



**STUDENT  
EEICT**

**BRNO UNIVERSITY OF TECHNOLOGY  
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION**

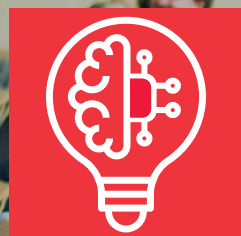
**PROCEEDINGS I OF THE 28<sup>TH</sup> STUDENT EEICT 2022**  
ELECTRICAL ENGINEERING, INFORMATION SCIENCE,  
AND COMMUNICATION TECHNOLOGIES  
**GENERAL PAPERS**

**April 26, 2022  
Brno, Czech Republic**



**BRNO FACULTY OF ELECTRICAL  
UNIVERSITY ENGINEERING  
OF TECHNOLOGY AND COMMUNICATION**





# STUDENT EEICT

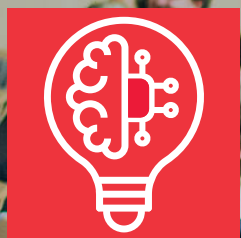
<b>Title</b>	Proceedings I of the 28 <sup>th</sup> Conference STUDENT EEICT 2022
<b>Supervisor</b>	Prof. Vladimír Aubrecht
<b>Editor</b>	Assoc. Prof. Vítězslav Novák
<b>Publisher</b>	Brno University of Technology, Faculty of Electrical Engineering and Communication
<b>Year released</b>	2022 1 <sup>st</sup> Edition

Content and language issues are the responsibility of the authors.  
ISBN 978-80-214-6029-4



BRNO FACULTY OF ELECTRICAL  
UNIVERSITY ENGINEERING  
OF TECHNOLOGY AND COMMUNICATION





**STUDENT  
EEICT**

## **SPONSORS**

**Diamond tier**

# **Honeywell**

**Golden tier**

**Garrett**  
ADVANCING MOTION

**onsemi**

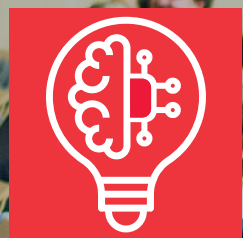
**ThermoFisher**  
SCIENTIFIC

**Valeo**



BRNO FACULTY OF ELECTRICAL  
UNIVERSITY ENGINEERING  
OF TECHNOLOGY AND COMMUNICATION





# STUDENT EEICT

## SPONSORS

Silver tier



**SAAB**



Honored tier



**ŠKODA**



## TECHNICAL SPONSORS & MEDIA PARTNERS



BRNO FACULTY OF ELECTRICAL  
UNIVERSITY ENGINEERING  
OF TECHNOLOGY AND COMMUNICATION



# Contents

## Foreword

## Secondary school papers

<b>V. Mautner, D. Olenišák, and M. Jurák</b>	
Robotic arm control using MQTT protocol.....	13
<b>V. Mautner, D. Olenišák, and M. Jurák</b>	
A Line Following car using OpenMV camera .....	17
<b>V. Glombíček</b>	
Scientific Calculator .....	21
<b>S. Matoušek</b>	
Wirelessly Controlled Robotic Swarm .....	25



## Bachelor papers

### B1 – Power Electronics and Energy

<b>Š. Chmelař</b>	
Potential production and development of photovoltaic panels .....	29
<b>M. Martinček</b>	
Spot Welding Machine for Li-ion Battery Cells.....	33
<b>P. Krejčířík</b>	
Community energy for the municipality with population 3 000.....	37

### B2 – Systems Modeling, Simulation, Control and Measurement

<b>J. Čelko and O. Mihálik</b>	
McRuer Models for Human-Machine Systems .....	43
<b>S. Novík and O. Mihálik</b>	
Analysis of Car Drivers' Behaviour and Driving Style .....	47
<b>O. Skalský, S. Pikula</b>	
System for verification of the signal quality dependence of a laser vibrometer .....	51
<b>A. Pala, P. Vyroubal</b>	
IoT Weather Station .....	55
<b>M. Černík and O. Mihálik</b>	
Heat Exchanger Regulation .....	59
<b>T. Kříčka, L. Langhammer</b>	
Design of 3rd-Order Electronically Reconfigurable Filter by Inverse Chebyshev Approximation.....	63

### B3 – Microelectronics and Technology

<b>P. Prášil, P. Petyovský</b>	
Universal asynchronous receiver/transmitter implementation in VHDL .....	68
<b>A. Jonáš</b>	
Electrochemical etching of graphite tips for potential use in nanoscience .....	72
<b>J. Drška</b>	
Automated firesport stopwatch.....	76
<b>T. Homolka, J. Libich</b>	
Sodium-ion batteries: research of carbon-based negative electrode materials .....	80
<b>J. Jarušek, I. Gablech</b>	
Transparent and conductive titanium nitride thin films for implantable bioelectronics deposited at low temperature using two Kaufman ion-beam sources.....	84

### B4 – Industrial Automation, Robotics and Artificial Intelligence I.

<b>J. Boch, M. Husák</b>	
Batch process in connection with industry 4.0 .....	89
<b>J. Nosek, and J. Arm</b>	
Testbed for virtual commissioning of PLC system .....	94
<b>V. Houdek, O. Baštán</b>	
Asset Administration Shell for the operator .....	98
<b>P. Černocký, and J. Arm</b>	
Demonstrative Manufacturing System Controlled by MES Utilizing AAS .....	102
<b>M. Mařas, T. Sýkora</b>	
Ice cube feeder.....	106

### B5 – Industrial Automation, Robotics and Artificial Intelligence II.

<b>R. Čambál, T. Sýkora</b>	
Automatic check for the certification label .....	112
<b>B. Zezula, P. Marcoň</b>	
Acoustic detection of starlings through artificial intelligence .....	116

## B6 – Bioengineering and Bioinformatics

<b>V. Plocková and K. Sedlář</b>	
Properties of Current Signals in Nanopore Sequencing .....	120
<b>T. Šilerová</b>	
Analyzing parameters of drawing .....	124
<b>A. Kostova, J. Schwarzerova</b>	
Concept Drift Detection in Prediction Classifiers for Determining Gender in Metabolomics Analysis .....	128

## B7 – Communication Systems and Network Security

<b>M. Mitěv and L. Pohl</b>	
Kernel latency analysis on Raspberry Pi OS .....	133
<b>M. Rickwood, and V. Oujezský</b>	
A proposal of a method to detect spam from information messages .....	137
<b>J. Rajm, J. Janoušek</b>	
Communication Network for a Swarm of Unmanned Aerial Vehicles .....	141
<b>M. Sečkář, and S. Ricci</b>	
Secure Two-Party Computation for weak Boneh-Boyen Signature .....	145
<b>V. Janos, R. Juran</b>	
Measuring Station Firmware for LoRaWAN Wireless Data Collecting .....	149



## Master papers

### M1 - Power Electronics and Energy

#### T. Řehořek

Role of battery capacity in charging habits of battery electric vehicle users ..... 154

#### M. Los, K. Jaššo

Study of the ecological impact of production, operation and disposal of various types of cars ..... 158

### M2 – Industrial Automation, Robotics and Artificial Intelligence

#### S. Svědřih, L. Žalud

Mobile Robot with Telepresence Control ..... 162

#### D. Ríčanek, P. Honec

Eulerian Video Magnification in Surveillance Systems ..... 166

#### J. Bittner, and T. Beneš

Advanced SCARA robot control system ..... 170

#### M. Koupil and J. Odstrčilík

Biometrics using face recognition ..... 175

#### M. Pončák and P. Petyovský

Innovation of the laboratory exercises for course Embedded Systems ..... 179

### M3 – Microelectronics and Technology

#### M. Šimek

Analysis of critical battery parameters for space applications ..... 185

#### V. Král

Saving area and power consumption in 65 nm digital standard cell library ..... 190

#### A. Šimek

Negative electrode for sodium-ion batteries ..... 194

#### A. Györi, D. Smékal

Post Quantum Cryptography on FPGA ..... 198

#### N.N. Pham, J. Leuchter, and H.Q. Dong

Wireless Battery Management System for Military Unmanned Vehicles ..... 202

### M4 – Systems Modeling, Simulation, Control and Measurement

#### O. Kolář, M. Kubiček

Proportional counter measurement system ..... 208

#### D. Piskoř, O. Mihálik, T. Sýkora, and J. Arm

A Fuzzy Controller for a Wind Tunnel Model ..... 212

#### Z. Vyležich, and M. Kyselák

Using wavelength multiplex and optical circulators for power-supply temperature fiber-optic polarization sensor ..... 216

#### N.N. Pham, J. Leuchter, and J. Václavík

Test Bench for Experimental Analysis of Conductive Composite Materials for Aerospace Applications ..... 220

#### M. Šimek, J. Arm

Horticulture IoT module ..... 225

#### O. Kašpárek, J. Fialka

Measurement System for Determining Piezoelectric Charge Coefficient ..... 229

#### J. Mička, J. Láčík, M. Martinec

A patch antenna array for IFF band ..... 233

### M5 – Bioengineering and Bioinformatics

#### M. Nohel, R. Kolář

Detection of optic disc in retinal fundus image ..... 239

#### K. Novotný, J. Mekyska

Assessing Movement of Articulatory Organs in Patients with Parkinson's Disease ..... 243

<b>J. Vývoda, R. Jakubíček</b>	
Comparison of full-size and patches-based learning approaches for aneurysm segmentation in TOF-MRI data.....	247
<b>P. Nemčerková and J. Schwarzerová</b>	
Dynamic metabolomic prediction based on genetic variation for <i>Hordeum vulgare</i> .....	251
<b>K. Širůčková, P. Solár, P. Marcoň</b>	
Analysis of brain tumors based on line of interest.....	255
<b>M6 – Communication Systems and Signal Processing</b>	
<b>J. Michálek, and V. Oujezský</b>	
An app to demonstrate the risks of granting permissions in Android OS.....	261
<b>O. Zelený and T. Frýza</b>	
Traffic Analysis Using Machine Learning Approach.....	265
<b>V. Barta, L. Polak, P. Cidl</b>	
Planning of Single Frequency Networks for DVB-T2 .....	269
<b>M. Kovařík and T. Caha</b>	
An Anti-spam Software Plugin For E-Mail Classification Based On Geographical Location Of The Sender .....	273
<b>L. Malček, P. Musil and P. Mlýnek</b>	
Evaluation of broadband couplers for noise injection into powerlines.....	277
<b>M7 – Network Services and Security</b>	
<b>A. Krivulčík and S. Ricci</b>	
An Implementation of Lattice-based Proof-of-Work on Blockchain .....	283
<b>P. Ryšavá and S. Ricci</b>	
Paillier Cryptosystem Optimisations for Homomorphic Computation.....	287
<b>F. Kamenář and Z. Martinásek</b>	
PTDOS:Tool to support DoS testing .....	291
<b>M. Liska and M. Štůsek</b>	
Network Media Player with Built-in Class-D Amplifier.....	295
<b>M. Tran and P. Dzurenda</b>	
NFT Minting System on Cardano Blockchain .....	299
<b>D. Kondys and D. Smékal</b>	
FPGA Digital Circuit for up to 400 Gbps Transfers over Ethernet .....	303
<b>M. Mahút and P. Musil</b>	
Impact of devices connected to powerline on selected BPL topologies .....	307

## Doctoral papers

### D1 – Power Electronics, Electric Drives and Electrical Machines

<b>V. Sizonenko, O. Vitek</b>	
Literature Overview of Fault Tolerant Electrical Machines .....	311
<b>D. Pribulla and O. Vitek</b>	
Design of a High-speed Solid-rotor Asynchronous Generator for an Expansion Turbine .....	316
<b>J. Kaser and J. Valenta</b>	
Mathematical Model of RCCB Sensitive to Pulsating Direct Currents .....	321
<b>P. Klíma and O. Vitek</b>	
Influence of the magnet shape on the rotor eddy-current losses in the outer rotor permanent magnet synchronous motor .....	326
<b>M. Mrajca and R. Cipin</b>	
Axial Channel Effect on Leakage Reactance in Split-Winding Transformer .....	331
<b>D. Růra and J. Bárta</b>	
An Analysis of Conical Homopolar Electrodynamic Bearing .....	336

### D2 – Energy and Distribution Systems

<b>T. Cerny and M. Ptacek</b>	
A Nodalization Study of Phebus FPT-1 Experiment in MELCOR 2.2 .....	341
<b>F. Novák and P. Baxant</b>	
Plunging Brno into darkness: Measuring the contribution of public lighting to the total luminance of the night sky .....	347
<b>J. Burian</b>	
Experimental investigation of sodium chloride neutron induced reactions for nuclear reactors cooled by molten salts .....	352

### D3 – Systems Modeling, Simulation, Control and Measurement

<b>K. Jurik and P. Drexler</b>	
An analysis of the birdcage resonators using magnetic field probes, for validation of optimal aspect ratio .....	358
<b>J. Diblík and E. Korobko</b>	
Vanishing solutions of a second-order discrete non-linear equation of Emden-Fowler type .....	363
<b>P. Raichl, P. Marcon</b>	
Testing of UWB Sensors For Use in Multi-robot Relative Localization Systems and Design of UWB Plugin for ROS/Gazebo Simulation Environment .....	368
<b>V. Mancl, P. Beneš</b>	
Optimization of electrostatic simulation by utilization of distribution equations with emphasis on design of capacitance sensors of inhomogeneity distance .....	373
<b>M. Mačák, P. Vyroubal</b>	
Multiscale Modelling of Li-Ion Battery Volume Expansion .....	378
<b>P. Dvorský, O. Baštán, and P. Fiedler</b>	
Complex systems resilience qualification .....	383
<b>M. Hemzal, V. Aubrecht</b>	
Study of the effect of tissue conductivity on electroporation .....	388

### D4 – Microelectronics and Technology

<b>M. Šedina and T. Kazda</b>	
History of EVs and measuring battery under external force .....	393
<b>O. Klvač, T. Kazda, Y. Fam and L. Novák</b>	
SEM Micromanipulator as Current Collector for Battery In-situ Characterization .....	398
<b>Z. Košelová</b>	
Nanoporous Enzyme-Based Impedimetric Biosensor .....	403

<b>D. Misiurev</b>	
Lithographic thin-film structures based on Electrochromic materials: Case study by scanning probe microscopy .....	409

## D5 – Bioengineering and Bioinformatics

<b>J. Schwarzerová and D. Čejková</b>	
Identification of horizontal genes transfer elements across strains inhabiting the same niche using pan-genome analysis .....	416
<b>A. Hamza, M. Mezl, and S. Zafar</b>	
Calculating the Cortical Thickness of Brain by using electrical and mathematical models for Optimal Transcranial Direct Current Stimulation .....	421
<b>D. Kuřátko, and Z. Raida</b>	
Anisotropic Forward Modeling of Electroencephalogram in MATLAB - CST Studio Suite Interface .....	427
<b>M. Ezati, I. Zumberg, V. Cmiel</b>	
Mitochondria morphology and membrane potential under stress conditions .....	432
<b>R. Ředina</b>	
Optimization of wavelet transform in the task of intracardiac ECG segmentation .....	437
<b>M. Nohel, J. Jan, and J. Chmelik</b>	
Characterization of osteolytic lesions by low-dose spectral CT in myeloma patients .....	442
<b>S. Kouakouo, V. Böhm, and K. Zimmermann</b>	
Analyses of apedal locomotion systems based on ferroelastomers .....	447

## D6 – Analog and Digital Signal Processing

<b>M. Rujzl</b>	
Design of hybrid computer for modelling of dynamic systems .....	454
<b>S. Bilik and K. Horak</b>	
SIFT and SURF based feature extraction for the anomaly detection .....	459
<b>A. Tomašov and T. Horváth</b>	
Performance comparison of a signal processing pipeline execution using CPU and GPU .....	465

## D7 – Communication Systems and Network Security I.

<b>E. Holasova, K. Kuchar, and R. Fujdiak</b>	
Network Supervision via Protocol Identification in the Network .....	470
<b>Shujairi Murtadha, and V. Škorpil</b>	
SVM Algorithm Training for DDoS on SDN Networks .....	475
<b>M. Rusz</b>	
Multi-purpose PLC gateway for management, control and visualization .....	480
<b>J. Pospisil, R. Fujdiak</b>	
Radio spectrum analysis with a focus on WiFi bands using software-defined radio .....	485

## D8 – Communication Systems and Network Security II.

<b>D. Kohout</b>	
Application for Testing Smart Meters .....	491
<b>A. Boháčik, R. Fujdiak, and J. Mišurec</b>	
Network probe: Network monitoring and management tool .....	496
<b>L. Benešl, P. Mlýnek and D. Harman</b>	
Pilot testing of LoRaWAN technology in disconnecting fuse boxes .....	501
<b>T. Stodulka, Z. Martinasek</b>	
Introduction of Cyber Arena .....	506
<b>E. Anderson</b>	
Performance Metrics Evaluations for Selected Proactive Routing Protocols in Smart Grid Neighbourhood Area Wireless Mesh Network .....	511





# STUDENT EEICT

## OPENING WORD OF THE DEAN

These Proceedings contain papers presented during the **28<sup>th</sup> annual STUDENT EEICT conference**, held at the Faculty of Electrical Engineering and Communication, Brno University of Technology, on April 26, 2022. The fruitful tradition of joining together creative students and seasoned science or research specialists and industry-based experts was not discontinued, providing again a valuable opportunity to exchange information and experience.

