového prúdu sa ustálila na hodnote 5.62 A. Najvyššia impulzná hodnota poruchového prúdu bola 6.606 A. Obr. 6 b) znázorňuje rýchlu Fourierovu transformáciu (FFT) napätia zo fotovoltaického systému pre kvalitatívnu analýzu. Výstupné napätie modelu malo hodnotu celkového harmonického skreslenia (THD) 0.16 %. Ako je znázornené na Obr. 6 b), je možné vidieť typické vyššie harmonické a medzi harmonické spôsobené meničom spolu s jednosmernou zložkou, pričom najvyššia hodnota harmonickej nepresiahla hodnotu 0.13 %.

E. Model batériového úložiska

Model batériového systému skladovania energie bol založený na rovnakom princípe ako model fotovoltaického systému. Model rovnako využíva zvyšovací transformátor na vyvedenie elektrickej energie do siete. Batériový modul sa môže pridať aj do fotovoltaického systému na zmiernenie výkyvov spôsobených vplyvom počasia. Na Obr. 7 je znázornený schematický model BESS. Obr. 8 a) zobrazuje prevádzkové parametre bežného stavu so stavom poruchy. Hodnota menovitého prúdu bola menšia ako 3.7 A [8].



Obr. 7 Schéma modelu batériového systému [8]

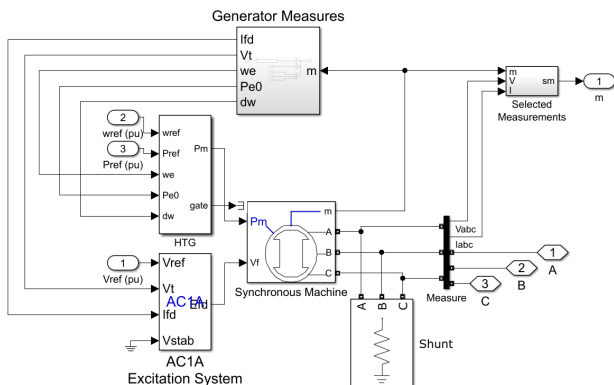


Obr. 8 a) Hodnoty prúdu modelu batériového systému počas poruchového stavu [8]; b) FFT analýza výstupného napätia modelu batériového systému [8]

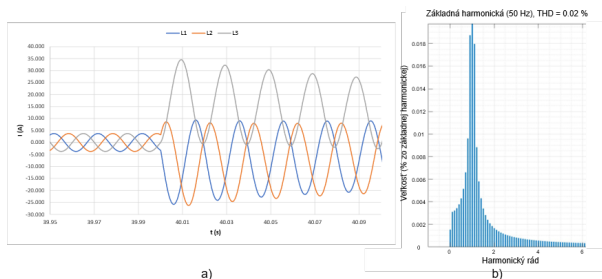
Simulovaná hodnota poruchového prúdu sa ustálila na 5.567 A. Najvyššia impulzná hodnota poruchového prúdu v blízkosti zdroja bola 6.621 A. Obr. 8 b) znázorňuje FFT analýzu napätia modelu batériového systému. Napätie tohto modelu malo hodnotu THD 0.10 %. Ako je znázornené na Obr. 8 b) je možné vidieť výskyt rovnakých harmonických, ako pri analýze modelu fotovoltaického systému z dôvodu použitia rovnakej konštrukcie, pričom jednotlivé hodnoty týchto nepriaznivých deformačných častí nepresahujú hodnotu 0.071 % [8].

F. Model fotovoltaického systému

Model vodnej elektrárne vychádzala z rovnakej konštrukcie ako model dieselového generátora a tepelnej elektrárne. Podobne, ako predchádzajúce modely je určená na priame pripojenie k sieti. Na výrobu elektrickej energie sa v tomto modeli použil synchronný stroj. Na Obr. 9 je znázornený schematický model vodnej elektrárne. Obr. 10 a) zobrazuje porovnanie prevádzkových parametrov bežného stavu so stavom poruchy.



Obr. 9 Schéma modelu vodnej elektrárne [8]



Obr. 10 a) Hodnoty prúdu modelu vodnej elektrárne počas poruchového stavu [8]; b) FFT analýza výstupného napätia modelu batériového systému [8]

Hodnota simulovaného menovitého prúdu bola menšia ako 3.723 A. Simulovaná hodnota poruchového prúdu sa ustálila na hodnote 12.8 A. Najvyššia impulzná simulovaná hodnota poruchového prúdu v blízkosti zdroja bola 34.65 A. Na Obr. 10 b) je znázornená analýza FFT bežného stavu napätia. Napätie tohto modelu malo hodnotu THD 0.02 %. Ako je znázornené na Obr. 10 b) je vidieť niektoré harmonické v blízkosti základnej hodnoty frekvencie 50 Hz, pričom jednotlivé hodnoty týchto rušivých prvkov nepresahujú hodnotu 0.019 % [8].

G. Zhodnotenie výsledkov simulácie

Zo simulovaných výsledkov môžeme zhodnotiť, že vplyv rôznych typov elektrární v tomto prípade modelov majú vplyv na elektrizačnú sústavu, pričom pri využití veľkého počtu fotovoltaických systémov alebo batériových systémov vznikajú problémy s udrжанím sústavy pričom inteligentné siete s rozpadom siete nemajú problémy, pretože rozdelené časti majú byť riadené pomocou systému mikrosiete. Pri používaní tepelných a vodných elektrární doplnenými dieselovými generátormi dosiahneme stabilnejšiu sústavu. Z tohto dôvodu je potrebné určiť vhodnú kombináciu zdrojov energie pre mikrosiete [9].

IV. VÝSKUM NOVÉHO SYSTÉMU CHRÁNENIA MIKROSÍETÍ

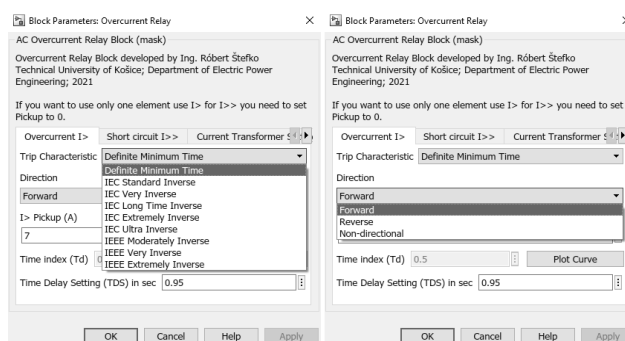
Návrh nového systému chránenia značne ovplyvňuje použitá topológia siete k nej pripojené výrobné zdroje elektrickej energie a samotné profily spotreby elektrickej energie pre miestne záťaž. Pričom ešte stále nie je navrhnutá

jednotná metóda riadenia lokálnych zdrojov pre mikrosiete a aktívne distribučné siete.

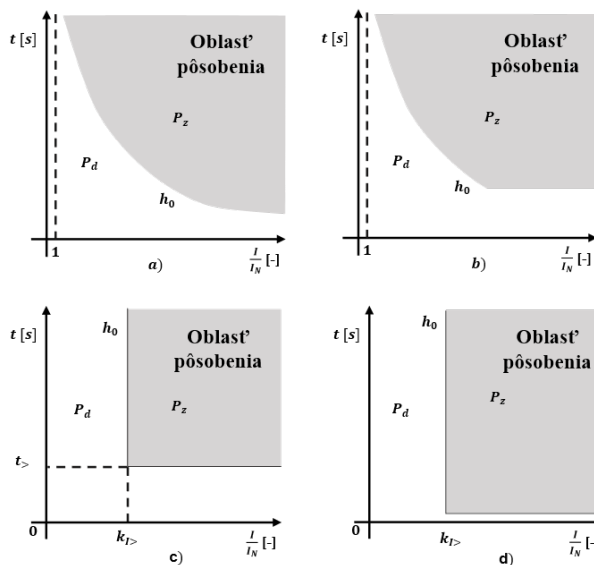
Modely jednotlivých ochranných relé boli vytvorené v programe Simulink (v9.12.0.2039608, MathWorks, CA, USA), ktorý je súčasťou programu Matlab. Jednotlivé modely rôznych typov ochranných relé boli vybrané podľa ich výskytu v distribučných sieťach pre napäťové hladiny VN a NN. Z analýzy výskytu jednotlivých typov boli vybrané štandardné vývodové ochrany (nadprúdová funkcia a smerová nadprúdová funkcia) a rozdielové ochrany.

A. Vývodová ochrana

Pri vývodovej ochrane uvažujeme pre prvý prípad len s využitím nadprúdovej funkcie pre prvý a druhý stupeň s možnosťou časového oneskorenia. Prípadne pre možnosť využitia ďalšieho stupňa, ako varovného stavu pri prekročení dovolenej nominálnej hodnoty prúdu. Pre jednoduchšie nastavenie nadprúdovej ochrany bolo vytvorené aj grafické rozhranie, ktoré je zobrazené na Obr. 11.



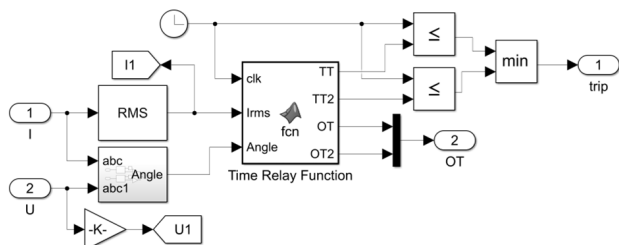
Obr. 11 Grafické rozhranie vývodovej ochrany [10]



Obr. 12 Vypínacie charakteristiky nadprúdovej ochrany a) Časovo závislá, b) Polo závislá, c) Časovo nezávislá, d) Okamžite pôsobiaca [11]
Legenda: Pd - dovolená oblasť; Pz - zakázaná oblasť; h0 - hranica pôsobenia, t> - časové oneskorenie; kI> - prúdové vysunutie

Ochranné relé disponuje možnosťou nastavenia vypínacej charakteristiky podľa normy IEC a IEEE. Toto nastavenie je možné použiť ako pre prvý stupeň tak aj pre druhý stupeň. Strmosť jednotlivých kriviek je možné jednoducho prispôsobiť pomocou rozbehového prúdu, časového násobiteľa alebo oneskorenia v závislosti od zvolenej funkcie. Jednotlivé vypínacie krivky sú zobrazené na Obr. 12. Model tiež disponuje možnosťou nastavenia smerovosti pre nadprúdovú ochranu, pričom túto funkciu je možné jednoducho vypnúť.

Model vývodovej ochrany pozostáva z niekoľkých modulov dostupných zo základnej knižnice programu Simulink, ako aj zobrazuje Obr. 13. Algoritmus ochrany sa vyhodnocuje v module Time Relay Function, ktorého výstup je podľa nastavenia čas vypnutia, ktoré sa porovnáva s časom simulácie. Pri aktivácii smerovosti ochrany (doprednej alebo spätnej) bude ochrana vyhodnocovať smer podľa veľkosti uhla medzi napätím a prúdom.

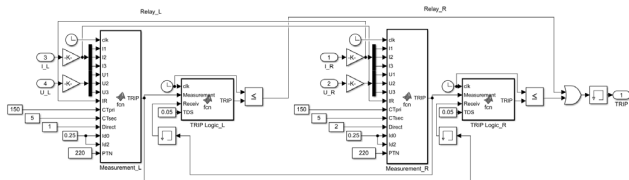


Obr. 13 Model vývodovej ochrany v programe Simulink [10]

Ochrana a riadenie napájačov pomocou REF-615 od spoločnosti ABB sú obľúbené v celosvetovej aplikácii vo všetkých priemyselných odvetviach a aj distribúcii energie. V nastavení sme nastavili zapínací prúd na 10 mA s časom oneskorenia 40 ms, pričom sme zvolili typ vypínacej charakteristiky pre inverznú časovo nezávislú.

B. Smerová vývodová ochrana

Pre doplnenie vzájomnej komunikácie ochranných relé bola potrebná úprava samotného modelu, ktorá bola rozšírená o funkciu „TRIP Logic“ obdobne ako je to zobrazené na Obr. 14. Vzájomná komunikácia využíva komunikačný protokol GOOSE správ s časovou známku. Z dôvodu zistenia výrazných rozdielov v samotnom vyhodnocovacom algoritme pre detekciu poruchového stavu od najrozšírenejších výrobcov. Pre lepšie porovnanie bol vytvorený nový model smerového nadprúdového relé v programe Simulink od spoločnosti SEL, ktoré v posledných rokoch získava na popularite v počte aplikácií.