The EEICT involves multiple corporate partners, collaborators, and evaluators, whose intensive support is highly appreciated. Importantly, the competitive, motivating features of the conference are associated with a practical impact: In addition to encouraging students to further develop their knowledge, interests, and employability potential, the forum directly offers career opportunities through the affiliated PerFEKT JobFair, a yearly job-related workshop and exhibition complementing the actual EEICT sessions. In this context, the organizers acknowledge the long-term assistance from the Ministry of Education, Youth and Sports of the Czech Republic, which has proved essential for refining the scope and impact of the symposium.

In total, 174 peer-reviewed full papers distributed between 23 sessions were submitted, before examining boards with industry and academic specialists. The presenting authors exhibited a very high standard of knowledge and communication skills, and the best competitors received prize money and/or small gifts.

Our sincere thanks go to the sponsors, experts, students, and collaborators who participated in, contributed to, and made the conference a continued success.

Considering all the efforts and work invested, I hope that the **28<sup>th</sup> STUDENT EEICT (2022)** has been beneficial for all the participants.

I believe that the inspiration gathered during the event will contribute towards a further rise of open science and research, giving all the attendees a chance to freely discuss their achievements and views.

**Prof. Vladimír Aubrecht**

Dean of the Faculty of Electrical  
Engineering and Communication



BRNO FACULTY OF ELECTRICAL  
UNIVERSITY ENGINEERING  
OF TECHNOLOGY AND COMMUNICATION

# Robotic arm control using MQTT protocol

V. Mautner<sup>1</sup>, D. Olenišák<sup>1</sup>, and M. Jurák<sup>1</sup>

<sup>1</sup>Střední odborné učiliště elektrotechnické Plzeň, Czech Republik

E-mail: [vojtulam@email.cz](mailto:vojtulam@email.cz), [kdansubs@gmail.com](mailto:kdansubs@gmail.com), [jurak@souepi.cz](mailto:jurak@souepi.cz)

**Abstract**— The aim of this work is to create a simple system that will serve mainly as a teaching tool for demonstrating communication in a simple IoT system. The proposed system should allow easily add different devices to the system and establish communication between them and create individual communication commands between devices. The proposed system uses the MQTT protocol, which is widespread in most IoT systems. This protocol is simple and therefore easy to implement even on microcontrollers that can be connected to the system.

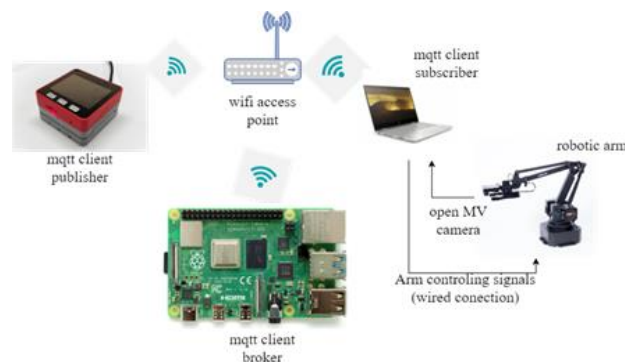
**Keywords**— robotic arm control, MQTT protocol, OpenMV camera,

## 1. INTRODUCTION

MQTT (Message Queuing Telemetry Transport) is a basic protocol used in IoT systems. The advantages of the protocol lie in its simplicity and simple implementation (it can be easily implemented on microcontrollers). The aim of our work was to create a simple educational IoT system that would allow to understand the basic MQTT communication. We chose a robotic arm supplemented by a computer vision module as the basic device of the whole system. This arm is then remotely controlled by other system devices (microcontroller, mobile phone etc.). All devices communicate with each other via MQTT messages and therefore it does not matter in which programming language the device control application is created. It is possible to easily add additional devices to the system and define new MQTT messages for them. This paper describes a simple IoT system in which a robotic arm can be remotely controlled. The structure of the article is as follows. Chapter two deals with the basic technical implementation of the whole system and a description of individual components, Chapter three describes the MQTT protocol, and the format of messages used in the system, Chapter Four describes the results of our work and topics that we would like to implement in the future. Chapter 5 shows the possibilities of practical use of the proposed system

## 2. TECHNICAL REALIZATION

The basis of the whole system is a commercially produced robotic arm uArm Swift Pro, which is connected to the control computer (Fig. 1). A computer vision module is attached to the arm to scan the workspace. The robotic arm can be controlled from various devices connected to the system via the MQTT protocol.



**Figure 1:** Robotic hand control system wiring diagram

One of the devices that can be used to control the robotic arm is the M5Stack module (Fig. 2a). It is a microcontroller based on ESP32, to which a display is connected. The microcontroller has the option of connecting to WIFI and it is possible to create an MQTT communication protocol on it. The M5stack module is programmed in the micro-Python programming language. In our case, the module serves as an MQTT client (sends commands to control the robotic arm and camera and receives information from camera module).



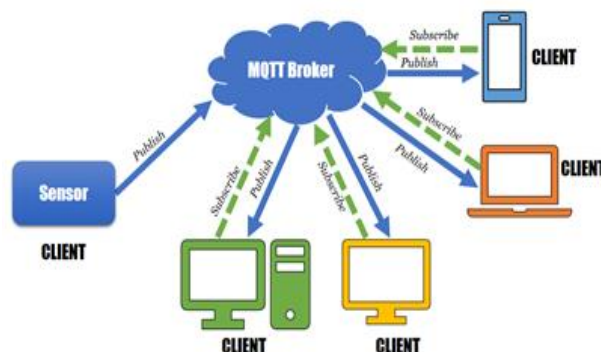
**Figure 2:** a) M5Stack module b) OpenMVH7 Plus camera

Another module used in the system is the computer vision module based on the OpenMV H7 Plus camera (Figure 2b) [1] (a simple OV5640 image sensor with 2592x1944 pixels connected to a relatively powerful microcontroller STM32H743II ARM Cortex M7 running at 480 MHz). The camera can be connected to the MQTT network via a suitable WiFi module [2] and can process commands from a broker. OpenMV camera is programmable in micro-Python, which implements some advanced computer vision features that allow you to easily create the necessary applications. In our system, OpenMV is involved as an MQTT client (publisher and subscriber).

The Python programming language was used to create the application for controlling the robotic arm.

### 3. MQTT PROTOCOL

The MQTT protocol is the most used protocol in IoT systems (Fig. 3). The basic device of the whole communication is the MQTT server called the so-called broker, which receives messages from clients of publishers (Publishers), sorts them and passes them to clients to subscribers (Subscribers). Messages are sorted and forwarded based on topic, i.e., the publisher sends a message with a given topic to the intermediary, who receives it, and sends it to subscribers who are subscribed to the given topic.



**Figure 3:** MQTT communication network scheme

In our system, the Mosquitto message server running at the Raspberry Pi microcomputer was used. M5Stack module, a computer vision module and an Android mobile phone were used as clients. Messages in the MQTT protocol consist of two parts – topic (it can be text string) and own message (it can be a text string, byte array, image, etc.) Clients - subscribers join the subscription of the given topic

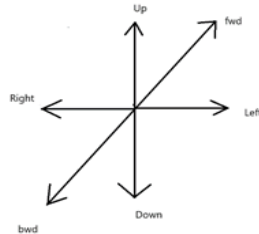
In our case, we have used 3 topics so far:

- **RAC** (Robotic Arm Control) - a control computer (PC) is connected to the subscription of this topic, to which a robotic arm is connected, the messages control the movement of the arm.

The following text strings are used as the messages, to inform the control computer that it should perform the elementary movement of the arm in individual directions (see Figure 4) or that it should lift or place the object:

**"left", "right", "up", "down", "fwd", "bwd", "set\_pos r fi z", "pump on", "pump off"**

The strings "pump on", "pump off" are used to hold the object and to store it (we used an air vacuum compressor, located at the end of the arm)



**Figure 4:** basic directions of robotic arm movement

- **CAM** (Camera) - the camera is attached to the subscription of this topic; the messages contain a question whether the object of the given color is in the working field of the arm. The following strings are used:

**"red", "green", "yellow"**

- **ESP** - the M5Stack microcontroller is connected to this topic, the messages are used to inform the microcontroller whether the object was found or not and in what position. The following strings are used

**"find\_obj x y ", "not found"**

It is possible to add other messages and topics to the system as needed (e.g., to find other objects with the camera, etc.)

#### 4. RESULTS AN FUTURE WORK

Figure 5 shows a system for remote control of a robotic arm. So far, we have implemented the following functions:

- Creating MQTT communication between devices.
- A control application in Python language was created for the robotic arm.
- An application in micro-Python language that finds an object of a given color has been created for the OpenMV camera.
- An application in micro-Python which allows you to build a three-color tower from cubes was created on the M5stack module,
- An application (in Java) was created on an Android mobile phone, which enables control of the robotic arm in basic directions and enables the transfer of objects.

In our future work, we would like to implement one of the following points:

- Image transfer of the captured scene to the LCD display of M5stack, PC or on a mobile phone so that it is possible to monitor what objects are in the camera workspace,
- use of brainwaves to control the robotic arm - you need to use a suitable EEG headband, which was not yet available,
- using the Leap motion controller (a device used in VR) to control robotic arm by more complex gestures.





**Figure 5:** Robotic Arm control system

## 5. CONCLUSION

A simple educational system showing the possibilities of the MQTT protocol, was created, However, the resulting system could find other practical uses in the following areas:

- Neuro-rehabilitation - instead of a robotic arm, a rehabilitation robot would be used, in which a rehabilitated arm would be hung.
- Industrial applications such as search for an object and transfer / transport it to a specified location - in this case, a mobile robotic arm would be used, which would allow searching for a specified object (e.g., a package of a given color, with a corresponding barcode, etc.) and transfer it to the destination position.

## REFERENCES

- [1] OpenMV camera H7 description. Accessed March 03, 2022. [Online]. Available: <https://openmv.io/products/openmv-cam-h7>
- [2] WiFi shield for OpenMV camera description. Accessed March 03, 2022. [Online]. Available: <https://openmv.io/products/wifi-shield-1>
- [3] OpenMV MicroPython documentation 1.15. Accessed March 03, 2022. [Online]. Available: <https://docs.openmv.io/>
- [4] uArm-Python-SDK Manual, Accessed March 03, 2022. [Online]. Available: <https://github.com/uArm-Developer/uArm-Python-SDK>
- [5] uArm-SwiftPro-User-Manual Accessed March 03, 2022. [Online]. Available: [https://cdn.shopify.com/s/files/1/0012/6979/2886/files/uArm\\_Swift\\_Pro\\_User\\_Manual-V1.1.23.pdf?v=1615280407](https://cdn.shopify.com/s/files/1/0012/6979/2886/files/uArm_Swift_Pro_User_Manual-V1.1.23.pdf?v=1615280407)
- [6] M5stack Fire IoT development Kit Manual Accessed March 03, 2022. [Online]. Available: <https://m5stack.hackster.io/products/m5stack-fire-iot-development-kit-psram-2-0>

# A Line Following car using OpenMV camera

V. Mautner<sup>1</sup>, D. Olenišák<sup>1</sup>, and M. Jurák<sup>1</sup>

<sup>1</sup>Střední odborné učiliště elektrotechnické Plzeň, Czech Republik

E-mail: [vojtulam@email.cz](mailto:vojtulam@email.cz), [kdansubs@gmail.com](mailto:kdansubs@gmail.com), [jurak@souepi.cz](mailto:jurak@souepi.cz)

**Abstract**— Most cars that follow the line use IR sensors for tracking. The accuracy of line tracking then depends greatly on the number and type of sensors. The car we designed uses an OpenMV camera to monitor the line. The OpenMV camera is a small camera connected to a microcontroller and in conjunction with the motor control board allows you to design a simple solution whose accuracy is much higher compared some types of IR sensors.

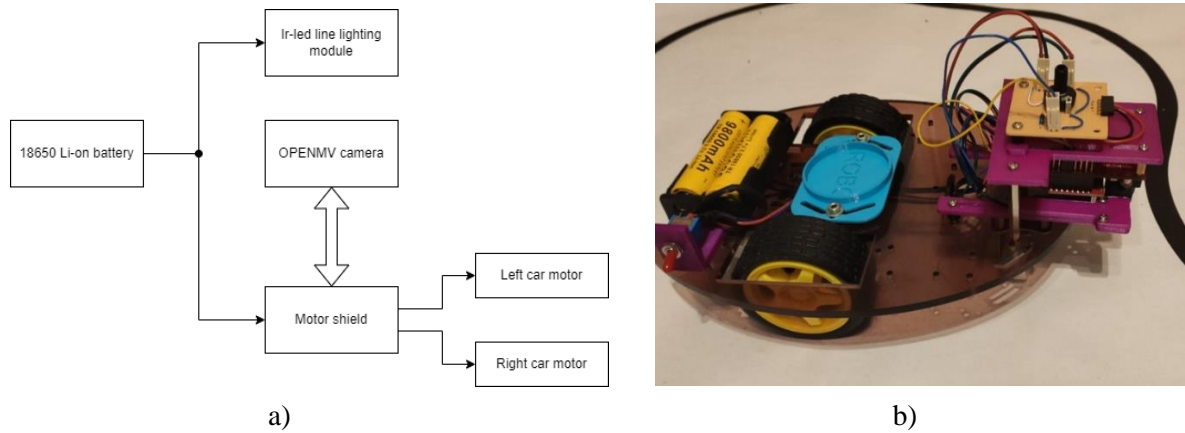
**Keywords**— line following car, OpenMV camera, image processing,

## 1. INTRODUCTION

Most cars that follow the line use IR sensors for tracking. The accuracy of line tracking then depends greatly on the number and type of sensors. Some IR sensors designed for line tracking are purely digital and only allow you to determine whether the sensor is on a black line or not. In contrast, analog sensors make it possible to determine more precisely the extent to which the sensor is on the line, thus enabling more precise control. The disadvantage of IR sensors is that a larger number of usually analog sensors are needed for accurate monitoring. Another disadvantage is that the sensors detect only a narrow part of the line and do not allow to predict what the next behavior will be (occurrence of a turn, etc.). In this article, we deal with the construction and operation of a car that uses an OpenMV camera to monitor the line. OpenMV is a camera created within the startup project. It is a camera chip that is connected to a microcontroller. In this case a microcontroller based on ARM Cortex M7 is used. In conjunction with the simple motor controller that comes with the camera, it is possible to design simple hardware using a minimum number of components. Moreover, the OpenMV camera uses micro-python for programming. The interpreter of micro-python language is in the microcontroller's memory, so creating an application is simple and does not require tedious translations into machine code as with other microcontrollers. In addition, the micro-python library [3] contains many basic operations for processing the image obtained by the camera, which greatly simplifies the development of the line tracking application. This article deals with the design of a car that uses the mentioned OpenMV camera and the description of motor control is based on the information obtained from the camera. The structure of the article is as follows. Chapter two shows the basic hardware diagram for car control and provides a more detailed description of each part. Chapter three shows a flowchart of a car control application and the basic formula used to control engines. Chapter four evaluates the achieved results.

## 2. TECHNICAL SOLUTION

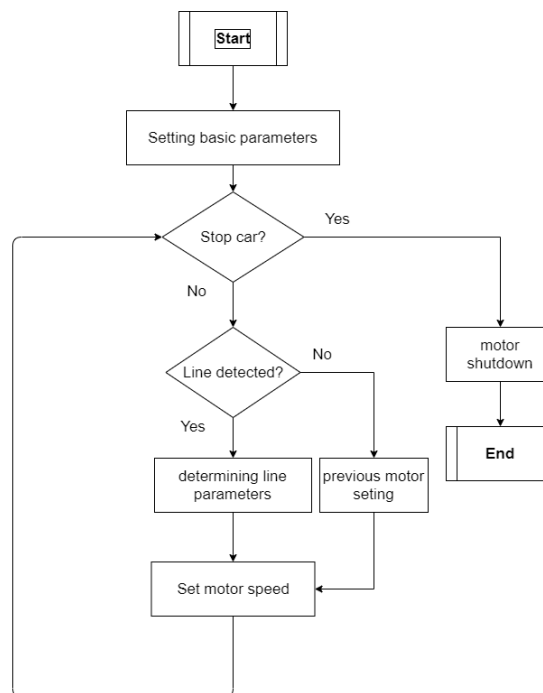
Figure 1 shows the technical solution of the whole car. The basis is the OpenMV camera [1], which was mentioned in the introductory chapter. It is an OV 7725 camera chip with a maximum resolution of 640x480 pixels, which is connected to a STM32H743VI microcontroller (ARM Cortex M7), which runs at a frequency of 480MHz, which ensures sufficient speed for running image processing applications (especially at lower resolutions). A motor shield [2] connected to the camera provides a convenient way of powering and controlling low-power motors. It features a 5V linear regulator to power of OpenMV camera and a motor driver chip to power two less than 2A low-power linear motors from the same battery. The camera with motor shield is mounted on a standard chassis for robotic cars and two li-ion 18650 batteries are used to power the entire system. Infrared LEDs are located near the camera, which serve as additional illumination of the monitored line (the car can reliably monitor the line even in very low light). In addition, this additional lighting serves to eliminate shadows, which greatly affected the line detection algorithm, and the car often left the line.



**Figure 1:** Technical solution of the car – a) block scheme of the system, b) technical realization

### 3. SOFTWARE SOLUTION

A micro-python application was created to control the car's engines. The flowchart of the application is shown in Figure 2. First, the camera is initialized (resolution settings, color system, etc.) and at the same time the basic speed of the motors is set. Pulse width modulation (PWM) is used to control motor speed



**Figure 2:** Flowchart of the application

Subsequently, the motors are started, and the application code goes into a cycle in which it remains until it is decided to switch the motors off (in our application, the motors were switched off when the monitored line was crossed out vertically by another line). Each time the cycle is passed, the camera image is recorded (in resolution 360x240 pixels), and a line is detected in the image by thresholding

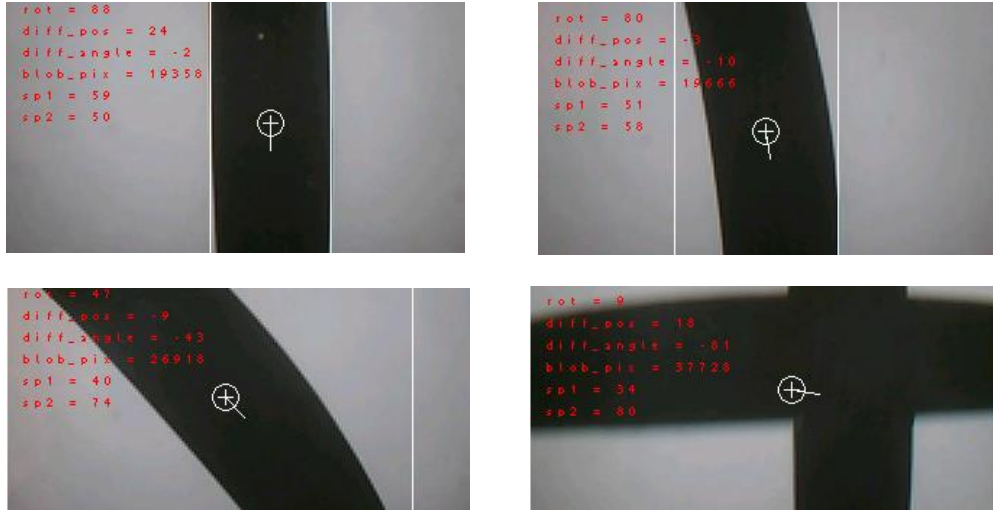
methods. The micro-python library function - *find\_blobs* () is used to detect the line. One of the input parameters of this function is the range of brightness levels of the detected object (line), that we want to find. The Threshold Editor utility, which is part of the OPENMV IDE, can be used to determine this range. The above-mentioned function returns the parameters of the found area (bounding rectangle, center of the object and its slope). Based on these parameters, the speed of the individual motors was adjusted so that the center of the detected line is in the vertical axis of the scanned image and its slope corresponds to an angle of approximately 90 degrees (the detected line is parallel to the vertical axis of the scanned image). Each time the center of the detected line deviates from the vertical axis of the image by more than 10 pixels, or the line angle is less than 85 degrees or greater than 95 degrees, speed of the motors is corrected so that the car turns to the line direction and these deviations are eliminated. To speed correction the following equations are used.

$$M_{left} = SPEED + difference + L\_corr \quad (1)$$

$$M_{right} = SPEED - difference + R\_corr \quad (2)$$

$$difference = a * diff_{pos} + b * diff_{angle} \quad (3)$$

where  $M_{right}$  and  $M_{left}$  denote left and right motor speed (in interval 0-100),  $SPEED$  is basic car speed (this value was set to 65 in our case),  $diff_{pos}$  is the difference between center of line and center of the image (in horizontal direction),  $diff_{angle}$  is difference between slope of the line and 90 degrees,  $a$  and  $b$  are the constants that were set during the test phase ( $a=0.4$ ,  $b=0.6$ ).  $L\_corr$  and  $R\_corr$  are correction constants that are used when the center of the line is too far from the center of image and a larger car rotation is required. Both constants were set to the same value ( $L\_corr = R\_corr = 5$ ).



**Figure 3:** Images from the camera recorded while driving

Figure 3 shows images recorded by the camera while driving. The pictures show the detected line and its parameters and deviations from the required position and speed setting of individual motors. The picture on the bottom right is a picture of a line that is used to stop the car's engines. The stop parameter is the size of the black area *blob\_pix*.



#### 4. CONCLUSION

The designed line following car was tested within the International Competition of Robotic Vehicles ROBO 2022 held in Pilsen [4]. One of the disciplines was to cross a given track (see Figure 4) with a container half filled with water. The water was not allowed to spill while driving. In our case, the container was placed in a blue holder in the middle of the car (see Figure 1 b). The designed vehicle won the Best Ride Award category and passed the route shown in 33 seconds. Line tracking accuracy was also good compared to other trucks.

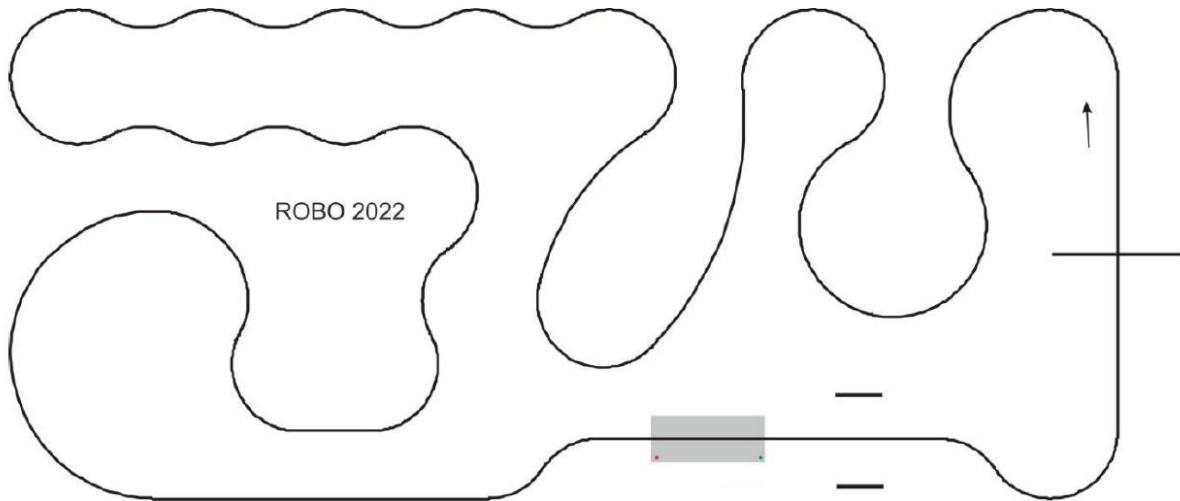


Figure 4: Picture of track shape (size of the track is 3 by 2 meters)

#### REFERENCES

- [1] OpenMV camera H7 description. Accessed March 03, 2022. [Online]. Available: <https://openmv.io/products/openmv-cam-h7>
- [2] Motor shield for OpenMV camera description. Accessed March 03, 2022. [Online]. Available: <https://openmv.io/products/motor-shield>
- [3] OpenMV MicroPython documentation 1.15. Accessed March 03, 2022. [Online]. Available: <https://docs.openmv.io/>
- [4] The ROBO 2022 Competition Rules. Accessed March 03, 2022. [Online]. Available: [https://www.souepl.cz/wp-content/uploads/2021/12/Pravidla\\_souteze\\_ROBO2022\\_ver7\\_EN.pdf](https://www.souepl.cz/wp-content/uploads/2021/12/Pravidla_souteze_ROBO2022_ver7_EN.pdf)

# Scientific Calculator

V. Glombíček<sup>1</sup>

<sup>1</sup>Gymnázium Josefa Božka, Český Těšín, Czech republic

E-mail: [glombicek.vojtech@gmct.cz](mailto:glombicek.vojtech@gmct.cz)

**Abstract**— This work deals with designing and constructing a miniature scientific calculator. The thesis describes the design of the electronic circuit, programming of the microcontroller and mechanical design of 3D printed parts.

**Keywords**—Scientific calculator, Printed Circuit Board, STM32, Arduino, Keypad, 3D Print, Battery, IC (Integrated Circuit)

## 1. INTRODUCTION

The goal of this work is to create a portable battery-powered platform which can be used for implementing a basic scientific calculator based on a popular microcontroller STM32F103C8T6 (also commonly known as bluepill) programmed using the Arduino framework. The code was written in Visual Studio Code running PlatformIO which is very advanced as opposed to the fairly basic Arduino IDE.

## 2. PRINTED CIRCUIT BOARD

In order to achieve a very small size of the device a PCB (Printed circuit board) was created. The board was designed in an online application EasyEDA [1] and manufactured by a chinese prototype board house JLCPCB.

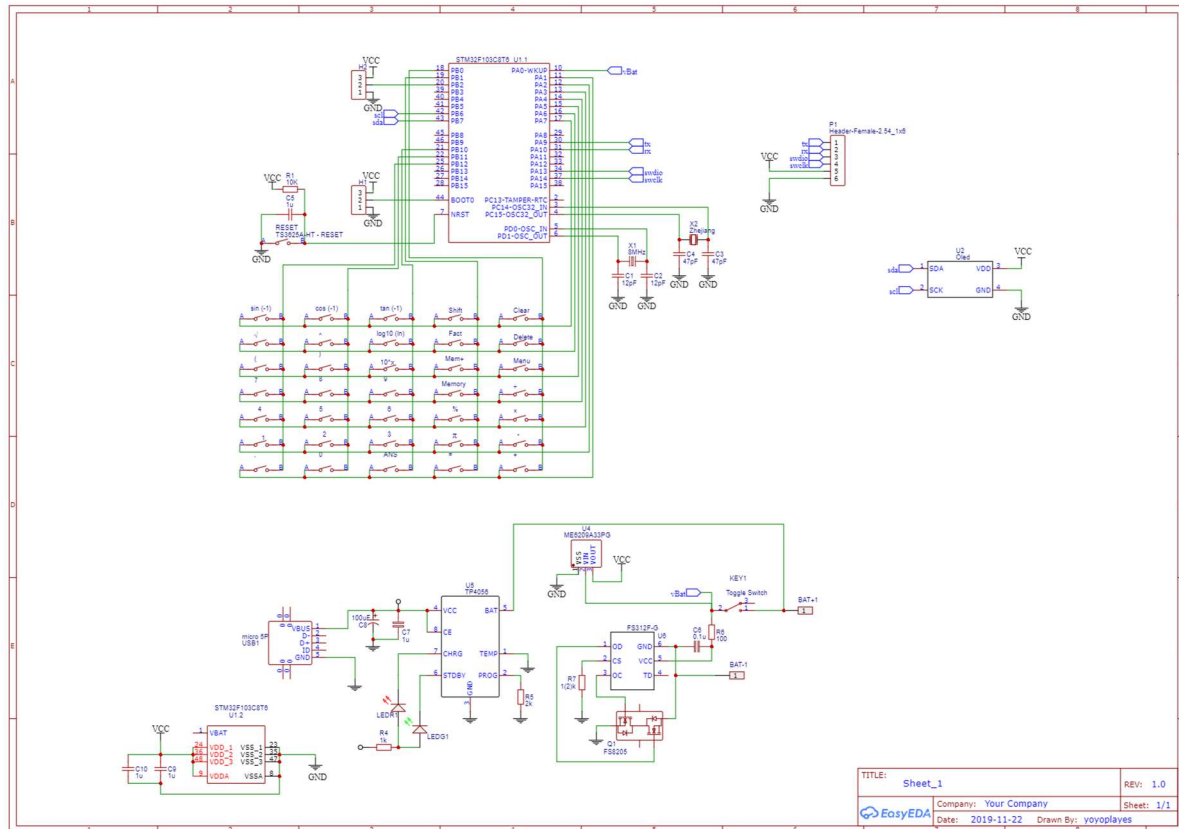
Most of the electronic parts are of the SMT type (Surface-mount technology) to further reduce the size.

The board consists of three sections: power, microcontroller and buttons.

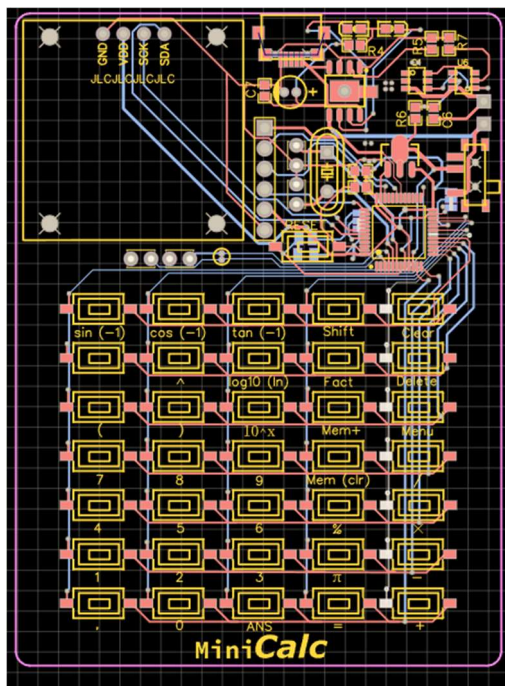
The power section includes a small lithium-polymer cell and a corresponding TP4056 charger IC with signal LEDs, which can be connected to a power source via a micro USB connector. The battery voltage is routed through a power switch into a linear voltage regulator which powers the device. Originally the power circuitry also involved two ICs for undervoltage and overcurrent battery protection, but these needed to be bypassed due to fake chips.

The microcontroller section is composed of a small OLED display connected the MCU itself via I<sup>2</sup>C and other supporting parts such as decoupling capacitors, 8MHz crystal, a reset button and breakout pins for UART and ST-LINK interface. These are used to connect to a computer for programming and debugging.

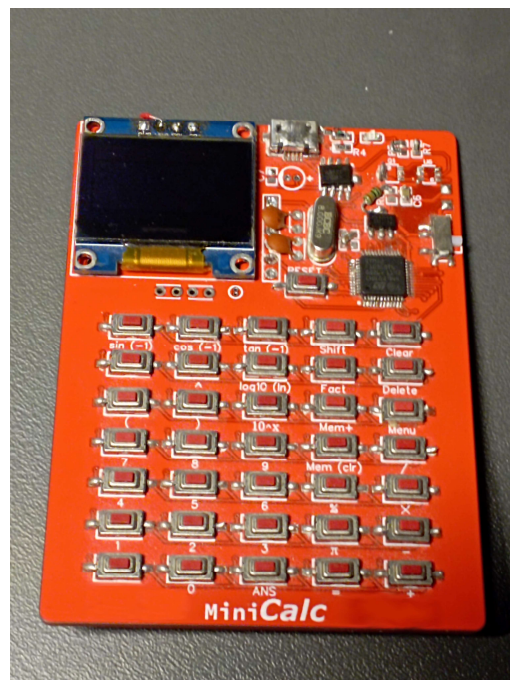
The keypad section is a 5x7 array of tactile pushbuttons which are connected to the microcontroller in a matrix fashion.