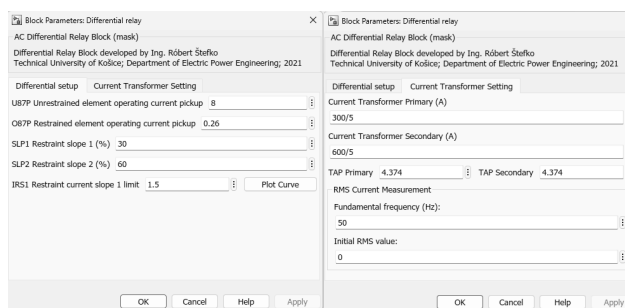


Obr. 14 Model vývodových ochrán s vzájomnou komunikáciou pomocou GOOSE protokolu v programe Simulink

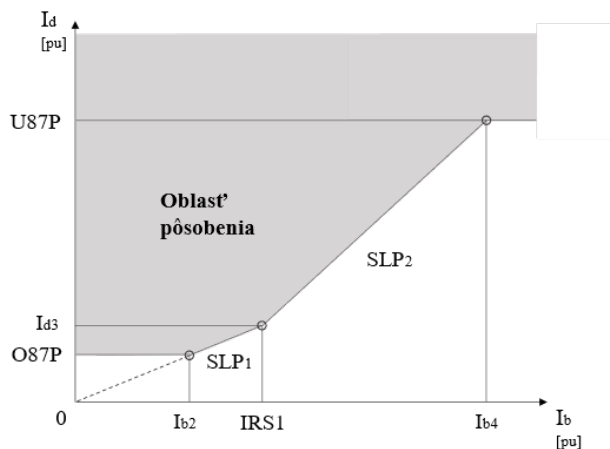
Podľa výsledkov analýzy prevádzky sietí v prípade zemnej poruchy výrazne narastá netočivá zložka poruchového prúdu, pričom sa súčasne mení aj spätná zložka poruchy a práve tento fakt je výrazne zohľadnený pri detekcii vzniku poruchy v danom algoritme. Zatiaľ, čo pri medzifázových poruchoch výrazne narastá súsledná a spätná zložka, ktorá je zohľadnená v samostatnom algoritme pri lepsiú detekciu. Logika algoritmu vyhodnocuje typ poruchy hneď z niekoľkých metód, ktoré následne zohľadňuje a rozhoduje či ide o medzifázovú poruchu. Algoritmus rovnako uvažuje aj s neúčinne uzemnenými sieťami, v tomto prípade model využíva detekciu pomocou netočivej zložky a zvýšenie nerovnomerného zaťaženia fáz.

C. Rozdielová ochrana

Pri rozdielovej ochrane máme jednoduchšie nastavenie, ktoré nám už menej ovplyvňuje hodnota nizkeho príspevku skratového prúdu z OZE. Pričom stále máme možnosť využitia nadprúdovej ochrany pre monitorovanie stavu využitia chráneného vedenia pre monitorovanie preťaženia, ako bolo spomenuté v predošlej podkapitole. Pre jednoduchšie nastavenie rozdielovej ochrany bolo vytvorené rovnako grafické rozhranie, ako aj pri nadprúdovej ochrane, ktoré je zobrazené na Obr. 15. S týmto typom rozdielovej ochrany budeme uvažovať v druhom prípade.

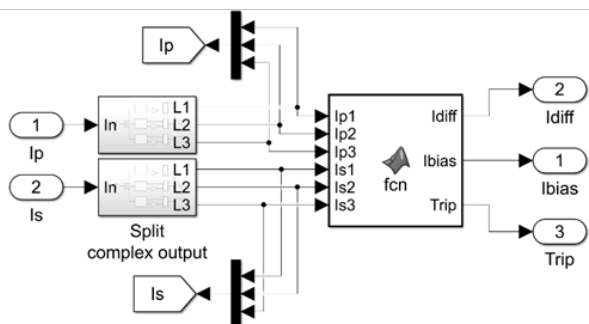


Obr. 15 Grafické rozhranie rozdielovej ochrany [10]



Obr. 16 Vypínacia charakteristika rozdielovej ochrany [11]

Rozdielové relé disponuje možnosťou nastavenia vypínacej charakteristiky jednoducho pomocou zopár bodov („O87P“, „IRSI“, „SLP1“, „SLP2“ a „U87P“). Zvyšné chýbajúce body krivky (Id3, Ib2 a Ib4) podľa Obr. 16 si ochrana dopočíta automaticky. Kompenzáciu prístrojových transformátorov prúdu je možné jednoducho doladiť pomocou „TAP Primary“ a „TAP secondary“, ktorý sa bežne používa na kompenzáciu nepresnosti samotných prevodov PTP a pri transformátoroch pre rôzne prevody prúdov a hodinového uhla. Model rozdielovej ochrany pozostáva z niekoľkých modulov dostupných zo základných knižníc programu Simulink, ako je zobrazené na Obr. 17.



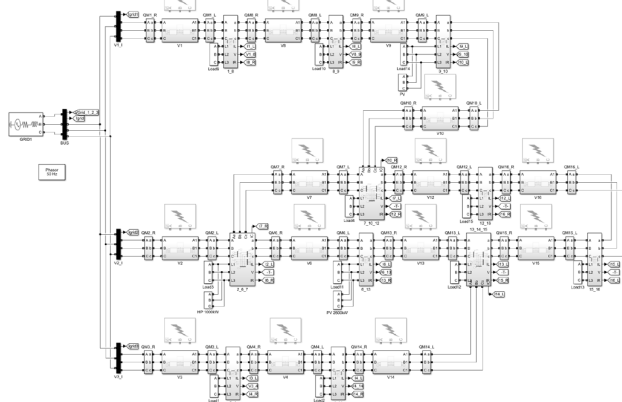
Obr. 17 Model rozdielovej ochrany v programe Simulink [10]

V. OVERENIE FUNKČNOSTI SYSTÉMU CHRÁNENIA

Overenie funkčnosti zvolených ochranných relé bol najprv otestovaný v prostredí programu MATLAB a Simulink. Z tohto dôvodu bola navrhnutá testovacia topológia pre mikrosieť, v ktorej sme uvažovali s rôznou kombináciou zdrojov OZE. Ako konkrétne distribuované zdroje elektrickej energie sme zvolili najnasadzovanejšie zdroje z aktuálne prebiehajúcich alebo už dokončených projektov mikrosietí zhodnotených v prvej kapitole, kde práve batériový systém a fotovoltaická elektrárne majú najväčšie zastúpenie.

A. Topológia testovanej mikrosiete

Topológia mikrosiete pozostáva z troch okružných sietí a štrnástich vedení, ako je znázornené na Obr. 18 pre už vytvorený model v programe Simulink.



Obr. 18 Topológia mikrosiete v programe Simulink

Ďalej pozostáva z dvanástich elektrických staníc, v ktorých sú umiestnené meranie a samotné ochranné relé pre chránenia každého vedenia z oboch strán. Poloha umiestnenia distribuovaných zdrojov elektrickej energie bola zvolená pre dve miesta, a to pre elektrickú stanicu 2_6_7 a elektrickú stanicu 9_10 tak, ako to zobrazuje Obr. 31. V elektrickej stanici 2_6_7 bola umiestnená vodná elektrárň (HP), ktorá bude slúžiť, ako základná báza pre mikrosieť popri napájaní zo siete. V elektrickej stanici 9_10 bude umiestnená pre dva rôzne scenáre buď fotovoltaická elektrárň (PV) alebo batériový systém (BESS). Samozrejme pre aplikovateľnosť ostrovného režimu takejto mikrosiete bude potrebné doplniť aj ďalšie zdroje elektrickej energie, čo povedie k zmene prevádzkových pomerov. Z tohto dôvodu budeme uvažovať len s dvoma typmi zdrojov pre najhorší možný scenár. V danej topológii bolo určených 14 poruchových miest pre každé vedenie.

B. Overenie funkčnosti systému

Reálne overenie výsledkov simulácie bolo overené pre reálne používané ochranné relé konkrétne pre vývodové ochrany terminály SEL-351A, SEL-751A, ABB REF-615, ktorých výskyt je aktuálne v distribučných sieťach pomerne veľký a rozdielovú ochranu ABB REM-545.

- Nadprúdové ochranné relé sa bežne používajú pre chránenie VN vedení, pričom do budúcnosti už budú zohrávať značne menšiu úlohu v systémoch chránenia, ako v súčasnosti z tohto dôvodu je potrebné pri návrhu nového systému chránenia zohľadniť aj základné princípy na chránenie samotných zariadení alebo vedení pre hraničné alarmové stavy a základné funkcie samotného chránenia pre poškodením.
- V navrhnutej topológii uvažujeme, že všetky vedenia sú rovnakého typu. Z tohto dôvodu budú využívať rovnaké nastavenia a, ako základnú nadprúdovú ochranu pre alarmové hlásenia pre preťaženia budeme uvažovať s 5 % dovoleným preťažením. Tak, ako už bolo spomenuté v podkapitole 4.2 pre model vedenia nominálny prúd určený pre použitý typ vodiča je 160 A. Podľa vzorca (3) si vypočítame rozbehovú hodnotu alarmového stavu pre prvý stupeň.

Následné sa určí podľa vzorca (4) nastavená hodnota alarmového stavu, ktorá podľa výpočtu z vzorca (3) bude mať hodnotu 5.89 A. Obdobne postupujeme aj pre nastavenie druhého nadprúdového stupňa pre vypínanie podľa vzorca (5). Kde nastavená hodnota vypínacieho stavu bude mať hodnotu 6.2 A. Pre časové oneskorenie pôsobenia prvého stupňa alarmového stavu uvažujeme s nastavením 0 s a pre časové oneskorenie vypnutia druhého stupňa s nastavením 0.5 s.

Meranie smerových nadprúdových ochrán bolo vykonané pomocou meracieho zariadenia Omicron CMC 256 plus ktorý disponuje dvoma prúdovými výstupmi a jedným napäťovým výstupom ako náhrada silovej časti merania cez PTP. V systéme smerových nadprúdových ochranných relé sa uvažuje o možnosti využitia komunikácie medzi relé na vzájomné blokovanie. Pomocou takejto komunikácie prostredníctvom protokolu GOOSE (IEC 61850) môžeme vzájomne blokovat

smerové relé, pokiaľ nie je splnená podmienka smeru toku poruchového prúdu pre obe ochranné relé. Z dôvodu použitia obojsmernej komunikácie medzi ochrannými relé bude aplikácia smerových ochranných relé v mikrosieťových systémoch kritická. Vzájomná komunikácia prebiehala cez switch Phomix contact pričom sa využili rôzne prenosové média ako cez FTP kábel tak aj optiku. Vzájomné prepájanie vodičov bolo realizované pomocou štandardných meracích káblov s banánovým konektorom.

a) *Meranie smerových nadprúdových ochrán SEL-351A a SEL-751A*

Ochranné relé SEL-751A bola napájaná priamo zo striedavej siete 230 V, pričom SEL-351A už vyžadovala autotransformátor na zníženie hodnoty napätia z 230 V na 120 V. Pri meraní bola použitá len jedna hodnota vzťažného meraného napätia z dôvodu nedostatkov výstupov z dostupného zariadenia CMC 256 plus.

Na nastavenie požadovaných parametrov sa používa softvér AcSELEator QuickSet a pre vzájomnú komunikáciu pomocou protokolu GOOSE je potrebná inštalácia doplnkového softvéru AcSELEator Architect. Obe tieto aplikácie sú postavené na registrovom type zadávania hodnôt pričom majú spoľahlivé a stabilné softvérové prostredie.

Na testovanie funkčnosti sa použila funkcia smerového prvku. V nastavení sme nastavili rozbehový prúd na 0.5 A s okamžitým pôsobením. Obe relé sú vzájomne blokované počiatočnými hodnotami funkcie smerového prvku oboch relé, z ktorých jedna používa dopredné smerovanie a druhé spätné smerovanie.

b) *Meranie smerových nadprúdových ochrán REF-615*

Ochranné relé REF-615 boli napájané priamo zo striedavej siete 230 V. Pri meraní bola použitá len jedna hodnota vzťažného meraného napätia z dôvodu nedostatkov výstupov z dostupného zariadenia CMC 256 plus. Na nastavenie požadovaných parametrov pôsobenia aj vzájomnej komunikácie pomocou protokolu GOOSE používa rovnaký softvér PCM600. Táto aplikácia využíva kombináciu registrového typu zadávania hodnôt a bločkových prepájaní jednotlivých modulov pre nastavovanie pôsobenie logiky samotnej ochrany, pričom nemajú najspoľahlivejšie a stabilné softvérové prostredie, o čom som sa presvedčil, aj počas samotného nastavovania.

Na testovanie funkčnosti sa použila funkcia smerového prvku „DEFHPDEF“ a „DPHHPDOC“. V nastavení sme nastavili rozbehový prúd na 0.1 A s časovým oneskorením pôsobenia 40 ms. Obe relé sú vzájomne blokované počiatočnými hodnotami funkcie smerového prvku oboch relé, z ktorých jedna používa dopredné smerovanie a druhé spätné smerovanie.

c) *Meranie rozdielovej ochrany REM-545*

Ochranné relé REF-545 bolo napájané priamo zo striedavej siete 230 V. Na nastavenie požadovaných

parametrov pôsobenia sa použilo priame nastavenie z predného panelu ochrany, pričom je možné pre nastavenie použiť aj program CAP505. Pre nastavenie parametrov prejdeme do zložky „Protection“ a v nej pokračuje následne do zložky „Diff6G“ a do „Setting group1“.

Na testovanie funkčnosti sa použila štandardná rozdielová funkcia. V nastavení sme nastavili minimálnu rozdielovú hodnotu pre „O87P“ alebo „Basic setting“ na hodnotu 26 %. Výpočet rozbehovej hodnoty prúdu je v rozsahu 5 – 50 %. Setting group1“.

Pre oblasť začiatku prvého zlomu „Ib2“ alebo „Turn-point I“ sme určili stabilizačnú hodnotu $0.867 \times I_n$. Strmost' prvého zlomu „SLP1“ alebo „Starting ratio“ sme nastavili na 30 %. Možnosť nastavenia rozsahu strmosti je pre túto oblasť je 10 – 50 %, zatiaľ, čo rozsah nastavenia začiatku prvého zlomu je 0 – $1 \times I_n$. Setting group1“.

Oblasť druhého zlomu „IRS1“ alebo „Turn-point 2“ má stabilizačnú hodnotu $1.5 \times I_n$, pričom stúpanie „SLP2“ v tejto oblasti má pevne nastavenú hodnotu 100 %. Možnosť nastavenie rozsahu pre začiatok druhej oblasti je 1 – $3 \times I_n$.

Pre zabezpečenie vypnutia pri vysokých rozdielových nadprúdoch použijeme nastavenie „U87P“ alebo „Inst. setting“ pre hodnotu $8 \times I_n$ z možnosti rozsahu nastavenia 5 – $30 \times I_n$.

Pri rozdielnych porovnávacích hodnotách, ochrana vyhodnocuje či sa porucha nachádza vnútri chránenej oblasti a pri vnútornej poruche dáva impulz k vypnutiu. Rozdielové ochrany monitorujú iba svoj ochránený objekt alebo úsek, nepotrebujú sa časovo prispôbiť ďalším ochranám a patria ku základným rýchlym ochranám bez času oneskorenia.

d) *Zhodnotenie možnosti ďalšieho vývoja*

Z priebehu merania a vytvárania simulácie prostredia mikrosiete bol potvrdený predpoklad o vysokej náročnosti pre návrh nového systému chránenia s, čo najväčším využitím už aktuálne nasadených ochranných relé v prevádzke hlavne v distribučných sieťach pre napäťové hladiny VN a NN. Z tohto dôvodu je potrebné do budúca navrhnuť nové zariadenia pre napäťovú úroveň NN, ktoré budú zabezpečovať úlohu predovšetkým monitorovania a centrálného spracovania dát s následným vyhodnotením miesta poruchy pomocou umelej inteligencie. Tieto nové jednotky budú predstavovať zjednodušené zariadenia aktuálnych ochranných relé, ktoré budú disponovať len potrebnými funkciami. Základný predpoklad je, že budú disponovať meraním prúdov a napätí s možnosťou ovládať vypínacie zariadenia pre dané úseky či už vzdušného vedenia alebo káblového vedenia. Pričom problematika centrálného spracovania údajov naráža na odpor pre spoľahlivosť takéhoto systému. Problém je čiastočne riešiteľný využívaním viacerých lokálnych správcom dát.

Aktívne podieľanie sa spotrebiteľov na výrobe elektrickej energie, a teda aj na riadení elektrizačnej sústavy využívajúc hlavne fotovoltických systémov. Ich každoročný nárast a prenikanie do distribučnej siete dosiahne bodu kedy už aktuálne využívaný systém chránenia bude nedostatočný. Z tohto dôvodu budú potrebné zmeny, a to hlavne v distribučnej sieti, ktorej sa tieto zmeny budú najviac týkať. Aktuálny výskum naznačuje, že, ako centralizovaný tak aj

decentralizovaný systém ponúka možnosti pre nasadenie v mikrosieťach. Pri uvažovaní o decentralizovanom systéme značne klesajú prenosové požiadavky na komunikáciu medzi jednotlivými ochrannými relé. Keďže vzájomná komunikácia prebieha len medzi určitými ochrannými relé a aj to len v malom objeme dát, ktoré nevyžadujú až takú robustnú komunikačnú sieť.

VI. ANALÝZA VÝSLEDKOV MERANIA

Výsledky analyzované z testovania všetkých poruchových stavov pre všetky uvažované miesta na jednotlivých vedeniach v programe Simulink poukazujú na spoločný problém vyhodnotenia miesta poruchového stavu, konkrétne podrobnejšie budú dané poruchové miesta opísané v tejto kapitole. Výsledky získané zo simulácie pre navrhnutú topológiu v programe Simulink boli otestované pre ochranné relé SEL-751A, SEL-351A, REF-615 a REM-545 tak, ako bolo zhodnotené v predošlej kapitole. Vo viacerých testoch však bola úspešnejšia logika výrobcu SEL.

a) Vyhodnotenie pôsobenia poruchy na vedení V1

Pre testované poruchové stavy pre vedenie V1 boli použité rôzne hodnoty impedancie poruchy pričom je možné zhodnotiť podľa Tabuľka 2, že daná oblasť nie je 100 % spoľahlivá pri čisto využití len smerových nadprúdových ochrán so vzájomnou komunikáciou. Ako vyplýva z Tabuľka 1 pre trojfázový nekoveý skrat pri uvažovaní s odporom poruchy každej fázy 50 Ω. Spätné ochranné relé pre daný úsek nebolo schopný vyhodnotiť smerovanie z dôvodu nameranej nižšej hodnoty, ako bola nastavená rozbehová hodnota. Zatiaľ, čo pri nižšej hodnote odporu 5 Ω ochrana vyhodnotila smerovanie správne podľa činného výkonu.

Obdobné výsledky boli zistené aj pri trojfázovej poruche v spojení so zemou cez odpor, čo predstavovalo prerhnutie vedenia vo všetkých troch fázach. Pri všetkých ostatných typoch porúch ochrana vyhodnotila smerovanie správne. Napriek tomu podľa týchto výsledkov nie je možné použiť danú ochranu na chránenie takej to rozľahlej mikrosiete, keďže jej úspešnosť bola len 90.91 %. Obdobné výsledky sa zopakovali aj pre vedenie V3.

TABUĽKA 1 TESTOVANÉ HODNOTY PORUCHOVÉHO STAVU PRE TROJFÁZOVÝ SKRAT PRE VEDENIE V1

V L1-E	V L2-E	V L3-E
52.48 V	52.48 V	52.48 V
79.88 °	-40.13 °	-160.10 °
I späť. L1	I späť. L2	I späť. L3
0.131 A	0.132 A	0.131 A
-7.05 °	-127.10 °	112.90 °
I dopr. L1	I dopr. L2	I dopr. L3
12 A	12 A	12 A
77.93 °	-42.07 °	-162.10 °

Zatiaľ, čo ochrana REF-615 využíva konvenčnejší spôsob vyhodnocovania smeru jej výsledky sú v porovnaní z ochranou SEL značne horšie, keďže ochrana nereagovala ani

pri medzifázových skratoch podľa hodnôt z Tabuľka 3, Tabuľka 4 a Tabuľka 5. Z tohto dôvodu sa ochrana REF-615 už vôbec neodporúča pre využitie chránenia mikrosietí.

TABUĽKA 2 TESTOVANÉ TYPY PORÚCH PRE NAVRHNUTÚ TOPOLOGIU PRE LOGIKU SEL:

Vedenie	Typ Poruchy	Impedancia poruchy Ron	Impedancia poruchy Rg	Dopredné pôsobenie	Spätné pôsobenie
V1	L1-G	0.001 Ω	10 Ω a 100 Ω	P, Q, G a P	P, Q, G
	L2-G	0.001 Ω	10 Ω a 100 Ω	P, Q, G a P	P, Q, G
	L3-G	0.001 Ω	10 Ω a 100 Ω	P, Q, G a P	P, Q, G
	L1-L2	10 Ω a 100 Ω	-	P, Q, G a P	P, Q, G
	L1-L3	10 Ω a 100 Ω	-	P, Q, G a P	P, Q, G
	L2-L3	10 Ω a 100 Ω	-	P, Q, G a P	P, Q, G
	L1-L2-G	0.001 Ω	10 Ω a 100 Ω	P, Q, G	P, Q, G
	L1-L3-G	0.001 Ω	10 Ω a 100 Ω	P, Q, G	P, Q, G
	L2-L3-G	0.001 Ω	10 Ω a 100 Ω	P, Q, G	P, Q, G
	L1-L2-L3	5 Ω a 50 Ω	-	P	Pa -
	L1-L2-L3-G	5 Ω	10 Ω a 100 Ω	P	Pa -
	L1-L2-L3-G	50 Ω	10 Ω a 100 Ω	P	-
V8	L1-G	0.001 Ω	10 Ω a 100 Ω	P, Q, G	P, Q, G
	L2-G	0.001 Ω	10 Ω a 100 Ω	P, Q, G	P, Q, G
	L3-G	0.001 Ω	10 Ω a 100 Ω	P, Q, G	P, Q, G
	L1-L2	10 Ω a 100 Ω	-	P, Q, G	P, Q, G
	L1-L3	10 Ω a 100 Ω	-	P, Q, G	P, Q, G
	L2-L3	10 Ω a 100 Ω	-	P, Q, G	P, Q, G
	L1-L2-G	0.001 Ω	10 Ω a 100 Ω	P, Q, G	P, Q, G
	L1-L3-G	0.001 Ω	10 Ω a 100 Ω	P, Q, G	P, Q, G
	L2-L3-G	0.001 Ω	10 Ω a 100 Ω	P, Q, G	P, Q, G
	L1-L2-L3	5 Ω a 50 Ω	-	P	Pa P
	L1-L2-L3-G	5 Ω	10 Ω a 100 Ω	P	Pa P
	L1-L2-L3-G	50 Ω	10 Ω a 100 Ω	P	Pa P

Legenda: P – činný výkon; Q – jalový výkon; G – zemná porucha

TABUĽKA 3 TESTOVANÉ HODNOTY PORUCHOVÉHO STAVU PRE DVOJFÁZOVÝ SKRAT L1-L2 PRE VEDENIE V1

V L1-E	V L2-E	V L3-E
55.95 V	53.40 V	56.05 V
84.83 °	-33.56 °	-152.10 °
I späť. L1	I späť. L2	I späť. L3
0.80 A	0.76 A	1.04 A
65.01 °	-31.36 °	-161.30 °
I dopr. L1	I dopr. L2	I dopr. L3
4.04 A	4.11 A	1.04 A
106.30 °	-59.03 °	-160.60 °

TABUĽKA 4 TESTOVANÉ HODNOTY PORUCHOVÉHO STAVU PRE DVOJFÁZOVÝ SKRAT L1-L3 PRE VEDENIE V1

V L1-E	V L2-E	V L3-E
53.33 V	55.97 V	55.88 V
86.54 °	-32.02 °	-155.10 °
I späť. L1	I späť. L2	I späť. L3
0.74 A	1.03 A	0.8 A