**Figure 1: Circuit schematic**



**Figure 2: PCB Layout**



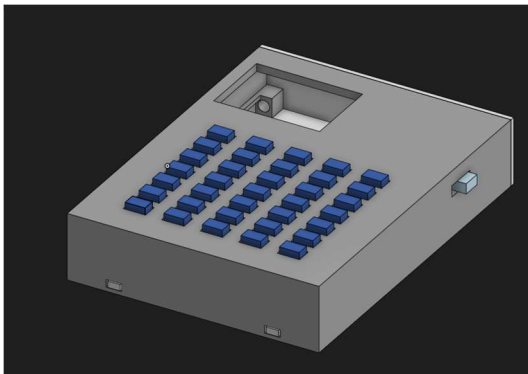
**Figure 3: Populated PCB**

### 3. ENCLOSURE

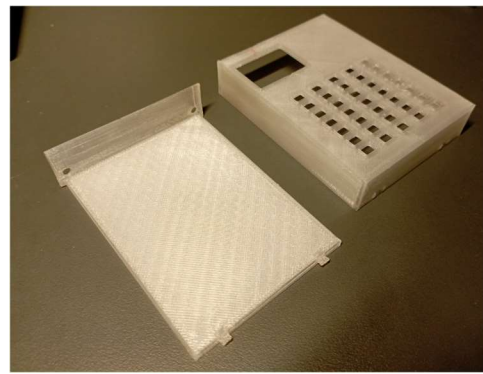
The enclosure is 3D printed on a filament-based 3D printer. It's two part design is fairly simplistic, needing only two M3 bolts to be put together. The model was created in an online application OnShape [2].

The PCB is held in place by integrated guide rails which ensure correct spacing between the board and the enclosure.

A set of 35 keycaps must be installed beforehand. These keycaps are 3D printed as well, with a paper label stuck on them. Due to the PETG plastic's properties regular stickers don't work, so a stronger Chemoprén type of glue had to be used.



**Figure 4:** 3D model of enclosure



**Figure 5:** Printed parts



**Figure 6:** Completed calculator

#### 4. SOFTWARE

The core component of the software is a shunting yard algorithm [3]. This algorithm was invented by Edsger Dijkstra and its purpose is to convert infix notation (which is commonly used as input to most calculators) to postfix (a.k.a. reverse-polish notation) which is much more convenient to further process. This allows the calculator to solve compound expressions with parentheses.

The calculator also allows for utilizing trigonometric functions by replacing them in the input string by their value and passing it on.

Examples:

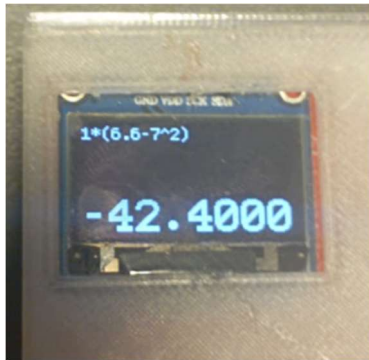


Figure 7:  $1 \times (6.6 - 7^2) = -42.4$



Figure 8 and 9:  $\sin(50) + \cos(40) = 1.532080$

#### 5. CONCLUSION

This work is mainly experimental as the final calculator is uncomfortable and impractical to use when compared to mass-produced calculators. It serves as a very good exercise in coding, 3D modeling in CAD software and designing circuit boards.

#### REFERENCES

- [1] EasyEDA [online]. [cit. 2022-03-11]. Available from: <https://easyeda.com/>
- [2] OnShape [online]. [cit. 2022-03-11]. Available from: <https://www.onshape.com/en/>
- [3] NORVELL, Theodore. Parsing Expressions by Recursive Descent [online]. 1999 [cit. 2022-03-11]. Available from: [https://www.engr.mun.ca/~theo/Misc/exp\\_parsing.htm](https://www.engr.mun.ca/~theo/Misc/exp_parsing.htm)



# Wirelessly Controlled Robotic Swarm

**Sebastian Matoušek**

Gymnázium Matyáše Lercha

E-mail: [seb.matousek@gmail.com](mailto:seb.matousek@gmail.com)

Supervised by: prof. Ing. Roman Maršálek, Ph.D.

E-mail: [marsaler@vut.cz](mailto:marsaler@vut.cz)

**Abstract**—The focus of this paper is to create a team of four robots that have a common goal and a central unit – an Android app – coordinating them. Each of the robots has a wide range of sensors at its disposal, like a gyroscope, an accelerometer, encoders or a range sensor that help the robots orientate in the terrain in which they are operating. An Android app was created for controlling the robots. The application also displays maps which the robots create, and plot the graphs from their data.

**Keywords**—Robot Swarm, ESP32, Wireless Communication, WiFi

## 1. INTRODUCTION

One of today's trends in robotics are so-called multi-agent systems, which are the networks of individual robots communicating with each other, trying to achieve a common goal. The potential utilization of such systems could be in company warehouses, in situations after a natural disaster, delivering mail and so on.

The presented paper deals with this particular subject, mainly with design, construction, and programming of a group of robots, which have a common goal – to map the area in which they are operating. All robots communicate wirelessly with an Android application on the users mobile phone, presenting the user with the data that the robots collect, which includes for example their battery status, the surrounding temperature, etc.

## 2. MAIN GOALS

The result of the project should be a group of four moving, wirelessly communicating robots. The group was supposed to be expandable, so all robots were meant to have almost identical construction and software. Every robot should be fitted with sensors that would collect data about its state and the state of its surroundings, and this information was supposed to be communicated to the user using a mobile application. Requirements for the robots constructed were as follows:

1. Wireless communication with the Android mobile application
2. The ability of said robot swarm to map the footprint of the area in which it operates
3. Compact dimensions of the robots (max. 15x15x15 cm)
4. Simple construction of individual robots
5. A sensor measuring battery voltage
6. A temperature sensor
7. Option for robots to be equipped with a camera (only one robot was equipped with the camera due to limited financial resources)

## 3. ROBOT ELECTRONICS

All electronic peripherals the robots consist of are shown on a block diagram in Fig. 1. Only the most important peripherals will be discussed in detail. All functions of the individual peripherals are summarized in Table I.

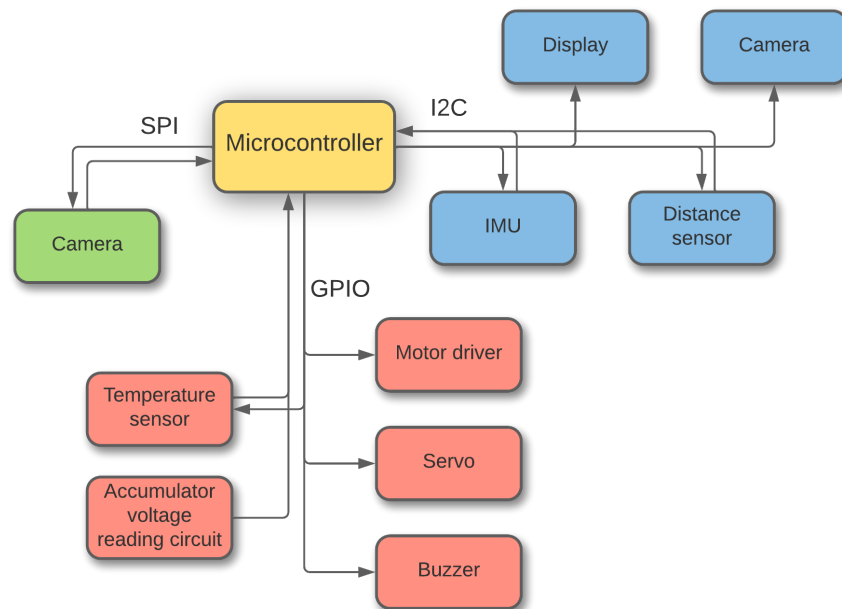


**ESP-32** Every robot is controlled by the microcontroller ESP-32. This microcontroller is equipped with WiFi, which the robot uses for communication with the Android app, and therefore the robot doesn't have to be fitted with another module for WiFi. The ESP-32 has two Tensilica Xtensa 32-bit LX6 processor cores, which allow the robot to stream video from its camera, while controlling its motors and mapping the area.

**Camera** All robots have a terminal for the camera Arducam 5MP+ on their PCB, which means any robot can be fitted with this camera. It allows the user to see from the robot's perspective, as the image from the camera is transferred to the Android app.

**Distance Sensor** Robots are equipped with the VL53L0X distance sensor. They use this sensor to orientate themselves in the area and to prevent crashing into walls. Thanks to the compact dimensions of this sensor, it can be attached to a MG-90S servo and turned, so that the robot can "look" around.

**Motors** The robot movement is driven by two DC motors equipped with encoders. These motors aren't visible in the block chart (Fig. 1.), as they are not directly connected to the microcontroller, but rather connected to and controlled by the motor driver. The encoders grant precise control over the movement of the robots.



**Figure 1:** Block chart of robot electronics

**PCB** A custom Printed Circuit Board (PCB) was created for the robots, so that the construction would be simple and all robots would have near-identical electronics. The PCB also helps with cable management of the robots.

#### 4. ROBOT MECHANICAL COMPONENTS

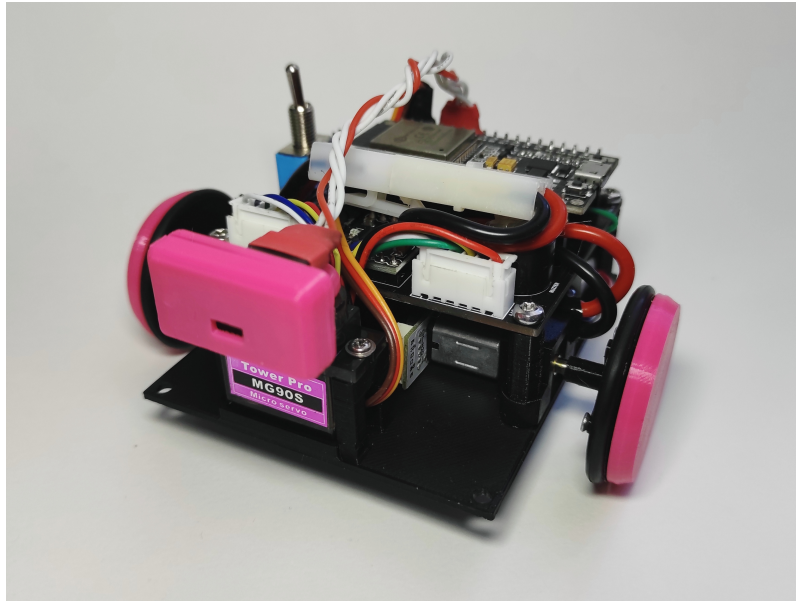
Most of the robot's mechanical hardware components are 3D printed. That means that all robots are almost identical and that all the parts are rigid, and can be manufactured quickly and cost-effectively.

The material used for 3D printing the parts is Polylactic acid (PLA), which is a nice compromise between flexibility and strength. The primary color of the robots is black, and each robot has a different detail color (green, pink, blue, gold) to distinguish the robots.

An image of an assembled robot (without its cover, so that the inner parts are visible) can be seen on Fig. 2.

**Table I:** Functions of individual robot peripherals

Periphery	Function
ESP-32	Controlling the robot and communicating with the Android application
Motor Driver DRV-8833	Controlling both motors of the robot
Servo MG-90S	Turning the distance sensor to a specific angle
Buzzer	Audio communication with the user
Temperature sensor DS18B20	Measuring the surrounding temperature
Accumulator voltage reading circuit	Accurate measuring of accumulator state
Inertial measuring unit (IMU) GY-521	Reserved for future use
Distance sensor VL53L0X	Measuring the distance to surrounding objects and orientation in space
Display	Basic communication with the user through text and pictographs
Camera Arducam 5MP+	Showing video from the robots' perspective to the user

**Figure 2:** Assembled robot

## 5. SOFTWARE

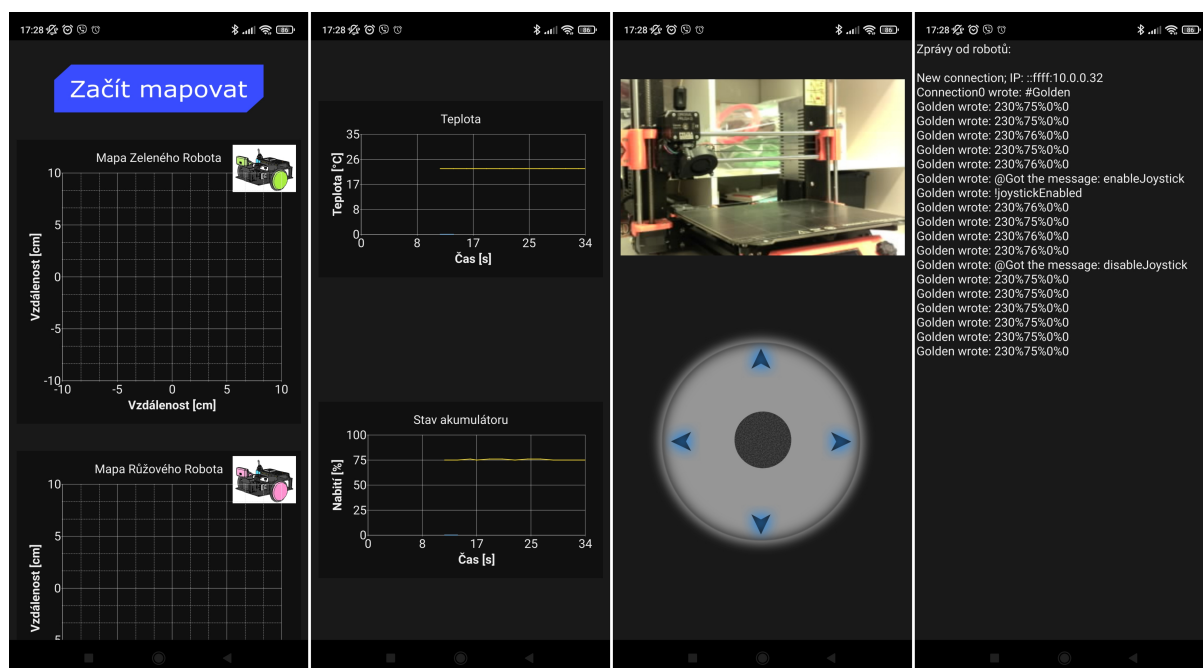
The programming was divided into two parts: the robot software and the Android application. Both were coded in the programming language C++.

**Robot software** The software of the robot controls all its peripherals and houses the algorithm for mapping the area. Each robot connects to the app upon startup, and then awaits the command to begin mapping. When connected to the app, it also sends data such as surrounding temperature and accumulator state to the application at regular intervals. When prompted, the robot equipped with the camera starts streaming the camera image to the user. If the user moves the joystick which is positioned under the

camera image in the app (Fig. 3., third image from the left), the robot will stop all actions and be fully controllable by the user.

**Android application** The main purpose of the application is providing a graphical user interface (GUI) to the user. Through this GUI, the user can then control the robots (start/stop mapping, show camera image, movement control through joystick, etc.) and view all the data and maps sent by the robots.

Each map is created based on the encoder data of one individual robot. Robots send the data to the app, which then interprets it and draws the corresponding map. Different parts of the application are shown in Fig. 3.



**Figure 3:** Parts of the Android application

## 6. CONCLUSION

The result of this project is a group of four robots capable of moving around the area and mapping it. Individual robots communicate with a mobile Android application over WiFi. The application acts as a GUI for controlling the robots and displaying the data sent by the robots.

## ACKNOWLEDGMENT

This thesis was made with financial support from Jihomoravský kraj.

# Potential production and development of photovoltaic panels

Š. Chmelař

Faculty of Electrical Engineering and Communication, Brno university of technology, Czech Republic

E-mail: [xchmel31@vut.cz](mailto:xchmel31@vut.cz)

**Abstract** – The aim of this article is to compare the changes in performance due to the evolution of technology over the period and to compare the most popular calculators, their advantages and disadvantages. In the first part, there are selected three buildings covering a large part of the types of sites for the construction of PV plants. A potential power calculator is applied to these buildings. In the second part, three calculators that are selected for simulating the performance of a model installation have been selected and described.

**Keywords** - photovoltaic panel, solar energy

## 1. INTRODUCTION

Nowadays, the focus of photovoltaic panels tends to be on a better price/performance ratio, thus the maximum output of the panels sold is not as important as the cost of manufacturing them. Power per unit area has been increasing only slowly since the second decade. The biggest breakthrough in PV panel performance was in the 1990s and eventually at the beginning of the millennium. This article compares the performance of PV panels in the context of the development of PV itself and will also focus on the comparison of different calculators calculating the performance of panels and the advantages or disadvantages of each calculator.