86.01 °	-43.43 °	-177.90 °
I dopr. L1	I dopr. L2	I dopr. L3
4.10 A	1.03 A	4.00 A
60.48 °	-42.70 °	-134.00 °

TABUĽKA 5 TESTOVANÉ HODNOTY PORUCHOVÉHO STAVU PRE DVOJFÁZOVÝ SKRAT L2-L3 PRE VEDENIE V1

V L1-E	V L2-E	V L3-E
56.02 V	55.92 V	53.38 V
87.91 °	-35.14 °	-153.50 °
I späť. L1	I späť. L2	I späť. L3
1.04 A	0.80 A	0.75 A
78.07 °	-55.94 °	-152.10 °
I dopr. L1	I dopr. L2	I dopr. L3
1.04 A	4.02 A	4.11 A
78.81 °	-13.79 °	-179.20 °

Rozdielová ochrana REM-545, ako jediná z testovaných ochrán spoľahlivo pôsobila na všetky typy testovaných porúch podľa výstupov z testovanej topológie. Z tohto dôvodu perspektívne do budúcnosti pre využitie pre mikrosiete je aplikácia rozdielových ochrán vyhovujúca.

b) Vyhodnotenie pôsobenia poruchy na vedení V2

Obdobné výsledky pre ochranu REF-615 boli zistené aj pri testovaní parametrov pre vedenie V2 pri medzifázovej poruche. Pri všetkých ostatných typoch porúch ochrana vyhodnotila smerovanie správne. Napriek tomu podľa týchto výsledkov nie je možné použiť danú ochranu na chránenie vedení, keďže jej úspešnosť bola len 72.73 %. Z tohto môžeme zhrnúť, že konvenčný spôsob vyhodnocovania smerovania nadprúdovej ochrany už nie je dostatočný pre novo nastupujúcu generáciu sietí. Chybné vyhodnotenie testovacích hodnôt touto ochranou sú zobrazené v Tabuľka 6, Tabuľka 7 a Tabuľka 8. Obdobné výsledky sa zopakovali aj pre vedenie V6.

TABUĽKA 6 TESTOVANÉ HODNOTY PORUCHOVÉHO STAVU PRE MEDZIFÁZOVÝ SKRAT L1-L2 PRE VEDENIE V2

V L1-E	V L2-E	V L3-E
56.80 V	54.50 V	56.50 V
85.74 °	-33.41 °	-151.85 °
I späť. L1	I späť. L2	I späť. L3
0.89 A	0.66 A	1.19 A
39.16 °	-41.46 °	-173.80 °
I dopr. L1	I dopr. L2	I dopr. L3
3.79 A	4.07 A	1.19 A
102.10 °	-61.08 °	-173.20 °

TABUĽKA 7 TESTOVANÉ HODNOTY PORUCHOVÉHO STAVU PRE MEDZIFÁZOVÝ SKRAT L1-L3 PRE VEDENIE V2

V L1-E	V L2-E	V L3-E
54.63 V	56.50 V	56.40 V
86.64 °	-31.83 °	-154.20 °

I späť. L1	I späť. L2	I späť. L3
0.65 A	1.19 A	0.90 A
75.84 °	-55.44 °	157.20 °
I dopr. L1	I dopr. L2	I dopr. L3
4.07 A	1.18 A	3.76 A
58.50 °	-54.85 °	-138.30 °

TABUĽKA 8 TESTOVANÉ HODNOTY PORUCHOVÉHO STAVU PRE MEDZIFÁZOVÝ SKRAT L2-L3 PRE VEDENIE V2

V L1-E	V L2-E	V L3-E
56.20 V	56.60 V	54.60 V
88.20 °	-34.10 °	-153.35 °
I späť. L1	I späť. L2	I späť. L3
1.18 A	0.91 A	0.64 A
63.68 °	-84.04 °	-165.80 °
I dopr. L1	I dopr. L2	I dopr. L3
1.18 A	3.75 A	4.07 A
64.29 °	-18.49 °	178.30 °

Rozdielová ochrana REM-545, ako aj smerové nadprúdové ochrany SEL-751A a SEL-351A z testovaných ochrán spoľahlivo pôsobila na všetky typy testovaných porúch podľa výstupov z testovanej topológie. Z tohto dôvodu perspektívne do budúcnosti pre využitie pre mikrosiete je aplikácia rozdielových ochrán vyhovujúca, pričom v určitých prípadoch pre malé mikrosiete je možné uvažovať aj so smerovými nadprúdovými ochranami. Pre všetky ostatné vedenia testované digitálne ochrany vyhovujú.

VII. ZÁVER

Zhodnotenie tohto systému chránenia ponúka niekoľko zaujímavých záverov a zistení týkajúcich sa ochranných relé. Meranie preukázalo určitý potenciál možnosti nasadenia smerových ochranných relé pre chránenie mikrosietí zatial, čo rozdielová ochrana si utvrdila svoje miesto v budúcnosti. Z meraní a simulovaných výstupov je zrejme, že rozhodujúcim parametrom samotného ochranného relé je jeho algoritmus vyhodnocovania, ktorý by mal byť neustále vyvíjaný. Porovnaním vyhodnotenia pôsobenia pre vedenie V1 s V3 je vidieť, že obe testované relé reagovali rôzne na rovnaké testované hodnoty. Pričom nezareagovanie ochrany od výrobcu SEL bolo spôsobené nedosiahnutím rozbehovej hodnoty. Zatial, čo v simulácii ak sme túto hodnotu znížili na požadovanú úroveň ochrana spoľahlivo zareagovala. Testovaná ochrana od ABB REF-615 vykazovala problémy v určených smerovosti z dôvodu malého vychýlenia fázora prúdu od fázová napätia. Čo predstavoval konvenčný spôsob vyhodnocovania smerovosti ochranného relé, pričom ochrana mala hlavne problém vyhodnotiť medzi fázové poruchy. Z tohto dôvodu pre jej aplikovateľnosť do praxe do budúcnosti bude potrebná inovácia algoritmu alebo jej úplné vyradenie pri prechode na mikrosiete, alebo pri jej vytváraní. Rozdielová ochrana v porovnaní z nadprúdovou smerovou ochranou spoľahlivo reagovala na všetky testované stavy pričom jej aplikovateľnosť je obmedzovaná práve z dôvodu vysokých investičných nákladov. Pričom pri zvažovaní aplikácie

WAMS samotné meracie jednotky vychádzajú cenovo priaznivejšie, ale je potrebné podotknúť potrebu centrálného spracovania údajov, ktorý bude disponovať vhodným algoritmom na ktorom vývoji sa ešte stále pracuje. Výskum tejto oblasti potvrdil využiteľnosť už nasadovaných ochranných relé aj pre mikrosiete pričom je potrebná kontrola využívaného algoritmu daného relé či spĺňa zistené skutočnosti. Výsledok experimentu poskytuje užitočný zistenia pre ďalší vývoj a testovanie možnosti využite pre ďalšie typy stavov pričom už je vytvorený realistický model, ktorého funkčnosť bola otestovaná v simulačnom prostredí Simulink.

POĎAKOVANIE

Práca bola podporená MŠVVaŠ v rámci projektu VEGA 1/0757/20 a Slovenskou agentúrou pre výskum a vývoj na základe zmlúv č. APVV-19-0576 a APVV-21-0312.

ZOZNAM POUŽITEJ LITERATÚRY

- [1] G.B. Narejo, et al., "Microgrids: Design, Challenges, and Prospects," 1. vyd. Boca Raton, USA: CRC Press. s. 1–314. ISBN 978-1-0004-5746-9.
- [2] R. Štefko, et al., "Špecifikácia vplyvov nepriaznivo pôsobiach na stabilitu prevádzky ES SR a ich eliminácia," in: ELEKTROENERGETIKA. Vol. 13, no. 1 (2023), s. 15–19. ISSN 1337-6756.
- [3] M. Mešter, "Výpočet skratových prúdov v trojfázových sústavách," 1. vyd. Košice: PRO. s. 1–94. ISBN 80-89057-10-1.
- [4] A. Pradhan, B. Panda, "A Simplified Design and Modeling of Boost Converter for Photovoltaic Sytem," in: International Journal of Electrical and Computer Engineering. Vol. 8, (2018), s. 141–149. ISSN 2088-8707.
- [5] A. Pradhan, B. Panda, "Analysis of ten external factors affecting the performance of PV system," in: 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS). Chennai, India: IEEE, s. 3093–3098. ISBN 978-1-5386-1887-5.
- [6] A. Pradhan, B. Panda, "Design of DC-DC converter for load matching in case of PV system," in: 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS). Chennai, India: IEEE, s. 1002–1007. ISBN 978-1-5386-1887-5.
- [7] M. HORÁK, "Systém chránenia a automatizácie distribučných elektrických sietí 22 kV," 1. vyd. PRO. s. 1–128. ISBN 978-80-89057-37-5.
- [8] R. Štefko et al., "Case Study of Power Plants in the Slovak Republic and Construction of Microgrid and Smart Grid," in: Applied Sciences. Vol. 11, no. 11 (2021), s. 1–22. ISSN 2076-3417.
- [9] S. Manson, E. Mccullough, "Practical Microgrid Protection Solutions: Promises and Challenges," in: IEEE Power and Energy Magazine. Vol. 19, no. 3 (2021), s. 58–69. ISSN 1558-4216.
- [10] R. Štefko et al., "Modeling of Protection Relays and Renewable Energy Sources for Microgrid Systems," in: Acta Electrotechnica et Informatica. Vol. 22, (2022), s. 9–17. ISSN 1335-8243.
- [11] R. Štefko, "Parametrizácia ochranných relé v elektrizačných sústavách," Košice: Technická univerzita v Košiciach. s. 1–115. ISBN 978-80-553-4068-5.

Skúmanie obmedzení prenosu dát v plne optických prenosových sieťach

Norbert Zdravecký

Department of Electronics and Multimedia
Communications,
Faculty of Electrical Engineering and Informatics,
Technical University of Košice,
Košice, Slovakia
norbert.zdravecky@tuke.sk

Ľuboš Ovseník

Department of Electronics and Multimedia
Communications,
Faculty of Electrical Engineering and Informatics,
Technical University of Košice,
Košice, Slovakia
lubos.ovsenik@tuke.sk

Abstrakt—Optické vlákno prenáša dáta vo forme svetelnej vlny, čo znamená, že signály nepodliehajú elektromagnetickým rušeniam, ktoré by mohli ovplyvniť ich prenosové vlastnosti.

Aby sa splnila výzva pre vyššiu kapacitu systému a vyššiu prenosovú rýchlosť používa sa hustý vlnový multiplex (Dense Wavelength Division Multiplexing - DWDM). Je to systém prenosu dát cez optické vlákno, ktorý využíva veľmi blízko rozmiestnené vlnové dĺžky svetla. V DWDM môže každý kanál prenášať dáta rôznej bitovej rýchlosti a bitového formátu, nezávisle od ostatných. Každý kanál má vyhradenú šírku pásma, všetky signály sa dostanú do prijímača v rovnakom čase. Škálovateľné DWDM systémy umožňujú poskytovateľom služieb prispôbiť sa požiadavkám spotrebiteľov. Ide o nákladovo efektívny spôsob zvýšenia kapacity existujúcej siete.

V tejto práci je analýza návrhu 64-kanálového DWDM systému, ktorý má prenosovú rýchlosť 40 Gb/s a prenosovú vzdialenosť do 210 km. Kanály sú s rovnakým rozstupom 100 GHz. Na optimalizáciu 64 kanálového DWDM systému sa používajú rôzne vstupné parametre. Návrh je simulovaný pomocou programu OptiSystem a analýza signálu je založená na bitovej chybovosti (BER), faktoru kvality (Q faktor) a diagramov oka prijateľného signálu. Práca sa ďalej zaoberá problematikou optických zosilňovačov v DWDM sieťach.

Kľúčové slová— DWDM, EDFA, modulácie, vysokorýchlostná optická sieť, UHC DWDM, WDM, zosilňovače

Abstract- Optical fiber transmits data in the form of a light wave, which means that signals are not subject to electromagnetic interference that could affect their transmission properties.