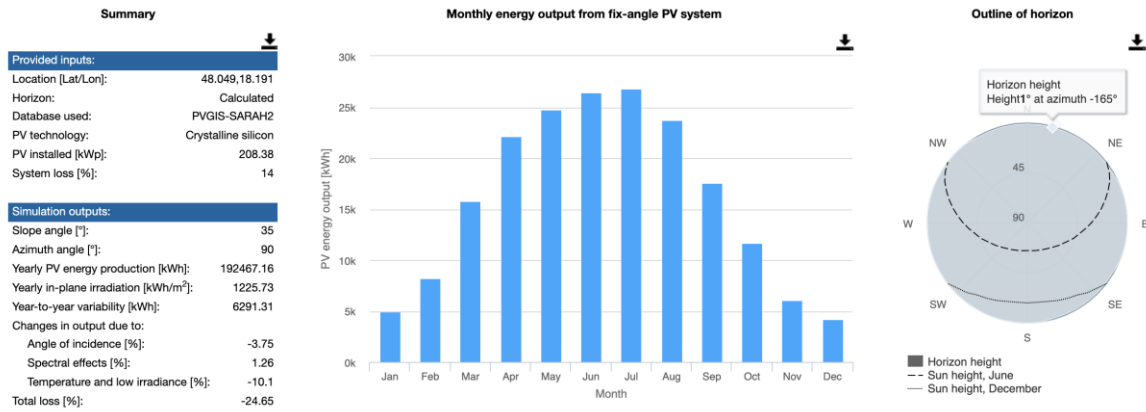
## 2. COMPARISON OF PERFORMANCE IN DEVELOPMENT

In a first step, according to the most suitable scenarios for the PV panels, three specific buildings were selected. The selected buildings belong to the Bánov municipality and were selected for more detailed analyzes for the bachelor thesis. By measuring and calculating the approximate area of roofs that can be used for PV panels, the expected power output on a given area can be determined using calculators based on long-term measurements and observations of ambient conditions such as the ratio of direct and diffuse radiation, the average number of hours the sun shines etc.

For this step, the PVGIS calculator was chosen as it best accounts for ambient conditions from long-term observations in given latitude bands. The input conditions have to include the power installed that is calculated from the average power of the panels and the area where the panels can be installed. Given the objective of comparing PV technology over the last three decades, only the average power of PV panels was changed. The average value of the maximum efficiency of PV panels according to the NREL source was chosen for the calculation. This is the average efficiency of different technologies in a given period. For houses with flat roofs, the south-facing orientation was selected, as it has the best efficiency. For gable roofs, the pitch is determined from maps and then entered into the calculator. The angle of the roof cannot be determined from the maps, so the angle that is common for roof construction is used, which is 35°. In the first case, the value is calculated from today's average PV panel parameters that have been set at 230 Wp/m<sup>2</sup>. For the model example, a single house with a partly east-facing and partly west-facing orientation was selected. The built-up area of the 1<sup>st</sup> house is 1 208 m<sup>2</sup> and its GPS coordinates are 48.9880136N, 17.7175203E. The other building has a gable roof-oriented half north and half south. The building area is 274 m<sup>2</sup> and its GPS coordinates are 48.9897842N, 17.7203292E. In the third case, a building with a flat roof and a built-up area of 339 m<sup>2</sup> was selected and its GPS coordinates are 48.9913717N, 17.7185711E. The building areas are not entirely accurate. They have been calculated on the basis of the maps. In this case, the area measurement tool from the mapy.cz website was used [1].

The built-up area of the first building was measured to 1 208 m<sup>2</sup> using maps. For houses with a gable roof, it is necessary to consider the increased roof area in relation to the built-up area. This increase in area is given by the roof typification coefficient T<sup>st</sup>. The value of this coefficient shall be considered at 1.5 for all houses with a pitched roof [1].

With a panel area utilization factor for the roof of  $K_v = 0.5$ , it follows that 906 m<sup>2</sup> of PV panels can be fitted on this roof. The total calculated output of this building is 192 467 kWh [2].



**Figure 1:** Results from the calculator for a given building [2].

The same calculation has to be made for the other selected years. For a flat roof, the calculation of the usable area differs only in the coefficient  $T_{st}$ , which in this case is equal to 1. The total power calculated by the calculator is equal to 273 624 kWh [5].

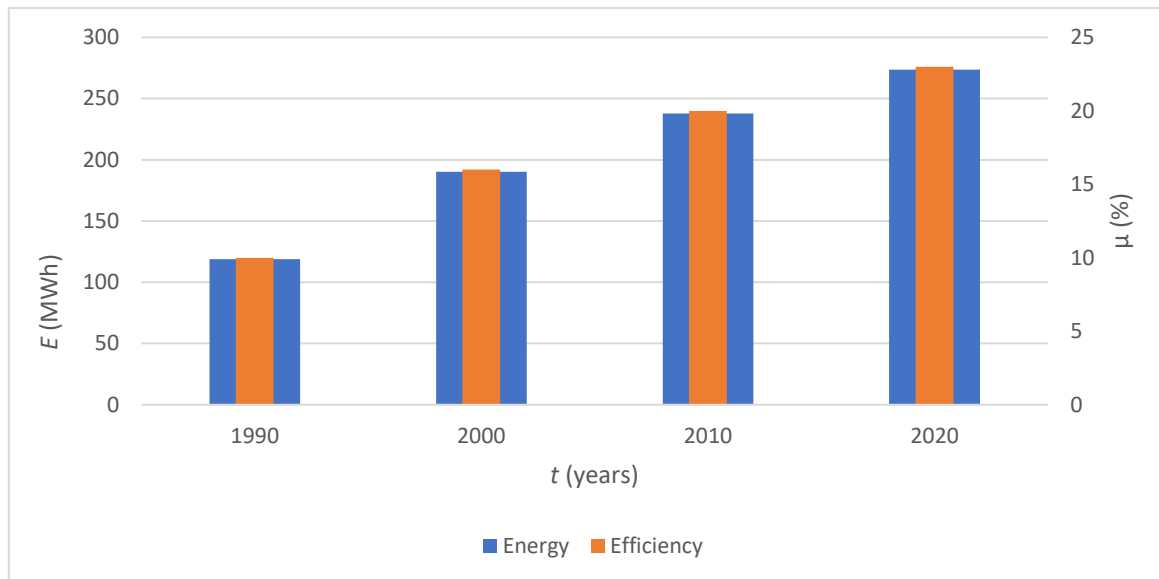
In the 1990s, the average efficiency of a monocrystalline panel was 10-12%. Thus, the known performance is around 100 Wp/m<sup>2</sup>. As for the theoretical power calculated by the PVGIS calculator, the power is equal to 118 967 kWh. The amount of electricity produced in the 1990s would be less than half of what it is today [5].

At the beginning of the millennium, the efficiency increased dramatically and was around 16% that means that the output was around 160 Wp. The total value of the energy produced would be 190 347 kWh according to the calculator [5].

In the last case in the 2010s, the efficiency of commercial panels was around 20%. The panel output was therefore 200 Wp per square meter. The value of the model buildings would be equal to 237,933 kWh [5].

**Table I:** Evolution of the performance of the buildings in relation to the efficiency of the panels (E...electricity generated) [2].

$t$	$E$
year	MWh
1990	118,967
2000	190,347
2010	237,933
2020	273,624



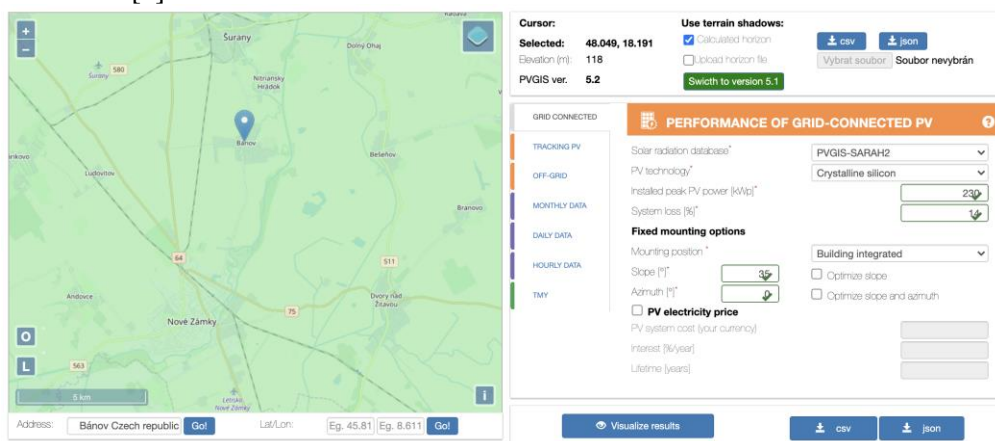
**Figure 2:** Evolution of the performance of the buildings in relation to the efficiency of the panels [2].

### 3. COMPARISON OF CALCULATORS

This chapter compares three different types of calculators, each with its advantages and disadvantages and each suitable for a different type of use.

#### PVGIS

The already mentioned calculator used for the first part of the work is the PVGIS calculator, which is suitable for measuring larger sites. The calculator works on a simplified and less accurate principle where an average value of the area that can be used for PV panels is calculated. It evaluates the average annual output divided into individual months. Also, it takes into account the ambient conditions that have been observed for a certain period of time in a given location. From the input data it is then necessary to know the installed capacity, the system losses, the tilt and rotation of the area on which the panels are installed [2].



**Figure 3:** PVGIS calculator [2].

#### EASY-PV

The EASY-PV calculator is more suitable for individual buildings, where you can select the panel and the roof on which the system is to be installed based on the specified parameters. The calculator can also be used for other areas. Once the area has been generated, the panels can be directly modelled and stacked. The advantage of the calculator is its accuracy. The disadvantage is the long modelling process and also unsuitability for modelling larger areas [3].

## PROJECT SUNROOF

The calculator was created by Google that uses data directly from its maps. It works on a similar principle to the PVGIS calculator. It can directly calculate the area available for the installation of panels and clearly calculates the electricity generated and the payback based on the data filled in. The calculator is only available in the US for now, but it is expected to be launched in Europe in the longer term [4].

## COMPARISON OF CALCULATORS

Each mentioned calculator offers different types of functions. The PVGIS calculator offers potential electricity production in the output data, taking into account the already measured data on ambient conditions. In the output data, the production is modelled on a month-by-month basis based on the ambient conditions during the period. The PV-CASE calculator offers the modelling of a custom PV plant on configured roof and can calculate the installed capacity value of the modelled installation based on the input data. The last calculator, Project Sunroof, combines the functions of the previous calculators and can calculate the possible installed capacity on a given roof using artificial intelligence and then calculate the theoretical value of the electricity produced from the map data.

## 4. CONCLUSION

Photovoltaic technology is constantly evolving and increasing in efficiency. The aim of this work was to compare the theoretical outputs that the panels could produce as the technology evolves. The calculators can be used for future area evaluations in case that lidar data is not available. They can serve as a general assessment of the magnitude of performance that could be achieved in a given area. Panel efficiency has not radically increased by over the last 20 years. Rather, the priority is to reduce the cost of production and improve the price/performance ratio of PV panels. Over the period studied, average panel efficiencies have increased by 5% per decade. The second part was focused on calculators that calculate the theoretical output and the electricity produced. The calculators account for ambient effects and have made great strides in this direction. Three types of calculators have been selected in this article. Each calculator has a specific purpose according to what the customer needs. The first calculator, PVGIS, is best used as a tool for a year-long theoretical overview of the energy produced relative to the specified output of the installation. The EASY-PV calculator is suitable for determining the installed power from a user-modelled installation. Google's latest calculator called Project Sunroof combines both functions of the previous calculators. The calculator uses input directly from maps and uses artificial intelligence to evaluate how much installed power can fit on a given roof and how much total energy the installation can generate.

The work seeks to gain insight into the future development of photovoltaic technology and how the electricity potential of a given area can be calculated. The values can be used to calculate how much electricity can be produced from renewable sources, which are essential for future electricity production.

## REFERENCES

- [1] ŠTEFEK, Martin. Potencionální produkce elektrické energie ze střešních fotovoltaických elektráren v obci do 3000 obyvatel. Brno: Brno University of Technology, Faculty of Electrical Engineering and Communication Technology, 2019. Martin Paar, Ph.D.
- [2] PVGIS calculator. <https://re.jrc.ec.europa.eu> [online]. EU: re.jrc.ec.europa, 2022 [cit. 2022-03-13]. Available from: [https://re.jrc.ec.europa.eu/pvg\\_tools/en/tools.html#api\\_5.2](https://re.jrc.ec.europa.eu/pvg_tools/en/tools.html#api_5.2)
- [3] EASY-PV. <https://pvcase.com> [online]. EU: PVcase, 2022 [cit. 2022-03-13]. Available from: <https://pvcase.com/ground-mount/>
- [4] PROJECT SUNROOF. <https://sunroof.withgoogle.com> [online]. USA: Google, 2022 [cit. 2022-03-13]. Available from: <https://sunroof.withgoogle.com>
- [5] Champion Photovoltaic Module Efficiency Chart. <https://www.nrel.gov/> [online]. USA: Alliance for Sustainable Energy, 2022 [cit. 2022-02-10]. Available from: <https://www.nrel.gov/pv/module-efficiency.html>



# Spot Welding Machine for Li-ion Battery Cells

M. Martinček<sup>1</sup>

<sup>1</sup>Brno University of Technology, Czech Republic

E-mail: [xmarti94@vutbr.cz](mailto:xmarti94@vutbr.cz)

**Abstract**—This article describes creation of spot welding machine prototype. Spot welding enables joining bigger number of Li-ion cells in order to create higher energy storage. Main objective of this article is to design diagrams and algorithms for the spot welding machine. Requirements to regulate current and minimize heat stress on battery cells are implemented in the solution.

**Keywords**— Spot Welding, Resistance Welding, Li-ion Cell, Battery Pack

## 1. INTRODUCTION

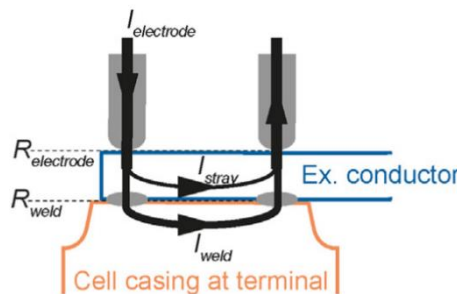
Along with rising need for high energy density, importance of maximizing energy storage capacity is coming to the forefront. This article describes a prototype of spot welding machine for Li-ion battery cells. Prototype of the device was developed with consideration of theoretical concepts and predetermined type of commercially available Li-ion cell. This device was mainly designed for assembling battery packs from single Li-ion cells, specifically cell size 18650.

It is required that spot welding machine creates a joint with low electric resistance that is capable of withstanding mechanical stress during assembly and maintenance. The joint needs to have electrical resistance as low as possible, in order to prevent unnecessary power losses. During the process of the welding, there should be minimum heat stress on cells, to avoid cell damage, as mentioned in Brand et al. [1] and Lee et al. [2].

## 2. SYSTEM CONCEPT

Spot welding is a type of resistance welding that works on principle of electric current generating a heat. In this case, a nickel strip is pressed against a terminal of a battery cell with two copper electrodes that are connected to electric current source.

According to Brand et al. [1], applied electric current consists of an effective current and a stray current, as illustrated in *Figure 1*.



**Figure 1:** Electrical resistances on battery cell during welding process

Stray current flows only through the nickel strip which increases power waste. Effective current flows through the nickel strip and through the cell terminal too. Due to the electrical resistance between nickel strip and the terminal, effective current creates voltage drop and generates heat according to the formula



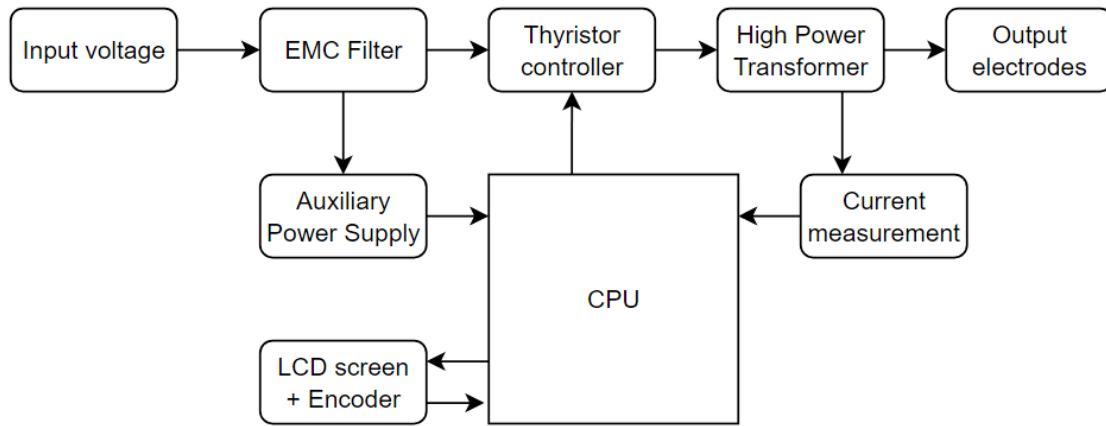
mentioned in Brand et al. [1]:

$$P = R * I^2, \quad (1)$$

where  $P$  represents power loss,  $R$  means electrical resistance, and  $I$  is electrical current. By changing the current flow it is possible to control power delivered to joint and therefore joint parameters.

### 3. DESIGN

Design of the welding machine consists of several modules that each must secure smooth welding operation and easy settings of welding parameters. Diagram of prototype with modules is described in **Figure 2**. Number of modules is reduced to necessary minimum in order to optimize cost-effectiveness of the solution while preserving effective function.



**Figure 2:** Diagram of prototype

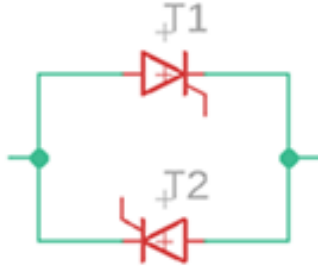
In this design, power regulation is based on using two thyristors as phase control regulator. Different power output can be achieved using different firing angle, which is a moment in the period in which the thyristors are triggered.

To ensure right timing for the thyristor controller, CPU (Central Processing Unit - microcontroller) monitoring waveform of the input voltage and switching gates of the thyristors at the right time is included. This provides appropriate amount of current to the primary winding of the high-power transformer. While primary winding has approximately 250 turns, secondary winding consists of only one turn, which then results in secondary winding current being approximately 250 times higher than the input current coming to the primary winding, according to formula:

$$\frac{N_1}{N_2} = \frac{U_1}{U_2} = \frac{I_2}{I_1}, \quad (2)$$

where  $N_1$  and  $N_2$  represents number of turns of the primary and secondary windings,  $U_1$  and  $U_2$  means primary and secondary electrical voltages,  $I_1$  and  $I_2$  represents primary and secondary electrical currents.

Main element of the circuitry is the thyristor regulator, which consists of two antiparallel thyristors, as illustrated in **Figure 3**. These two thyristors are both being triggered simultaneously by small isolating gate-drive transformer. The gate-drive transformer consists of trifilar winding and represents galvanic isolation between thyristors and logical circuits.



**Figure 3:** Circuit consisting of two antiparallel thyristors

Output current is being measured with a current conveyor with a Hall effect sensor. The current conveyor returns linear voltage according to the amount of the monitored current. This gives feedback to the CPU to adjust the firing angle to the correct value to achieve required current.

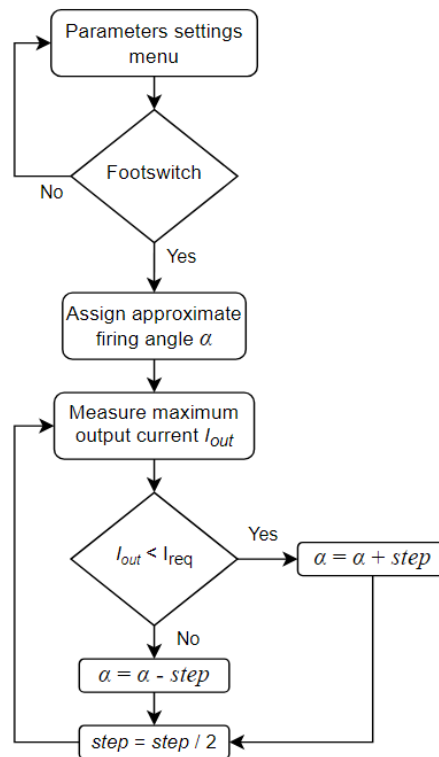
As a protection against electromagnetic interference, the EMC Filter is used. This blocks electromagnetic disturbance that is either being generated by device itself or coming from the external source.

LCD screen and a rotary encoder included in the design of the device enable adjustment of duration of welding process and setting desired output current.

#### 4. ALGORITHM

To achieve desired output current, an algorithm is introduced. This algorithm pictured in **Figure 4** uses a table assembled from an experimental measurement of output current at a different firing angle. Values from this table are used as a starting point to reduce number of iterations required for reaching proper firing angle.

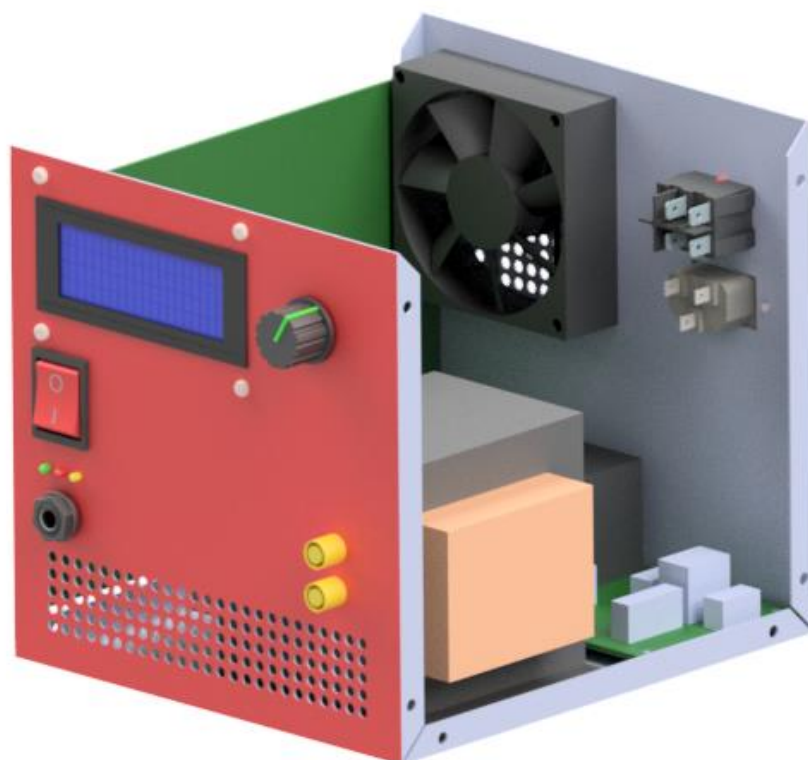
After the first waveform, the firing angle  $\alpha$  is increased or decreased by user defined step. This step is divided by two at each of the iterations, to narrow output current error. Following several iterations, output current error is less than 5 % and effort is made to maintain the error unchanged throughout the whole process of welding.



**Figure 4:** Current controller algorithm diagram

## 5. CONSTRUCTION

Main body of the prototype is a box that consists of top and base parts, base part holding all the elements mounted and top part acting as a covering lid. **Figure 5** illustrates base part including elements. Front panel contains output connectors, footswitch connector, indicating LED diodes, LCD screen and rotary encoder, as mentioned in Chapter 3. Front panel also contains perforated section, which, along with fan installed on the back panel, enables increased air circulation. The back panel of the base part contains input power connector with power switch and fuse. Inside of the base part includes high-power transformer, main circuit board and auxiliary power supply board. Main circuit board holds most of the circuits including thyristors, microcontroller, gate-drive transformer and relay, which connects the circuitry to the input voltage. Auxiliary power supply board provides low voltage source, needed for proper operation.



**Figure 5:** 3D Model of the prototype

## 6. CONCLUSION

This article describes a prototype of spot welding machine created for assembly of battery packs using single Li-ion cells. Theoretical knowledge of concepts was applied to construct diagrams and algorithm for current controller. Constructed device provides pulsed current at the output with option to control peak current. The device creates joints with low electric resistance while not exposing battery to excessive heat stress and therefore preventing unnecessary power losses.

## REFERENCES

- [1] M. J. Brand, P. A. Schmidt, M. F. Zaeh and A. Jossen, "Welding techniques for battery cells and resulting electrical contact resistances," *Journal of Energy Storage*, vol. 1, no. 2015, p. 7–14, <https://doi.org/10.1016/j.est.2015.04.001>
- [2] S. S. Lee, T. H. Kim, S. J. Hu, W. W. Cai and J. A. Abell, "Joining Technologies for Automotive Lithium-Ion Battery Manufacturing: A Review," ASME 2010 International Manufacturing Science and Engineering Conference, 2010, pp. 541-549, <https://doi.org/10.1115/MSEC2010-34168>

# Community energy for the municipality with population 3 000

P. Krejčířík

Faculty of Electrical Engineering and Communication, Brno University of Technology,  
Czech Republic

E-mail: xkrejc67@vutbr.cz

**Abstract**— The article is focused on community energy and the possibility of use in small communities. It also presents a municipality located in the Zlín region in the Czech Republic and outlines possible community energy proposals for the municipality of Bánov.

**Keywords**— Community energy, community opportunities in small municipality

Nowadays, in the Czech Republic, but also in the world, there is an effort to produce electricity mainly from renewable sources and thus to achieve less carbon dioxide production into the air. One possibility is to bring customers together in communities to reduce greenhouse gas emissions by generating electricity from renewable sources. Projects can help decarbonise the energy system by having communities generate low-carbon electricity.

There is untapped potential in many communities in the form of free space for possible power generation plants such as the municipality of Bánov. It is a small municipality in the Zlín region, in the Uherské Hradiště district, with an altitude of 287 m above sea level. The municipality has 2 103 inhabitants and 673 houses. The entire municipality is fully supplied with drinking water, using a group water supply system. Sewage is discharged by sewerage to the local sewage treatment plant. The sewerage network has undergone partial repair during the construction of the sewage treatment plant. Gas is supplied via a of high- pressure pipeline and electricity via a 22 kV HV network.

There is also a Primary School in the municipality, the school has all 9 years of primary attendance and is a catchment school for children from surrounding villages. The primary school is preceded by a kindergarten in the municipality of Bánov with a capacity of 120 places and its own canteen. Health care in the municipality is provided by a health centre with two general practitioners and two dentists.  
[1]

There would be several options for the community, these would be options where buildings that are in a small area in the center of the community would be considered initially, followed by buildings that are out of the core area but still owned by the municipality. In the last proposal, selected industry would be involved in the community.

## 1. CORE AREA

One option for community energy could be to engage with municipality buildings that are in proximity, in the center of the municipality. In this case it would be mainly the primary and kindergarten and the health center, possibly also the sports complex. (Points 1-4 in Figure 1.) The connection between primary school, kindergarten, sports complex, and health center should be not problematic, as all buildings are in the immediate vicinity.

## 2. REMOTE BUILDINGS OWNED BY THE MUNICIPALITY

In addition, any buildings owned by the municipality would also be considered for inclusion in the community. to the buildings considered in the first proposal would be added the municipal office, post office and fire station. There are also two buildings owned by the municipality that serve as apartment buildings and are located in a relatively in close location for connection within the network. These are a Residential buildings with a capacity of 20 flats (number 8 in figure 1) and with a capacity of 5 flats (number 7).

## 3. PRIVATELY OWNED BUILDINGS

However, the community does not have to consist only of municipal buildings; selected industry in the municipality of Bánov can be considered as part of community in the future. For example, the joinery workshop located near the municipal office. Which would make it easy to connect to the local distribution network. The whole building also has a relatively large roof area with suitable orientation. The future scenario also can consider expanding to other citizens of the municipality of Bánov, but this case is not evaluated I the article.



**Figure 1:** Distribution of buildings in the municipality of Bánov



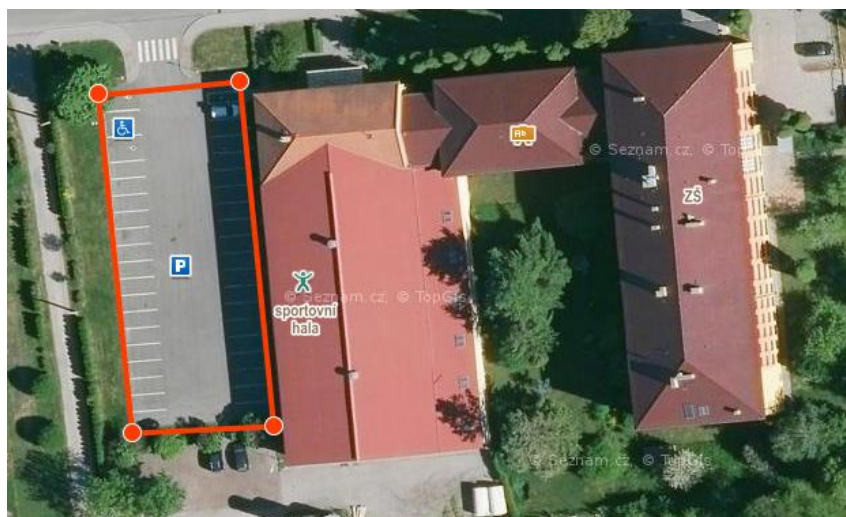
#### 4. CONNECTING THE WASTEWATER TREATMENT PLANT

There is a question if a sewage treatment is appropriate to include in the community, as the sewage treatment plant is located at a relatively long distance from the nearest community site and a long distribution network would have to be in case of local distribution network owner by municipality.

The roof of the building could be fitted with photovoltaic panels to ensure a sufficient supply of electricity and there is a possibility to use the space around the buildings and to place photovoltaic panels there as well, the area in question is approximately 250 m<sup>2</sup> according to the maps. As another option for the wastewater treatment plant, are looking at community connection, but the building would not be connected to the community via the local distribution network, but via the public network. However, the connection would depend on the arrangement with the distributor.

#### 5. ROOFING THE CAR PARK WITH PHOTOVOLTAIC PANELS

Near the primary school and the sports complex, there is a car park that could be roofed with photovoltaic panels. According to map measurements, the car park has an approximate area of 1 000 m<sup>2</sup>. The photovoltaic panels would provide energy to power the adjacent buildings and it would also be possible to build charging stations for electric vehicles. The building would also provide with protection from the sun on hot summer days in the car park. On the other hand, one of the disadvantages would be the location of the car park, which is adjacent to the sports complex building, and in the morning the building could shade the panels and reduce the efficiency of the PV panels.



**Figure 2:** Marked parking lot at the sports complex building.

Photovoltaic panels are useful for ensuring a sufficient supply of electricity during the summer months when there is huge amount of sunlight. Conversely, during the winter, the supply from PV panels might not be adequate.

#### 6. ROOF AREA SUITABLE FOR PHOTOVOLTAIC PANELS

To find out the approximate production of a photovoltaic power plant, it is necessary to know the area of the roof. The table below shows the calculated values of the roof area of the different buildings that should be involved in the community. The orientation of the roofs is also given. Using the PVGIS follow-up program, the production of PV plants located on the roofs of the buildings was calculated.



Buildings	Roof area [m <sup>2</sup> ]	Roof orientation	Electricity produced [kWh]
Primary School – east	444	east	79009
Primary School – south	81	south	16962
Sports Complex	518	east	91643
Health center	282	south	61142
Kindergarten	134	south	29053
Municipal Office	226	south	49000
Joinery workshop	203	south	44013
Residential building 8	650	east/west	59618
Sewage treatment plant	250	south	54200
Car park	1000	south	216816
Post office	39	south	8455

**Table1:** Roof area, orientation and electricity produced

## 7. CONCLUSION

The article presents possible proposals of communities in the municipality and also approximate areas of roofs with potential power production from PV panels was evaluated. The total roof area was calculated at 3 788 m<sup>2</sup>, of which 1 612 m<sup>2</sup> was oriented to the east and the remaining 2 176 m<sup>2</sup> was oriented to the south. Using the PVGIS software, the total value of the electricity produced was calculated to be 709 911 kWh. The article reflects the initial level of analysis focused on the energy flows in the distribution network affected by the designed community, further evaluations will be part of future research.