Dense Wavelength Division Multiplexing (DWDM) is used to meet the challenge of higher system capacity and higher transmission speed. It is a fiber optic data transmission system that uses very closely spaced wavelengths of light. In DWDM, each channel can transmit data of a different bit rate and bit format, independently of the others. Each channel has a dedicated bandwidth, all signals reach the receiver at the same time. Scalable DWDM systems allow service providers to adapt to consumer demands. It is a cost-effective way to increase the capacity of an existing network.

This work analyzes the design of a 64-channel DWDM system that has a transmission speed of 40 Gb/s and a transmission distance of up to 210 km. The channels are equally spaced at 100 GHz. Various input parameters are used to optimize a 64 channel DWDM system. The design is simulated using the OptiSystem

program and the signal analysis is based on Bit Error Rate (BER), Quality factor (Q factor) and eye diagrams of the received signal. The work also deals with the issue of optical amplifiers in DWDM networks.

Key words— DWDM, EDFA, modulations, high-speed optical network, UHC DWDM, WDM, amplifiers

I. ÚVOD

Každým rokom pozorujeme stále dramatickejšie zmeny v telekomunikačnom priemysle, ktoré uľahčujú náš životný štýl. Potreba väčšej kapacity siete a väčších prenosových rýchlostí sa neustále zvyšuje, kvôli obrovskému rastu využiteľnosti internetu. Hlavným faktorom je dostatočná šírka pásma a rýchlosť pripojenia pre každého používateľa. Odhady rastu sa v priebehu rokov značne menili. Dnešným dňom počítajú s nárastom okolo 50% ročne. 40 až 80 kanálové optické systémy s rýchlosťami 10 Gb/s už len s ťažkosťami spĺňajú nároky vysokorýchlostných sietí. Preto sú pre diaľkové prenosy nasadzované systémy s prenosovou rýchlosťou 40 a 100 Gb/s na kanál [1,2].

Podniky sa v súčasnosti spoliehajú na vysokorýchlostný prenos dát. Vysokorýchlostné siete sa používajú na prepojenie viacerých miest a využívajú sa na všetky typy trás. So zvýšením dopytu na rýchlosť a spoľahlivosť siete, rastú aj nároky na šírku pásma [3]. Technologickým pokrokom sa darí stále znižovať náklady na šírku pásma.

Dátová prevádzka už dlhší čas dominuje v sieťach súčasnosti. V minulosti boli siete navrhnuté efektívne na podporu hlasu viac ako na dátové prenosy. Trend širokopásmovej komunikácie ďalej tlačí infraštruktúru do väčšej kapacity a vyššej rýchlosti.

ITU-T štandardizovala optickú prenosovú sieť s použitím systému WDM a neskôr systému DWDM ako základnú infraštruktúru siete. Siete WDM sú schopné prenášať obrovské množstvo dát vďaka použitiu viacerých rôznych vlnových dĺžok v jednom optickom vlákne. Systémy UHC DWDM (Ultra-High-Capacity DWDM) sa uplatnili v transoceánskych optických sieťach. Napríklad linka TAT-14 bola spustená v roku 2001, meria 15428 km a spája USA s Anglickom a Francúzskom [4,5]. Súčasná kapacita prenosu je 3,2 Tb/s a využíva sa 40 kanálový 10Gb/s DWDM.

Pre dosiahnutie maximálneho potenciálu DWDM systémov je potrebné preskúmať prenosové vlastnosti optických vlákien a optimalizovať parametre sietí. Optické kanály sú ovplyvňované tvarom signálu, ale aj ich lineárnymi a nelineárnymi charakteristikami. Pre multi-kanálové vysokorychlostné systémy sú tieto vplyvy škodlivé. Dôležité je teda pochopenie problematiky, modelovanie a charakterizácia správania sa týchto systémov.

Hlavným cieľom tejto práce je skúmať a analyzovať vplyvy modulačným formátom a nelineárnymi javov v DWDM a UHC DWDM systémoch.

Znáмым problémom optických chrbticových sietí je konverzia signálu medzi elektrickou a optickou doménou. Tento problém značne obmedzuje celkovú rýchlosť systému. Zlepšenie vykazujú plne-optické siete [5].

V prvej časti tejto práce sa budeme venovať optickým zosilňovačom, ktoré sú využívané v optických sieťach. V kapitole 3 rozoberáme návrh 64 kanálového DWDM systému navrhnutého v programe OptiSystem.

II. OPTICKÉ ZOSILŇOVAČE

Optické zosilňovače patria medzi najspoľahlivejšie a najefektívnejšie riešenia kompenzácie útlmu a disperzie, ktoré sú hlavnými limitujúcimi faktormi prenosových vlastností optického vlákna. Nežiadúce dôsledky rozširovania impulzov vplyvom disperzie a poklesu optického výkonu vplyvom útlmu sa zväčšujú s rastúcou dĺžkou optickej trasy do bodu, kedy je potrebné signál regenerovať alebo zosilniť. Okrem použitia zosilňovačov, existujú viaceré zložité metódy ako napríklad metóda 3R (reshaping, retiming, retransmission), pri ktorej dochádza ku konverzii fotón-elektrón, aby sa následne na elektrickom signáli obnovili tvary pulzov a časovanie. Následne dôjde k retransmisii konverziou elektrón-fotón na regenerovaný optický signál [6,7]. Toto riešenie je komplikované a prináša zo sebou viaceré problémy, či už z hľadiska spoľahlivosti, nutnosti prerušenia optickej trasy alebo závislosti na prenosovej rýchlosti a signálovej modulácii.

Moderné optické zosilňovače predstavujú efektívnejšie a ekonomicky prijateľnejšie riešenie, ktoré poskytuje vysoký zisk optického výkonu nezávisle na vlastnostiach vstupného optického signálu. Optické zosilňovače najčastejšie fungujú na rovnakom princípe ako lasery, teda na princípe stimulovanej emisie. Z hľadiska parametrov optického zosilňovača je najdôležitejší zisk pri čerpaní optického výkonu do zosilňovača. Zisk ovplyvňuje mnoho faktorov, no najmä materiál, z ktorého je vyrobené médium v optickom zosilňovači, intenzita svetelného lúča na vstupe a taktiež jeho vlnová dĺžka. V ideálnych podmienkach je možné zisk matematicky popísať rovnicou ako:

$$g(\omega) = \frac{g_0}{1 + (\omega - \omega_0)^2 T_2^2 + \frac{P}{P_s}} \quad (1)$$

Pričom g_0 predstavuje najvyššiu hodnotu zisku, ω označuje veľkosť dopadajúceho signálu, ω_0 predstavuje prechodovú frekvenciu, P je optický výkon signálu, ktorý sa zosilňuje, Parameter T_2 označuje hodnotu relaxačnej doby dipólu a P_s označuje saturačný výkon, ktorý závisí na použítom zosilňovačom médiu konkrétneho optického zosilňovača [8].

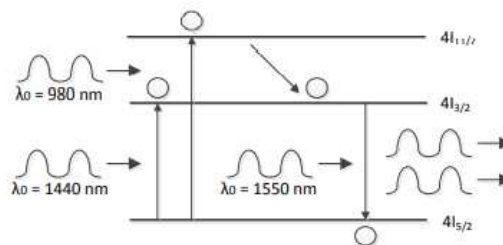
V homogénnych systémoch je úbytok zisku pri $\omega \neq \omega_0$ charakterizovaný na základe Cauchy-Lorentzianovho rozdelenia ako:

$$\Delta\omega_g = \frac{2}{T_2} \quad (2)$$

V praxi sa preferujú zosilňovače s veľkou šírkou pásma, aby bol zisk konštantný pri viackanálových signáloch na celej šírke pásma. Ktorá je daná šírkou polovičného maxima spektra zisku $g(\omega)$. V súčasnosti sa najviac používajú optické vláknové zosilňovače typu (EDFA, Brillouinov zosilňovač, Ramanov zosilňovač) a polovodičové optické zosilňovače (SOA – Semiconductor Optical Amplifier) typu (FOPA – Fiber-Optic Parametric Amplifier, PI FOPA – Polarization Intensive FOPA) [12].

A. Erbium dopovaný vláknový zosilňovač EDFA

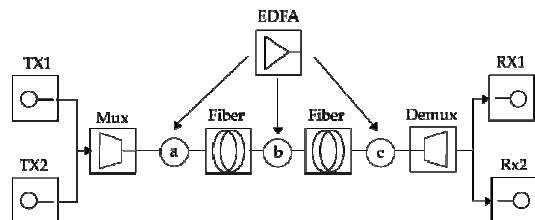
Jedná sa o typ optického zosilňovača s pomerne jednoduchou konštrukciou, ktorý dokáže pracovať v širokom intervale vlnových dĺžok a taktiež konštantne zosilňovať viacero separátnych kanálov v systémoch vlnových multiplexov WDM. Komerčne sa začali používať v roku 1993 v telefónnych systémoch. Vďaka svojej spoľahlivosti a jednoduchosti oproti elektro-optickým riešeniam ide o najpoužívanejší optický zosilňovač, ktorý sa používa najmä pri prenosoch na veľké vzdialenosti, v neprístupnom teréne alebo pri podmorských systémoch [9]. Okrem Erbium sa používajú ako dopanty vo vláknach aj iné vzácne zeminy ako napríklad Neodymium, Ytterbium alebo Thulium, pričom sa odlišujú rôznymi vyžarovacími vlnovými dĺžkami. Samotný princíp fungovania je pomerne jednoduchý a funguje vďaka vláknou, ktoré obsahuje Erbium a technikou inverzie populácie zosilňuje vstupný optický signál. V momente, keď na Erbium dopadne svetelný lúč s vlnovou dĺžkou 980 nm alebo 1480 nm tak zmení svoj stav. V prvej fáze prejde do prechodného, nestabilného stavu a následne v druhej fáze späť do základného stavu s vlnovou dĺžkou 1525-1565 nm. Doba nestabilného stavu je úmerná vlnovej dĺžke svetla, ktoré vstupuje do vlákna dopovaného Erbium. Pri vlnovej dĺžke 980 nm bude mať nestabilný, prechodný stav pomerne krátku dobu trvania, pokiaľ sa navráti do stavu stabilného [10]. Pri vstupnej vlnovej dĺžke 1480 nm bude toto svetlo priamo excitované do stabilného stavu. Pokiaľ bude na vstup prichádzať už existujúci signál, dôjde k jeho zosilneniu.



Obr. 1 Vstupná vlnová dĺžka a jej prechodný a stabilný stav Erbium dopovaný vláknový zosilňovač (EDFA) má niekoľko obmedzení na zabezpečenie zosilnenia s vysokým

ziskom a širokopásmovej prevádzky. Každý zosilňovač vykazuje svoje vlastné výhody, napríklad SOA využíva menší elektrický výkon, menej menších komponentov, zatiaľ čo EDFA je možné obohatiť pomocou rôznych kodopantných materiálov, ako je oxid kremičitý, oxid hlinitý, telurid, fosfát, bizmut a ďalšie, na zlepšenie zisku. Zosilňovače EDFA sú schopné pumpovať v závislosti na rôznych vlnových dĺžkach, od nízkeho zisku, na polarizácii svetla. Väčšina jednotlivých založených zosilňovačov vykazuje nižší zisk v porovnaní s hybridnou konfiguráciou. Kombinácia Ramanovho zosilňovača a EDFA dosahuje to najlepšie z parametrov hybridného zosilňovača. Pred skombinovaním oboch zosilňovačov je potrebné optimalizovať konštrukčné parametre, ako sú dĺžka zosilňovačov, vlnová dĺžka a výkon vstupného čerpadla oboch zosilňovačov. Optimalizáciu týchto parametrov je možné vykonať pomocou simulácie v simulačnom softvéri ako napríklad OptiSystem. Hybridný dvojstupňový zosilňovač Raman-EDFA môže poskytnúť plochý zisk L-pásma. Zvýšenie vstupného výkonu zvyšuje variáciu zisku v celej šírke pásma [10,11].

V rámci DWDM môže byť EDFA umiestnená na troch rôznych miestach optického systému. Používa sa ako výkonový zosilňovač na strane vysielača. Má za úlohu poskytovanie vysokého vstupného výkonu do celého vlákna. Ďalej ako in-line zosilňovač na kompenzáciu straty vlákna pri prenose a ako predzosilňovač na prijímači.