## 8. REFERENCES

- [1] *Obec Bánov* [online]. Bánov: Obec Bánov, 2019 [cit. 2022-01-05]. Available from: [https://www.banov.cz/download-file/PRO\\_Banov\\_final.pdf](https://www.banov.cz/download-file/PRO_Banov_final.pdf)
- [2] Obec Bánov. *Mapy.cz* [online]. Praha: ., 2021 [cit. 2022-03-13]. Available from: <https://mapy.cz/zakladni?x=17.7288598&y=48.9738422&z=12&source=muni&id=3276&ds=1>
- [3] PVGIS. *PVGIS* [online]. Brusel: ., 2020 [cit. 2022-03-26]. Available from: [https://re.jrc.ec.europa.eu/pvg\\_tools/en/#api\\_5.2](https://re.jrc.ec.europa.eu/pvg_tools/en/#api_5.2)
- [4] *Are Small Islands Feasible Prototypes for Future Smart Grids?* [online]. .: IEEE, 2019 [cit. 2022-03-26]. Available from: <https://smartgrid.ieee.org/bulletins/october-2019/are-small-islands-feasible-prototypes-for-future-smart-grids>



# PROČ ČEKAT NA BUDOUCNOST?

Vytvářej budoucnost společně s námi!  
V největším R&D centru Honeywellu v Evropě  
nabízíme studentské stáže, part-time či  
absolventské pozice.

Registruj se pomocí QR kódu a o nových  
nabídkách se dozvíš jako první!



**Honeywell**



# **PŘÍLEŽITOSTI PRO STUDENTY A ABSOLVENTY**

V Honeywellu věříme, že budoucnost je taková, jakou ji společnými silami vytvoříme. Zaměřujeme se na oblasti Aerospace a Safety & Productivity Solutions.

## **KOHO HLEDÁME?**

- Studenty a absolventy všech oborů FEKT VUT
- Zájem o dlouhodobou spolupráci a ochota dále rozvíjet znalosti z oboru
- Stáže: možnost pracovat 16–20 hodin týdně (flexibilní směny)
- Komunikativní znalost angličtiny

## **PŘÍKLADY POZIC**

- Electrical/Electronics Engineer
- Hardware Engineer
- PLC Controls Engineer
- Systems Engineer
- Prototype/Test Technician
- Technical Support



[www.honeywell.jobs.cz](http://www.honeywell.jobs.cz)

[careers.honeywell.com/brno](http://careers.honeywell.com/brno)

e-mail: [Martina.Hanusova@Honeywell.com](mailto:Martina.Hanusova@Honeywell.com)

# McRuer Models for Human-Machine Systems

J. Čelko and O. Mihálik

Department of Control and Instrumentation, Brno University of Technology, 61600 Brno, Czech Republic.

E-mail: [216785@vut.cz](mailto:216785@vut.cz), [173702@vut.cz](mailto:173702@vut.cz)

**Abstract**—This article is focused on modeling human-machine systems using McRuer models, particularly systems formed by human operator and a steering wheel. Result of the work are dynamical models of human operator corresponding to regulation of different systems and various system forcing functions. Different steering wheel stiffness was also measured to find its influence on the speed of human reactions.

**Keywords**—human-machine systems, human operator, McRuer models, steering wheel, stiffness

## 1. INTRODUCTION

According to [1], there are 3 main methods of modeling human operator. First are models based on control theory (McRuer models also belong here), second are models based on human psychology and finally models based on intelligence technologies. Within this paper, human operator will be described by control theory models as a linear time-invariable system. These models allow us to predict human behaviour.

## 2. MCRUER MODELS

McRuer models can be explained on a plain regulation loop shown on Figure 1, where  $F_R$  is the model of human operator and  $F_S$  is the regulated system.

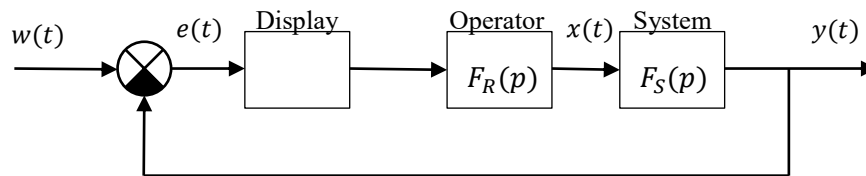


Figure 1: Regulation loop [2]

The principle of the McRuer models can be simply described as a process, where the human operator adapts and changes his frequency response in order to achieve transfer function

$$F_0(p) = F_R(p)F_S(p) = \frac{K_0}{p} \cdot e^{-\tau p}, \quad (1)$$

in the direct branch of the loop [2].

## 3. SIMULATOR AND EXPERIMENT

All measurements were taken with the vehicle driving simulator shown on Figure 2. The simulator was equipped with Logitech G920 steering wheel, through which the human operator was interacting with the system. Processing of the operator action (the rotation of the steering wheel in time), as well as implementation of the controlled system and indicator of the regulation error were realized in MATLAB Simulink environment. Before a measurement, operator tried the regulation of the given scenario a few times to get used to it. Subsequently 10 repeated measurements were made, from which an average

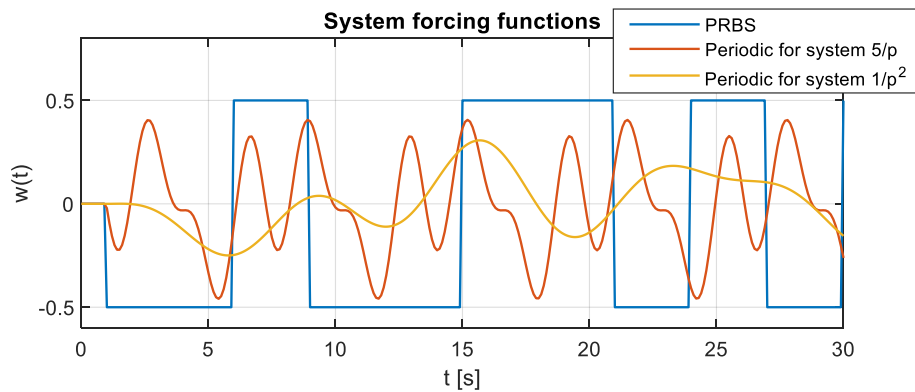
operator model was estimated. All experiments were taken only with one operator though.



**Figure 2:** Vehicle driving simulator [3] and regulation error indicator

#### 4. TESTED SYSTEMS AND FORCING FUNCTIONS

Two forcing functions were chosen for identification of the human operator: A *pseudo random binary sequence* (PRBS) and a periodic signal, in which several sin functions were combined. To allow the operator to prepare for the regulation, all forcing functions were multiplied by zero in the first second of the measurement. All measurements had the same length of 60 seconds. An example of the PRBS and periodic forcing functions are shown in Figure 3.



**Figure 3:** Forcing functions used for the measurements

**Table 1:** Tested systems

General transfer function	Particular transfer function
$K_S$	10
$\frac{K_S}{p}$	$\frac{5}{p}$
$\frac{K_S}{Tp + 1}$	$\frac{10}{p + 1}$
$\frac{K_S}{p^2}$	$\frac{1}{p^2}$

Operator's responses were tested with the transfer functions shown in the Table 1. All systems were measured with the PRBS forcing function. In addition, the  $5/p$  and  $1/p^2$  systems were also measured with a harmonic forcing functions. Furthermore, the  $1/p^2$  system was measured with two different stiffnesses of the steering wheel: with one that comes with the steering wheel by default and with no stiffness at all.

## 5. IDENTIFIED OPERATOR TRANSFER FUNCTIONS

For every scenario, the optimal transfer function describing the operator was found, i.e., the simplest mathematical description that could simulate operator's actions accurately enough. Table 2 lists the models of operator for various systems. These were obtained with default steering wheel stiffness. Responses to the system  $1/p^2$  (measured under different conditions) are shown in the Table 3.

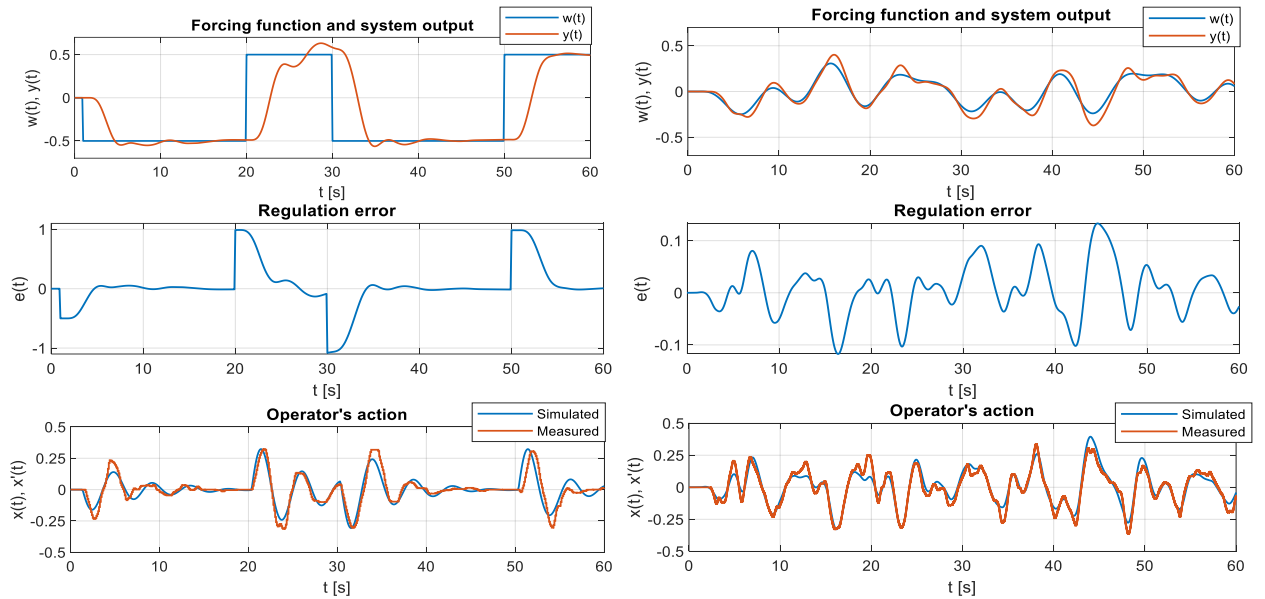
**Table 2:** Identified operator transfer functions

Controlled system	System forcing function	Identified operator transfer function	Theoretical operator transfer function [1]
10	PRBS	$\frac{0.1391}{p} \cdot e^{-0.325p}$	$\frac{K_R}{p} \cdot e^{-\tau p}$
$\frac{5}{p}$	PRBS	$0.1547 \cdot e^{-0.58p}$	$K_R \cdot e^{-\tau p}$
$\frac{10}{p+1}$	periodic	$0.3853 \cdot e^{-0.08p}$	
	PRBS	$\frac{0.08(1.178p+1)}{p} \cdot e^{-0.595p}$	$\frac{K_R(T_1p+1)}{p} \cdot e^{-\tau p}$

**Table 3:** Identified operator transfer functions for  $1/p^2$  system

Steering wheel stiffness	System forcing function	Identified operator transfer function (general)	Identified operator transfer function (particular)
Default	PRBS	$\frac{K_R(T_1p+1)}{T^2p^2+2\xi Tp+1} \cdot e^{-\tau p}$	$\frac{0.0213(17.3p+1)}{0.5939p^2+2 \cdot 0.269 \cdot 0.5939p+1} \cdot e^{-0.66p}$
	periodic	$\frac{K_R(T_1p+1)}{T_2p+1} \cdot e^{-\tau p}$	$\frac{1.1999(2.461p+1)}{0.2899p+1} \cdot e^{-0.325p}$
None	PRBS	$\frac{K_R(T_1p+1)}{T^2p^2+2\xi Tp+1} \cdot e^{-\tau p}$	$\frac{0.0338(11.5565p+1)}{0.5019p^2+2 \cdot 0.2398 \cdot 0.5019p+1} \cdot e^{-0.675p}$
	periodic	$\frac{K_R(T_1p+1)}{T_2p+1} \cdot e^{-\tau p}$	$\frac{0.8279(3.1383p+1)}{0.3162p+1} \cdot e^{-0.415p}$

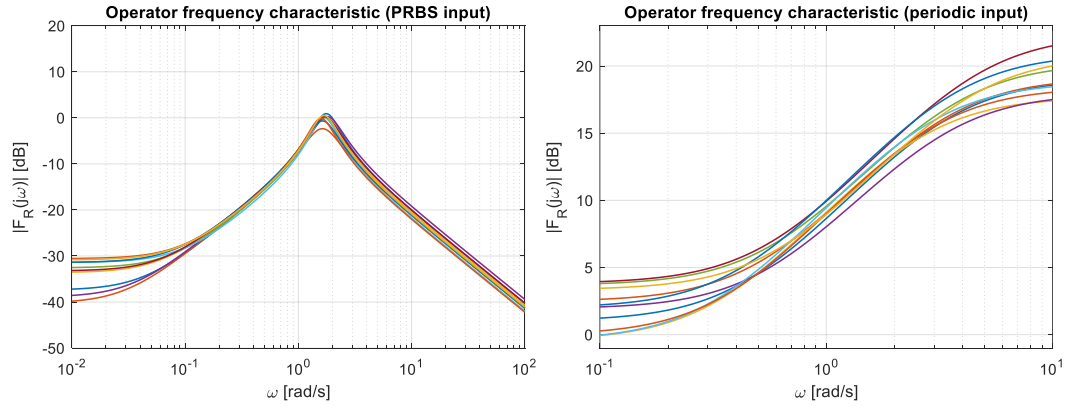
Measured and estimated data from one measurement for the system  $1/p^2$  are shown on Figure 4. Operator frequency characteristics for every measurement with this system are shown on Figure 5.



**Figure 4:** Example of measured and estimated data,  $F_S(p) = 1/p^2$  with PRBS and harmonic functions



Transport delays couldn't be identified with high accuracy. For example, with PRBS forcing function, the `tfest` function used for identification gave similar percent fit to estimation data (about 50–60%) for transport delays spreading from 0.3 to 0.7 seconds. Direct estimation of reaction delay by the `tfest` function was not reliable. Therefore, model was identified with several delays and the model with the smallest error of the fit was used.



**Figure 5:** Operator frequency characteristics when regulating system  $1/p^2$

## 6. CONCLUSION

For the simpler systems ( $K_S, K_S/p, K_S/(p + 1)$ ) observed operator responses were equal to McRuer theoretical assumptions. However, for the  $1/p^2$  system the responses were different. That could be caused by the fact, that the theoretical regulator is  $F_R(p) = K_R p \cdot e^{-\tau p}$ , which has no real-world equivalent.

Another discovery is that the optimal transfer function of human operator differs depending on the system forcing functions. For the  $5/p$  system only the numerical value of coefficients differs, whilst there was a completely different formula in the case of the  $1/p^2$  system.

With the different steering wheel stiffness experiment, another interesting result came out. There was a significant drop of the coefficients with the sinusoidal input and no stiffness. However, with the PRBS input, only a slight change of coefficients values was observed.

In general, identified operator transfer functions comply with McRuer models theory. However, individual coefficients are dependent on the system forcing function and the steering wheel stiffness. Next aim will be to estimate optimal wheel stiffness for given scenarios to achieve faster driver reactions and ensure higher safety in traffic.

## ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-20-6205 — “Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0” financially supported by the Internal science fund of Brno University of Technology.

## REFERENCES

- [1] Shuting Xu et. al. “Review of control models for human pilot behavior,” *Annual Reviews in Control*, Volume 44, 2017, Pages 274-291, [cit. 2022-3-4] ISSN 1367-5788, <https://doi.org/10.1016/j.arcontrol.2017.09.009>.
- [2] McRuer, Duane T. and Ezra S. Krendel. Mathematical models of human pilot behavior [online]. Neuilly sur Seine: North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development, 1974 [cit.2021-12-29]. <https://www.sto.nato.int/publications/AGARD/AGARD-AG-188/AGARD-AG-188.pdf>
- [3] Michalik, D., O. Mihalik, M. Jirgl a P. Fiedler. “Driver Behaviour Modeling With Vehicle Driving Simulator,” *IFAC-PapersOnLine*, Volume 52, Issue 272, 2019, Pages 180-185 [cit. 2022-3-3]. ISSN 24058963, <https://doi.org/10.1016/j.ifacol.2019.12.753>.

# Analysis of Car Drivers' Behaviour and Driving Style

S. Novik and O. Mihálik

Department of Control and Instrumentation, Brno University of Technology, 61600 Brno, Czech Republic.

E-mail: [186148@vut.cz](mailto:186148@vut.cz), (S.N.) [173702@vut.cz](mailto:173702@vut.cz) (O.M.)