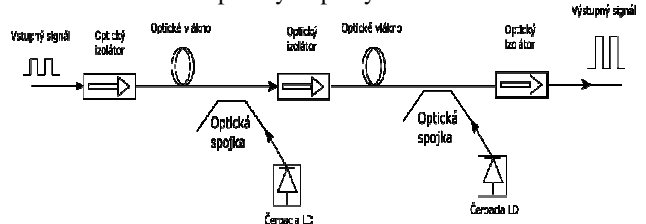


Obr. 2 Schéma zosilňovača EDFA s možnosťami umiestnenia v optickom systéme (a - výkonový zosilňovač, b – in-line zosilňovač, c – predzosilňovač)

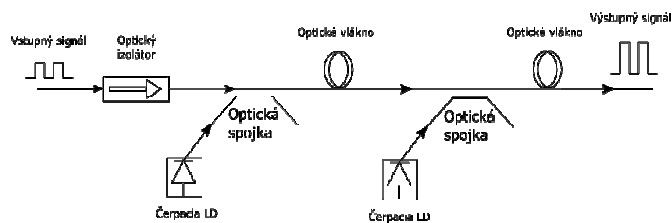
B. Ramanov zosilňovač

Ramanov optický zosilňovač je zariadenie používané v oblasti optických komunikačných systémov na zosilnenie optického signálu. Ide o populárny typ optického zosilňovača, ktorý používa stimulovaný Ramanov rozptyl, ktorého dôsledkom je vznik Ramanovho zisku. Výhodou tejto technológie je schopnosť zosilňovať signál v oboch smeroch, pričom jeho aktívne zosilňovacie médium je optické vlákno alebo krištáľ. V prípade optického vlákna nie je potrebné, aby bolo dopované vzácnymi zeminami, no pre dosiahnutie väčšieho zisku sa zvyknú používať špeciálne optické vlákna, ktoré sú zväčša dopované Germániom. Zosilnenie vzniká emitovaním fotónov z nižšieho stavu do vyššieho stavu vplyvom Ramanovho rozptylu. V Ramanovom rozptyle dochádza k interakcii fotónov s molekulami optického média [13]. Optické médium je potrebné vybudovať zdrojom svetla. ROA optický zosilňovač sa využíva v oblasti 1300 – 1600 nm. Typické zosilnenie pre túto oblasť je 15 dB a použiteľná šírka pásma je 48nm. Nevýhodou Ramanovho zosilňovača je

kolísanie výkonu čerpaceho zdroja, ktorý môže spôsobiť kolísanie amplitúdy a jitteru signálu. Poznáme dva typy ROA zosilňovačov a to dopredný a spätný ROA zosilňovač.



Obr. 3 ROA so spätným čerpacím zdrojom



Obr. 4 ROA s dopredným čerpacím zdrojom

V ROA je využitý tzv. Ramanov rozptylový posun, ktorý vzniká v dôsledku zmien v energii molekúl vlákna pri vzájomnom pôsobení s optickým signálom. Pri prechode optického signálu cez vlákno dochádza k čiastkovému prenosu energie z nosnej frekvencie signálu na iné frekvenčné zložky, čo vedie k zosilneniu signálu. Princíp fungovania ROA spočíva v tom, že do optického vlákna je vložená svetelná pumpa, ktorá poskytuje energiu potrebnú pre Ramanov rozptyl. Optický signál, ktorý chceme zosilniť, je potom vložený do rovnakého vlákna a prechádza cez neho. Pri kontakte s pumpovacím signálom dochádza k Ramanovmu rozptylu a zosilneniu optického signálu. Ramanov optický zosilňovač je široko používaný v optických komunikačných systémoch na zosilnenie signálov na rôznych vlnových dĺžkach. Jeho výhody zahŕňajú vysokú účinnosť zosilnenia, široký rozsah zosilňovanej vlnovej dĺžky a možnosť integrovať ho do rôznych komplexných optických systémov. Použitie Ramanových optických zosilňovačov prispelo k vývoju a rozšíreniu optických komunikačných sietí s vysokou priepustnosťou a spoľahlivosťou, čo umožnilo prenos veľkého množstva dát na veľké vzdialenosti. Hlavná výhoda ROA spočíva v jeho schopnosti zosilniť optický signál na rôznych vlnových dĺžkach. To je možné vďaka Ramanovmu rozptylu, ktorý umožňuje prenos energie z nosnej frekvencie na frekvenčné zložky signálu [14]. Vďaka tomu je ROA schopný pracovať s viacvláknovými systémami a zosilňovať signály v širokom spektre vlnových dĺžok.

ROA má vysokú účinnosť zosilnenia a môže poskytovať vysoký výkon. To je dôležité pre prenos optických signálov na veľké vzdialenosti, kde je potrebné zabezpečiť dostatočnú úroveň signálu. ROA sa často používa na zosilnenie signálov v dlhých optických vláknoch, ktoré majú veľkú stratu signálu. ROA môže byť použitý aj na iné účely okrem zosilňovania signálov v komunikačných systémoch. Napríklad sa využíva v laseroch na zosilnenie výstupného signálu. Tiež môže byť

súčasťou optických senzorov, ktoré vyžadujú citlivé a presné meranie optických parametrov [1].

Súhrn vlastností ROA:

- Široký rozsah zosilnených vlnových dĺžok: ROA je schopný zosilňovať optické signály v širokom spektre vlnových dĺžok. To znamená, že môže pracovať s rôznymi typmi optických signálov a rôznymi vláknovými systémami.
- Vysoká účinnosť zosilnenia: ROA poskytuje vysokú účinnosť zosilnenia optického signálu. To znamená, že dosiahne značné zvýšenie výkonu signálu pri relatívne malom množstve vlozenej energie.
- Vysoký výstupný výkon: ROA je schopný poskytovať vysoký výstupný výkon zosilneného signálu. Tento výstupný výkon je dostatočne silný na prenos optických signálov na veľké vzdialenosti bez významnej straty signálu.
- Nízky šum: ROA má nízky prídavný šum, čo je dôležité pre udržanie vysokého pomeru signálu ku šumu (SNR). Nízky šum zabezpečuje kvalitný a spoľahlivý prenos signálov.
- Možnosť integrovať sa do optických systémov: ROA môže byť ľahko integrovaný do rôznych optických systémov a sietí. Môže byť implementovaný ako samostatné zariadenie alebo ako súčasť optických prvkov, ako sú optické prepínače alebo multiplexory.

C. Polovodičový zosilňovač

Optické zosilňovače SOA (Semiconductor Optical Amplifier) používajú polovodič ako zosilňovacie médium, ktoré je navrhnuté na použitie vo všeobecných aplikáciách na zvýšenie výkonu optického výkonu a na kompenzáciu straty iných optických zariadení [11]. Polovodičové optické zosilňovače sa v telekomunikačných systémoch často používajú vo forme vláknových komponentov, ktoré pracujú pri vlnových dĺžkach od 0,85 μm do 1,6 μm a generujú zisky až do 30 dB. Polovodičový optický zosilňovač, dostupný vo vlnovej dĺžke 1310 nm, 1400 nm, 1500 nm, 1600 nm je možné použiť s jednovidovým vláknom alebo v polarizáciu zachovávajúcim vláknom. Šum SOA je väčšinou vyšší ako u EDFA. Z hľadiska výkonu sú SOA horšie ako EDFA a z hľadiska šírky pásma porovnateľné. Výhodou SOA sú rozmery zosilňovača. Všetky parametre, ktoré boli uvedené sú zhrnuté v TABLE I.

SOA zosilňovače zosilňujú dopadajúce svetlo prostredníctvom stimulovanej emisie. Stimulované fotóny majú rovnakú vlnovú dĺžku ako optický signál, čím signál zosilňujú. SOA môže fungovať v dvoch rôznych režimoch. V prvom prípade, ktorý sa nazýva režim Tavel-Wave (TW), je zabránené osciláciám, aby sa dosiahol zisk jedného prechodu.

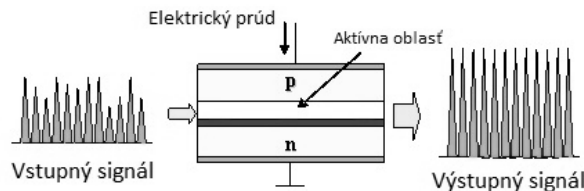
TABLE I Všeobecné parametre zosilňovačov

Vlastnosti	EDFA	ROA	SOA
Zisk (dB)	>40	>25	>30
Vlnové dĺžky (nm)	1530-1560	1280-1650	1280-1650
Šírka pásma (3dB)	30-60	Záleží na zdroji	60
Saturácia (dBm)	22	0.75x zdroj	18
Polarizačná senzitivita	Nie	Nie	áno
Šum (dB)	5	5	8
Výkon zdroja (dBm)	25	>30	< 26

Aktívne médium je opísané ako homogénne rozšírený dvojúrovňový systém. V takom prípade má koeficient zisku g(ω) tvar (3)

$$g(\omega) = \frac{g_0}{1 + (\omega - \omega_0)^2 T_2^2 + P/P_s} \tag{3}$$

Kde g₀ je špičková hodnota zosilnenia, ω₀ je prechodová frekvencia, P je optický výkon zosilneného signálu, P_s je saturačný výkon a T₂ ≤ 1ps je doba relaxácie dipólu.



Obr. 5 Princíp fungovania SOA

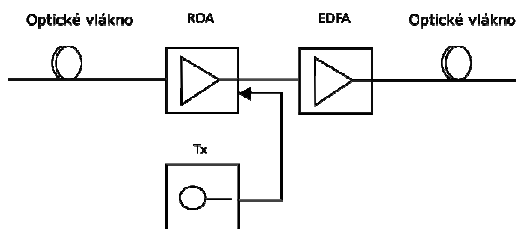
Parametrické zosilňovače (FOPA-Fiber Optical Parametric Amplifier) sa ukázali ako sľubná zosilňovacia technológia s vlastnosťami, ako je veľmi veľký zisk (70 dB) široké (>100 nm) a ploché spektrum zisku a fázovo citlivé zosilnenie umožňujúce šumové číslo <3 dB. FOPA sa stala dostupnou pre in-line amplifikáciu zavedením nových architektur FOPA na ktoré nemá vplyv polarizácia. FOPA sa spolieha na Kerrov efekt a umožňuje prakticky okamžitú dobu odozvy 0,1 fs. Schopnosť FOPA pracovať v ľubovoľných pásmach vlnových dĺžok ho robí použiteľným aj pre pasívne optické siete (PON) využívajúce pásma O, C a L. Toto otvára možnosti pre FOPA ako in-line zosilňovače pre rozšírenú pasívnu optickú sieť (PON).

Polarizáciu zanedbávajúce FOPA (PI-FOPA) zosilňujú kanály WDM s opačným šírením v obojsmernom prenose signálu v pásme C a L. Väčšina aplikácií v optických komunikáciách vyžaduje, aby zosilňovače mali zosilnenie zanedbávajúce polarizáciu. PI-FOPA boli obmedzené na šírku pásma <20 nm kvôli zložitosti optických schém [15,16].

D. Hybridná konfigurácia EDFA a Ramanovho zosilňovača

Vzhľadom na rastúci dopyt po vysokej prenosovej kapacite a systéme s najlepším výkonom komunikačného systému s optickými vláknami sa hybridné zosilňovače stávajú najlepšou možnosťou na vyriešenie problému. Polovodičový optický

zosilňovač (SOA), vláknový zosilňovač dotovaný erbiom (EDFA) alebo Ramanov zosilňovač, má niekoľko obmedzení na zabezpečenie zosilnenia s vysokým ziskom a širokopásmovej prevádzky. Niektoré nedostatky je možné vyriešiť nasadením hybridného systému. Kombináciou Ramanovho a EDFA zosilňovača získame nízkošumový a širokopásmový signál [13].



Obr. 6 Schéma hybridného zosilňovača

III. NÁVRH 64 KANÁLOVÉHO DWDM SYSTÉMU V PROGRAME OPTISYSTEM S RÔZNYMI ZOSILŇOVAČMI

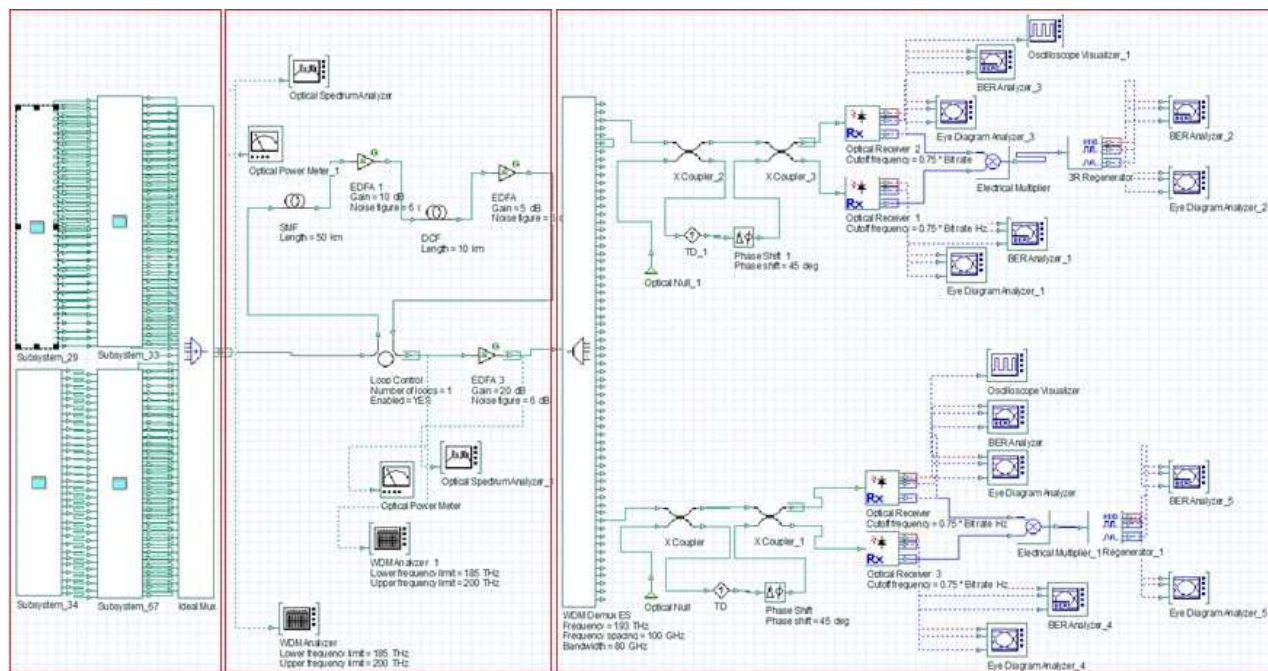
Simulovanie optických komutačných sietí je efektívny nástroj pre návrh ale aj implementáciu optických systémov. Simulácie značne skracujú čas potrebný na uvedenie sietí na trh. Simuláciou je možné overiť funkčnosť a efektivitu navrhnutých scenárov optickej siete. Program OptiSystem od spoločnosti Optiwave je vhodný pre simuláciu optických komunikačných systémov [18].

Navrhnutú schému môžeme rozdeliť na tri hlavné časti. Vysielaciu časť, prenosovú časť a prijímaciu časť. Vo vysielacej časti sú obsiahnuté laserové polia zložené z CW (Continuous Wave) laserov. Lasery sú nastavené od 193,1 THz s kanálovými rozstupmi 100 GHz. Kvôli počtu kanálov sme laserové polia vložili do subsystemu. Ďalším prvkom je subsystem modulácie. V každom návrhu bolo možné subsystem modulácie vymeniť a simulovať následné zmeny. Jednotlivé modulačné subsystemy sú zobrazené v ďalších kapitolách. Nasledujúcim prvkom optického reťazca je DWDM multiplex, ktorý má za úlohu spájať kanály do jedného jednovidového optického vlákna.

Na Obr. 7 môžeme vidieť celú schému simulovanú v programe OptiSystem.

Druhou hlavnou časťou je prenosová časť, ktorá obsahuje optické vlákno, v našom prípade SMF (jednovidové optické vlákno), zosilňovače podľa daného návrhu a DCF vlákno (vlákno s potlačením disperzie).

Treťou hlavnou časťou je prijímacia strana. Obsahuje DWDM demultiplexor, ktorý má za úlohu z jedného optického vlákna rozdeliť jednotlivé kanály. Podľa modulácie sú do schémy zaradené demodulátory a posledným prvkom je optický prijímač. Ďalšie prvky na prijímacej strane slúžia ako vizualizéry. Je použitý BER analyzer, ktorý vykresľuje aj diagram pre Q faktor. Analyzer diagramu oka vykresľuje graf diagramu oka a hodnotu výška oka.



Obr. 7 schéma 64x100 Gb/s DWDM v programe OptiSystem

V niektorých prípadoch bolo nutné využiť 3R regenerátor signálu. Architektúra a princíp 3R regenerátora s vylepšením

BER pozostáva z optického zdroja hodín, erbiom dopovaného vláknového zosilňovača, vysoko nelineárneho vlákna HNLF

(Highly Nonlinear Fiber) a pásmového filtra BPF (Band-Pass Filter). Hodiny vytvárajú impulzy kratšieho trvania ako impulzy na regeneráciu. Frekvencia opakovania hodín je zvolená tak, aby sa rovnala rýchlosti prenosu signálu.

A. Optimalizácia zosilňovačov

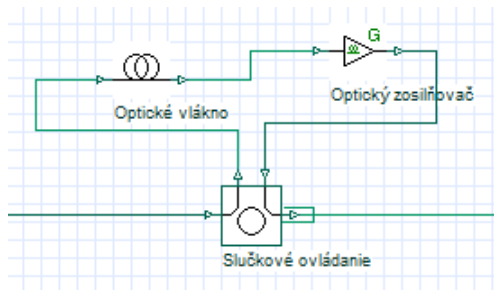
Poznáme rôzne typy optických zosilňovačov, ktoré ovplyvňujú kvalitu signálu v DWDM prenosovom systéme. Správny typ optického zosilňovača so správnym nastavením parametrov výrazne prispieje ku zlepšeniu prenosových vlastností siete. Všeobecné parametre navrhnutého systému sú zobrazené v Tab. 2.

TABLE II Všeobecné parametre

Zdroj Svetla	CW laser
Centrálna frekvencia	193,1 THz
Kanálové rozostupy	100 GHz
Typ modulácie	-3, 0, 3 dBm
Bitová rýchlosť	40 Gb/s
Disperzia	16,75 ps/nm/km
Zosilňovač	EDFA, ROA, SOA

1) Erbium dopovaný optický zosilňovač EDFA

64 kanálový DWDM systém pri prenosovej rýchlosti 40 Gb/s sme testovali so zmenami výkonu laserov vo vysielačej časti systému a to pre hodnoty -3 dBm, 0 dBm a 3 dBm pri jednoduchej modulácii NRZ na vzdialenosti od 70 do 210 km. Ak sme zväčšili dĺžku optického vlákna nad 210 km, a nezmenili iné vstupné parametre, systém nebol ďalej schopný generovať relevantné výsledky pre všetky kanály. Modulačný formát bol vybraný kvôli úzkemu optickému spektru čo spôsobuje minimalizovanie nežiadúcich javov pri prenose. V programe OptiSystem sme použili prvok „optický zosilňovač“, ktorý sme parametrovo upravili aby spĺňal podmienky EDFA zosilňovača zapojeného in-line. Zisk EDFA sme nastavili na hodnotu 20 dB.



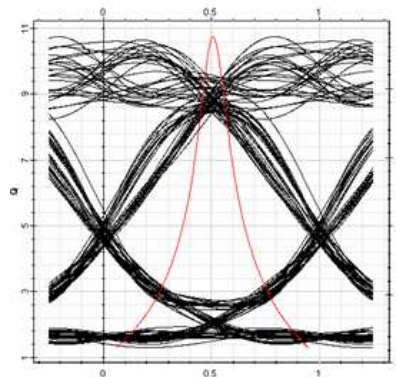
Obr. 8 optický zosilňovač

Kvôli časovej náročnosti výpočtov pre program OptiSystem a hardware sme simulovali iba prvý kanál a 59. Kanál v našom 64 kanálovom DWDM systéme pri prenosovej rýchlosti 40 Gb/s. Tieto kanály sme zvolili, kvôli dostatočnému kanálovému rozostupu a väčšej diverzifikácii výsledkov.

TABLE III Výsledky prvého kanála 40 Gb/s

Výkon lasera	Parametre	Výsledky
-3 dBm	Q faktor	11,01756
	Min BER	$2,13453 \times 10^{-24}$
	Výška oka	0,018254
0 dBm	Q faktor	11,5392
	Min BER	$4,18331 \times 10^{-31}$
	Výška oka	0,0246
-3 dBm	Q faktor	12,3977
	Min BER	$1,34412 \times 10^{-35}$
	Výška oka	0,02864

Z výsledkov pre prvý kanál môžeme pozorovať, že so zvyšujúcim sa výkonom lasera sa zvyšovali aj parametre Q faktor a min. BER. Najlepšie parametre na vzdialenosť 210 km systém dosiahol pre výkon lasera 3 dBm a to Q faktor 12,3977 a min BER $1,34412 \times 10^{-35}$.

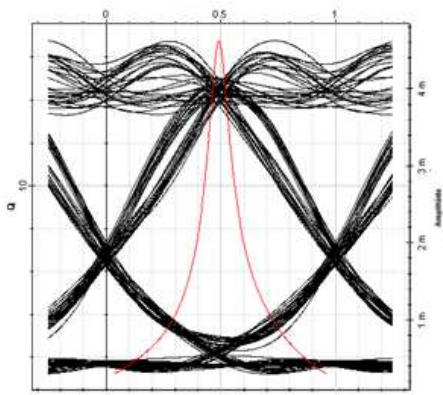


Obr. 9 diagram oka pre prvý kanál

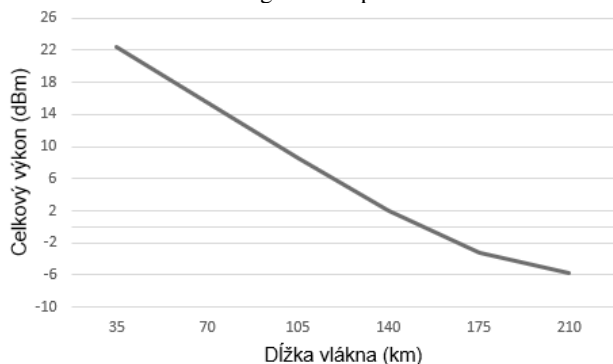
Do systému boli zaradené aj merače výkonu, ktoré sa umiestnili pred a za zosilňovač. Bolo zistené, že pri výkone lasera 3 dB a prenosovej dĺžke 70 km výkon za zosilňovačom dosahoval hodnotu 16,78 dB a pri dĺžke 140 km už iba hodnotu 2,72 dB.

TABLE IV výsledky 59. kanála 40 Gb/s

Výkon lasera	Parametre	Výsledky
-3 dBm	Q faktor	9,0752
	Min BER	$2,92 \times 10^{-19}$
	Výška oka	0,01498
0 dBm	Q faktor	10,5012
	Min BER	$4,18 \times 10^{-22}$
	Výška oka	0,01698
-3 dBm	Q faktor	11,0357
	Min BER	$1,34 \times 10^{-26}$
	Výška oka	0,01928



Obr. 10 diagram oka pre 59. kanál



Obr. 11 Závislosť výkonu od dĺžky vlákna

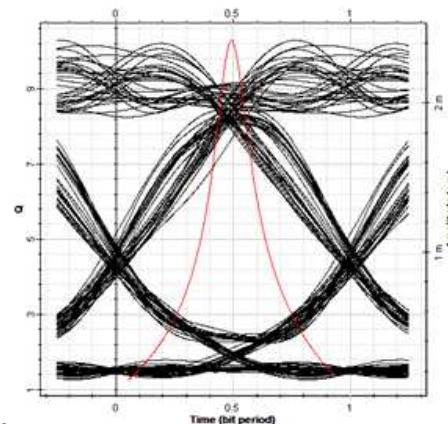
V nasledujúcej simulácii sa vykonala zmena prenosovej rýchlosti zo 40 Gb/s na 100 Gb/s. Výsledky dosiahli nižšie hodnoty ako pre 40 Gb/s DWDM systém. Systém sa rovnako simuloval pre 3 výkony lasera (-3 dBm, 0 dBm, 3dBm). Ostatné vstupné parametre ostali nezmenené.

TABLE V Výsledky prvého kanála 100 Gb/s

Výkon lasera	Parametre	Výsledky
-3 dBm	Q faktor	10,0547
	Min BER	$4,787 \times 10^{-20}$
	Výška oka	0,01651
0 dBm	Q faktor	10,7162
	Min BER	$4,994 \times 10^{-20}$
	Výška oka	0,01686
-3 dBm	Q faktor	11,3427
	Min BER	$1,145 \times 10^{-24}$
	Výška oka	0,01707

S nárastom prenosovej rýchlosti sa môže vyžadovať zvýšenie výkonu signálu, aby sa udržal prijateľný pomer OSNR. Z výsledkov môžeme pozorovať, že so zvýšením

výkonu sa min BER zlepšilo na vzdialenosť 210 km. Útlm vlákna však narastá s výkonom signálu, čo spôsobuje degradáciu signálu na dlhšie vzdialenosti, preto treba nájsť správny pomer pri zvyšovaní výkonu a vzdialenosti.



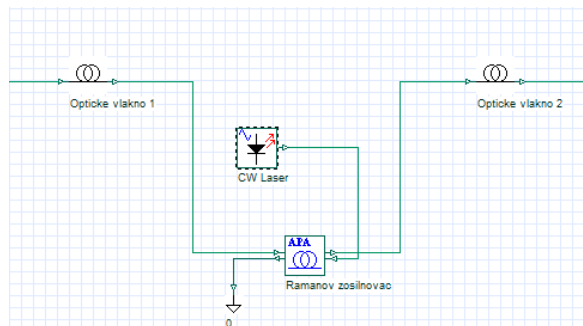
Obr. 12 diagram oka prvý kanál

TABLE VI Výsledky 59. Kanála 100 Gb/s

Výkon lasera	Parametre	Výsledky
-3 dBm	Q faktor	9,01682
	Min BER	$1,574 \times 10^{-19}$
	Výška oka	0,01426
0 dBm	Q faktor	9,8052
	Min BER	$6,152 \times 10^{-19}$
	Výška oka	0,01435
-3 dBm	Q faktor	10,0135
	Min BER	$1,228 \times 10^{-20}$
	Výška oka	0,01657

2) Ramanov optický zosilňovač ROA

Výhodou Ramanovho zosilňovača je fakt, že zosilnenie prebieha priamo vo vlákne určenom na prenos signálu. Pre zosilnení signálu sa využíva nelineárny jav nazývaný stimulovaný Ramanov rozptyl. ROA sa obdobne testoval pri rôznych výkonoch lasera v 64 kanálovom DWDM systéme. Prenosová rýchlosť 40 Gb/s a výkony lasera -3 dBm, 0 dBm a 3 dBm pri jednoduchej modulácii NRZ na vzdialenosti od 70 do 210 km. Kde sa vzdialenosť 210 ukázala ako hraničná pre navrhnutý systém. Dlhšie vzdialenosti CW laser napájajúci ROA bol nastavený na vlnovú dĺžku 1450 nm v ktorej Ramanov zosilňovač pracuje najlepšie a výkon pumpy 10 dBm.

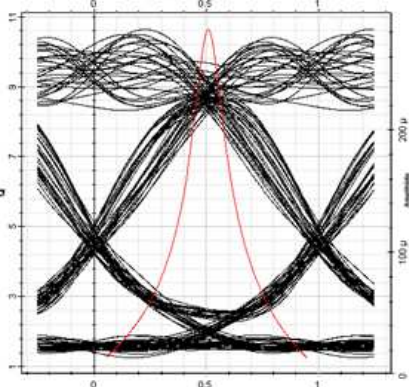


Obr. 13 Ramanov optický zosilňovač v programe OptiSystem

TABLE VII Výsledky prvého kanála pre 64 kanálový DWDM systém s ROA na vzdialenosť 210 km

Výkon lasera	Parametre	Výsledky
-3 dBm	Q faktor	9,01682
	Min BER	$1,574 \times 10^{-19}$
	Výška oka	0,01426
0 dBm	Q faktor	9,8052
	Min BER	$6,152 \times 10^{-19}$
	Výška oka	0,01435
-3 dBm	Q faktor	10,0135
	Min BER	$1,228 \times 10^{-20}$
	Výška oka	0,01657

Z výsledkov je možné pozorovať, že opäť so zvyšujúcim sa výkonom lasera stúpa Q faktor. Výsledky pre Ramanov zosilňovač sú horšie ako pre EDFA zosilňovač. Na Obr. 83 a) je diagram oka pre prvý kanál 64 DWDM s ROA a na Obr. 83 b) je 59. kanál 64 kanálového DWDM systému s ROA.



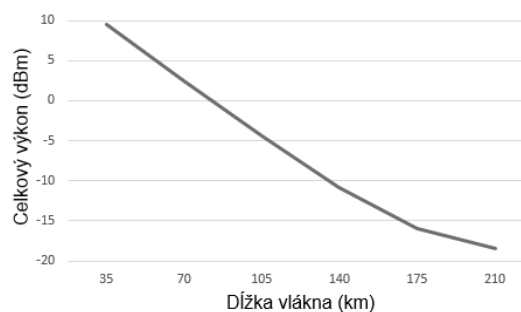
Obr. 14 diagram oka pre prvý kanál

TABLE VIII Výsledky 59. kanála DWDM systému s ROA

Výkon lasera	Parametre	Výsledky
-3 dBm	Q faktor	9,01682
	Min BER	$1,574 \times 10^{-19}$

Výkon lasera	Parametre	Výsledky
-3 dBm	Q faktor	9,01682
	Výška oka	0,01426
0 dBm	Q faktor	9,8052
	Min BER	$6,152 \times 10^{-19}$
	Výška oka	0,01435
-3 dBm	Q faktor	10,0135
	Min BER	$1,228 \times 10^{-20}$
	Výška oka	0,01657

Z výsledkov ROA môžeme usúdiť, že pri našich rovnakých vstupných podmienkach je zosilňovač ROA menej účinný ako EDFA čo je aj dôvodom prečo sa pri optimalizácii výsledného 64 kanálového DWDM systému pracovalo so zosilňovačmi EDFA.



Obr. 15 Závislosť celkového výkonu od dĺžky vlákna pre ROA zosilňovač

ZÁVER

Cieľom tejto práce bolo podľa najnovšej dostupnej literatúry navrhnuť a odsimulovať optické zosilňovače EDFA, ROA v DWDM systéme optickej siete v simulačnom nástroji OptiSystem skúmaním ich vlastností, výhod, nevýhod a ich vzájomným porovnávaním.

Analyzovali sme zosilňovače pre 64 kanálový DWDM pre dĺžku vlákna 210 km a výkon lasera -3 dBm do 3 dBm. Z výsledkov vyplýva, že systém EDFA DWDM dosahuje lepšie hodnoty BER ako ROA DWDM. Pri porovnaní optických zosilňovačov z hľadiska Q faktoru v závislosti na hodnote výkonu lasera sa javil ako najlepší EDFA zosilňovač. Schéma návrhu optickej siete so zapojením ROA zosilňovača sa vyznačuje širokým pásmom vlnových dĺžok oproti návrhu optickej siete so zapojením EDFA zosilňovača. Nevýhodou tejto siete je, že v porovnaní s návrhom optických sietí so zapojením EDFA zosilňovačov, dosahuje najnižší zisk výkonu. Optický zosilňovač EDFA je vhodný výber pre tento typ navrhnutých sietí. EDFA je veľmi účinný a stále viac a viac dostupný zosilňovač s možnosťou širokého využitia. Najlepší min BER pre 64 kanálový DWDM systém s EDFA pre prvý kanál vyšiel $1,34412 \times 10^{-35}$.

REFERENCIE

- [1] Agrawal, Govind. *Optical Communication: Its History and Recent Progress*. 10.1007/978-3-319-31903-2_8, (2016)
- [2] John M. Senior. *Optical Fiber Communication Senior - 3rd Edition*. ISBN: 978-0-13-032681, (2009)
- [3] Yi Zhang, Chowdhury, P, Tornatore, M and Mukherjee, B (2010), Energy Efficiency in Telecom Optical Networks. Retrieved from IEEE Communications, Tutorials and Surveys, Vol 12. No 4: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5522464
- [4] Hui R. *Introduction to Fiber-Optic Communication*, Academic Press, ISBN:978-0-128092-345-4, (2019)
- [5] Agrawal, G. P. *Nonlinear Fiber Optics*, 6th edition, Academic Press, San Diego, CA, (2019), ISBN: 9780128170427
- [6] Antoniadis, N. – Ellinos, G. – Roudas, I. *WDM Systems and Networks, Modeling, Simulation, Design and Engineering*, 2012, 549 p. ISSN: 1935-3839, DOI 10.1007/978-1-4614-1093-5.
- [7] Mohammed, Abd El-Naser A.; Rashed, Ahmed Nabih Zaki; Saad, Abd El-Fattah A. *Applications of Arrayed Waveguide Grating (AWG) in Passive Optical Networks*. 2009-5-02, [cit. 2011-11-11], s. 1-12.
- [8] Anurupa, Kaur, S., & Malhotra, Y. , Performance evaluation and comparative study of novel high and flat gain C+L band Raman+EYDFA co-doped fibre hybrid optical amplifier with EYDFA only amplifier for 100 channels SD-WDM systems. *Optical Fiber Technology*, 53, 2019, doi:10.1016/j.yofte.2019.102016
- [9] Singh, M. "A Review on Hybrid Optical Amplifiers. *Journal of Optical Communications*," 39(3), 2018, 267–272. doi:10.1515/joc-2016-0137
- [10] Huszanik, T., Turan, J., & Ovsenik, L. (2020). Optimization of Optical Amplification in the High Capacity DWDM System. 2020 21th International Carpathian Control Conference (ICCC). doi:10.1109/iccc49264.2020.9257245
- [11] S. J. Anderson Comparative Study of EDFA and SOA in WDM Optical Network Placed at Different Position, In: *International Journal of Photonics and Optical Technology*, Vol. 1, no. 1, 2015, p.4-6.
- [12] Zhu, T., & Menyuk, C. R. (2019). Fundamental limits of amplifiers in coherent systems. *Journal of Lightwave Technology*, 37(17), 4243-4253.
- [13] M. M. J. Martini, C. E. S. Castellani, M. J. Pontes, M. R. N. Ribeiro and H. J. Kalinowski, "Performance comparison for Raman+EDFA and EDFA+Raman hybrid amplifiers using recycled multiple pump lasers for WDM systems," 2015 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC), Porto de Galinhas, 2015, pp. 1-5. doi: 10.1109/IMOC.2015.7369103
- [14] Kuanlin Mu, Zhuoze Zhao, Zhengkang Wang, Jianmin Shang, Song Yu, and Yaojun Qiao, "Raman/EDFA hybrid bidirectional amplifier for fiber-optic time and frequency synchronization," *Opt. Express* 29, (2021)
- [15] J. Purina, S. Olonkins, V. Borovs and J. Porins, "Investigation of EDFA Performance in DWDM Transmission Systems Under Different Operating Conditions," *Latvian Journal of Physics and Technical Sciences*, vol.56, no.2, 2019, pp.42-53. <https://doi.org/10.2478/lpts-2019-0012>
- [16] Yu Yu, Xinliang Zhang, J. B. Rosas-Fernández, Dexiu Huang, R. V. Penty, and I. H. White, "Single SOA based 16 DWDM channels all-optical NRZ-to-RZ format conversions with different duty cycles," *Opt. Express* 16, 16166-16171 (2008)
- [17] T. K. Yaratha, N. Reddy and G. A, "WDM Optical Network Analysis using EDFA and RAMAN Amplifier," *Indian Journal of Applied Research*, vol. 5, no. 6, pp. 346-349, 2015.
- [18] Marzieh S. L.; Faramarz E.S. Design of a 32x5 Gb/s DWDM Optical Network over a Distance of 1000 km, DOI: 10.5923/j.optics.20160602.02, (2016) Dostupné na internete: <https://www.researchgate.net/publication/307422847_Design_of_a_32_5_Gbs_DWDM_Optical_Network_over_a_Distance_of_1000_km>

QuoVadis Research @ FEI

Vydavateľ: Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky

Sídlo: Letná 9, 042 00 Košice

Elektronická adresa: <http://quovadis.fei.tuke.sk>

Periodicita: 2-krát ročne

Založenie časopisu: September 2018.

Jazyk: V časopise sú články v anglickom aj slovenskom jazyku.

ISSN: 2585-9587