**Abstract:** Driving security remains one of the important issues. Nowadays, various assistance systems are implemented, such as the systems for analysis of control of a car by its driver. To understand the performance of the driver's control, a program was created to obtain valuable data and relevant characteristics. To obtain the data, we used an internally designed, laboratory-made vehicle driving simulator developed by D. Michalík [2]. Driver data were obtained using a proprietary vehicle driving simulator, and these were evaluated in the MATLAB environment via integral criteria and other calculated parameters, such as reaction delay. Features thus obtained were used as a training set for the machine learning, using LDA and QDA methods (linear and quadratic discriminant analysis). These methods reveal information concerning the importance of features for the task of driver's identity prediction based solely on the driving actions.

**Keywords**—Feedback control, driver analysis, integral criteria of quality, machine learning, MATLAB, vehicle driving simulator.

## 1. INTRODUCTION

Vehicle driving simulators come as an attractive alternative to a real-life driving, because they are both less expensive and driver can execute tests in safety. Simulators are an effective alternative in providing safe testing environment and simulate scenarios which could be way harder to find in real world. It is also a great tool to gather and measure various values from simulator.

Analysis of drivers was the subject of interest in this bachelor's thesis. Our aim was to investigate the possibility to classify the identity of a driver. The simulator developed by D. Michalík [2] was used to gather time series of steering wheel angle, pedal pressure, and car position. These were used to compute features (such as reaction delay and integral quality criteria) for classification of drivers' identity using artificial intelligence. In this process we also obtained information which features are more valuable than the others.

Section 2 covers the theoretical connection between human and simulator as a feedback loop. Also, a brief description of used tools taken from theory of control and regulation will follow. In Section 3, Block diagrams are provided to help understand and visually show how the flow of algorithm programmed in MATLAB is gathering, processing, and computing useful data. Then these data will used as input test set in artificial intelligence, which will help us determine which values can be used for identification of a driver. Section 4 discusses retrieved results and what was achieved in this work.

## 2. PRELIMINARIES

We can describe driver and simulator as a feedback scheme [1], where the driver is block  $F_R(p)$  and simulator is  $F_S(p)$ . It is a purposeful action on an object and the goal is to achieve a predetermined state.

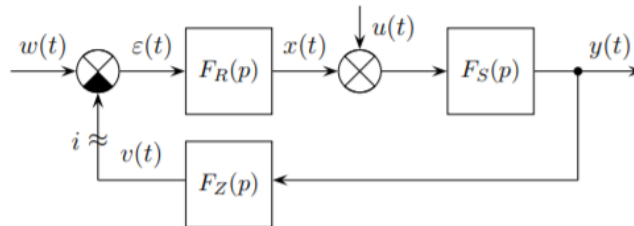


Figure 1: Simplified feedback block scheme

A general block diagram of feedback control is shown in **Figure 1**, where the quantity  $e(t)$  is the control deviation, which is defined as the difference between the desired value  $w(t)$  and the output value  $y(t)$ . The quantity  $e(t)$  characterizes the distance of the car from the centre of the desired lane. The quantity  $x(t)$  is called the action, in this case the steering wheel angle. Here it is the case of control where the attempt is to reduce the value of the deviation as fast as possible. As the main tool for evaluating the speed and quality of the control, we use the integral quality criteria (described below). Apart from these criteria, other parameters have been used which will be subsequently used as driver characteristics. In our case, we can use a simplified scheme where the feedback transfer  $F_Z(p)$  is equal to one.

As suggested by [2], a human, or in this case, a driver, can be described by transfer function:

$$F(p) = \frac{k \cdot p}{T^2 \cdot p^2 + 2 \cdot T \cdot \xi \cdot p + 1} e^{-p\tau} \quad (1)$$

where the parameters are time constant  $T$ , gain  $k$  and damping  $\xi$ . They are found by the MATLAB identification algorithm *tfest* (transfer function estimation).

### Integral criterion

Integral criteria characterize the quality of the regulatory process on an infinite time horizon. There are several of them. The choice of the appropriate criterion is therefore largely related to the type of process being controlled. It is well known that one of the least favorable cases that a control circuit must cope with are step changes in the signals entering it. Therefore, as a rule, we investigate the behavior of the control circuit under step changes in the desired value of the output (reference signal) or disturbance. If the control path is satisfactory for step changes in the input signals, we can expect it to be satisfactory for other limited signals that do not have a jump shape. Furthermore, it should be noted that even if the problem is implemented in finite time, theoretically the control takes place on an infinite time horizon, and some criteria are based on this theoretical assumption.

When designing controllers, we try to keep the individual integrals to minimum values. The lower the value, the better the regulatory process was executed. The common integrate criterions are of the form

$$J_{m,n} = \int_0^\infty t^m |e(t)|^n dt = \int_0^\infty t^m |w(t) - y(t)|^n dt. \quad (2)$$

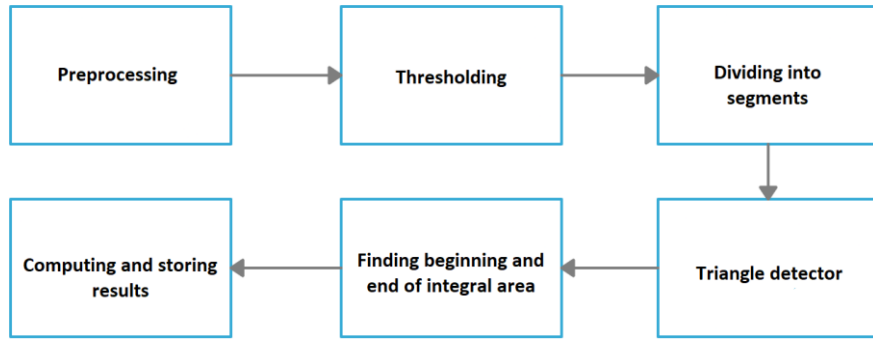
where  $m$  and  $n$  are constants selected by the user. We confined our analysis to the absolute criterion  $J_{0,1}$  and quadratic criterion  $J_{0,2}$ . The ITAE criterion  $J_{1,1}$  was intentionally avoided, chiefly due to its time-weighting factor  $t$ , which tended to amplify the noise present after the main part of the feedback loop step response.

### Driver's identity classification

As a tool for partitioning the feature space data (parameters such as integral criteria, reaction delay, etc.), we use the so-called quadratic discriminant analysis (QDA) [3], one of the methods of multivariate statistical analysis that tries to discriminate objects coming from a finite number of classes (categories) based on a given training set. In this case, the training set contains the extracted parameters, and the classes are the labels of the drivers (i.e., which parameters belong to a certain driver). The matrix from the given parameters is given below. All values were stored in a text file, and the necessary values were then loaded into MATLAB (each file represents one run performed). The processed parameters were stored in the matrix from all the runs performed. The output is a column vector showing which row belongs to a particular driver (class labels for the calculation). The coefficient  $n$  in the matrix represents the total number of rides uploaded to the program for calculation.

## 3. ALGORITHM

To convey the process of data processing, individual steps are illustrated by the flowchart in **Figure 2**. At the very beginning, discrete signals were obtained and sent from the simulator to MATLAB. First, the signals pass through a preprocessing stage which removes inappropriate signal detections (mostly start and end signals that capture the position of the steering wheel at the beginning and at the end, which can interfere with the subsequent calculation). This ensures correct calculation of the criteria and additional parameters. Subsequently a thresholding step follows, whereby signal peaks are detected.

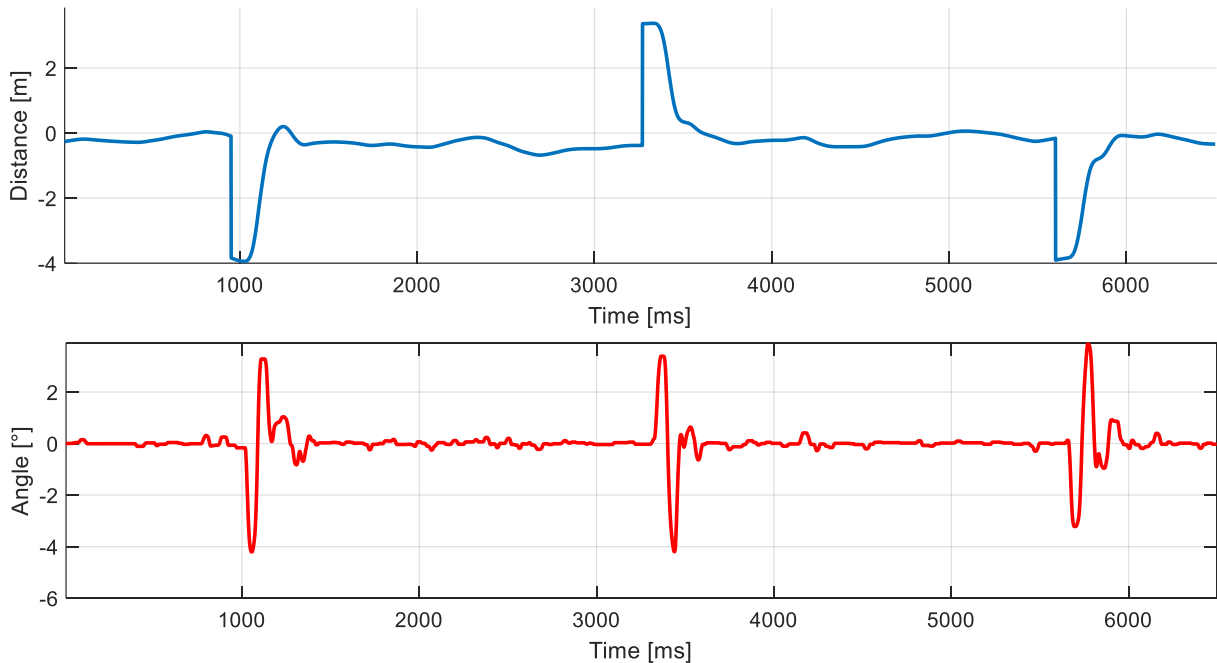


**Figure 2:** Block diagram of algorithm

The processed signal is then divided into segments (each segment contains a detected peak) and these are then sent sequentially to the triangle detector, which pinpoints the start and end of the peaks to calculate the integral criteria more accurately from the area. Along with the criteria, the reaction time, rise time and period of the peak are also calculated.

### Measured signals

The task of a driver is to change lane according to visual signal (green arrow). This corresponds to a measurement of a step response. Main two signals that were taken from the simulator are wheel angle and distance from the middle of a driving lane. The signals were then used for calculations of integral criteria and time values (reaction time, time of ascent...). **Figure 3** of the first plot shows when the signal to change lane was given, that's why it is the biggest distance from the middle lane, and below by the movement of driving wheel we can observe and compute how good was drivers reaction, how fast the lane was changed etc.



**Figure 3:** Graphs of measured signals

### Extracted features

Collected results were put into an input test set (which took form of a matrix, where all results were stored in rows and columns). This set was used by artificial intelligence and as a result, we got confusion matrix. A confusion matrix is a summary of prediction results on a classification problem. The number of correct and incorrect predictions are summarized with count values and broken down by each class.

Actual Values	2	18	3		1
	5		18	4	
	69		5	16	1
	74	1	5	1	15
		2	5	69	74
		Predicted values			

**Figure 4:** Confusion matrix of values

The columns represent the actual value of the predicted feature and the rows represent the classifier prediction. The cells of the matrix contain the frequencies of how many times given combination of actual and predicted values occurred on the training set under study. Cases on the diagonal of the confusion matrix are classified correctly, off-diagonal cases are errors. The numbers below represent each driver.

#### 4. CONCLUSION

A total of 4 driving sessions of different drivers were observed, where data were collected and then parameters were calculated from these data to describe their driving quality. Since performing test drives can carry considerable risk, the driving simulations were performed on a simulator developed at the Institute of Automation and Measurement Technology of the BUT. Control deviation signals and action variables were acquired. From all obtained and computed values it was determined that most useful data for driver identification were absolute and quadratic criterions, plus to these values were added time values such as react and rising time of a driver. The accuracy of used values as a input train set for QDA method is 76,1%.

#### ACKNOWLEDGEMENT

The completion of this paper was made possible by the grant No. FEKT-S-20-6205— “Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0” financially supported by the Internal science fund of Brno University of Technology.

#### REFERENCES

- [1] P. BLÁHA, P. VAVŘÍN. *Řízení a regulace I*. VUT v Brně, 2005.
- [2] MICHALÍK, David and JIRGL, Miroslav and ARM, Jakub and FIEDLER, Petr. *Developing an Unreal Engine 4-Based Vehicle Driving Simulator Applicable in Driver Behavior Analysis—A Technical Perspective*. VUT v Brně, 2021.  
Available from: [URL:<https://www.mdpi.com/2313-576X/7/2/25>](https://www.mdpi.com/2313-576X/7/2/25)
- [3] JAMES, Gareth, Daniela WITTEN, Trevor HASTIE a Robert TIBSHIRANI. *An introduction to statistical learning: with applications in R*. New York: Springer, 2013. Springer texts in statistics, 103. ISBN 978-1-4614-7138-7.

# System for verification of the signal quality dependence of a laser vibrometer

O.Skalský, S.Pikula

Brno University of Technology, FEEC, Department of Control and Instrumentation, Czech republic

E-mail: [xskals03@vut.cz](mailto:xskals03@vut.cz), [pikula@vut.cz](mailto:pikula@vut.cz)

**Abstract**— The paper describes impact of a laser vibration measurement parameters on the signal quality of vibration measurement using Laser Doppler Vibrometer (LDV). Among these parameters are distance, perpendicularity, material, roughness or polarization properties of the measured surface. Follows the presentation of the measurement system that will be able to measure the quality of the LDV signal affected by these influences for a specific surface, at different distances and angles. Based on the characteristics obtained, it will be possible to assess the suitability of the measured surface for LDV use.

**Keywords** — Laser Doppler Vibrometer (LDV), He-Ne Laser, LDV Signal Quality Dependence, Measurement System

## 1. INTRODUCTION

Vibrometry is a discipline that deals with the measurement of vibrations and their description using motion quantities. It is mainly used in industrial production, machine diagnostics, research and development and construction. Thanks to the current trend, laser Doppler vibrometers (LDV) are also becoming more affordable and their range of applications is expanding. Compared to conventional vibrometers, the accuracy of LDVs is orders of magnitude higher. It is also a non-contact measurement where the laser beam has virtually no effect on the object to be measured, which can be several metres away [1].

However, LDV measurements can be negatively affected by several factors and, if the measurement conditions are not chosen appropriately, it cannot be ensured that the measured data will always be valid. LDV manufacturers generally do not provide sufficient information about these parasitic effects and the customer cannot be sure in advance whether the LDV will be suitable for the application [1][2].

The aim of this work is to reveal and describe the most significant parasitic phenomena, to implement a measurement apparatus to verify the dependence of the LDV signal quality on these parameters and thus to create a measurement system to obtain additional characteristics about the particular LDVs and types of measured surfaces. Based on this knowledge, it will be possible to assess the suitability of the LDV for a particular application.

## 2. LASER DOPPLER VIBROMETER – SIGNAL QUALITY DEPENDENCE

Although vibration measurement with LDV is very accurate and sophisticated, the measured data can be completely invalidated due to unsuitable measurement conditions. There are several phenomena affecting the signal quality of the measurement. We will describe four most crucial parameters, which we will focus on further.

### **Reflectivity and surface roughness of the measured object**

The influence of material type and surface processing is an obvious factor affecting the quality of the output signal. For very glossy surfaces with little roughness, the quality is very susceptible to non-perpendicularity of the beam and the surface to be measured. On the other hand, for rough surfaces that disperse the light greatly, strong attenuation occurs and measurement over longer distances may not be possible. The laser head on such surfaces may also not detect the distance of the object correctly, the focus length is incorrectly set and autofocus may fail [1][3].

### **The angle of the beam to the surface of the measured object**

Another noticeable influence on the measurement quality is the variety of the angle between the laser beam and the surface of the material to be measured. As the normal vector of the plane deviates from



the laser beam, the signal quality of course decreases. The sensitivity of the measurement to changes in this angle is mainly determined by the surface to be measured [1][3].

### **Polarization properties of the surface of the measured object**

In order to avoid significant or complete absorption of the beam due to possible linear polarization properties of the measured surface, the laser beam of the sensor head is polarized in a circular orientation. The influence of the polarization properties of the measurement object can still have a negative effect on the measurement [3].

### **Destructive interference at certain distances from the measured object**

In the laser cavity of a Fabry-Perot resonator, standing waves can occur only at wavelengths  $\lambda_n$

$$\lambda_n = \frac{l}{n} \text{ where } n \in \mathbb{N} \quad (1)$$

where  $l$  is the optical length of the laser resonator cavity. The radiative transition of an electron to a lower energy level due to spontaneous emission emits a photon with a wavelength corresponding to the energy difference of this electron. Since the energy levels do not define the exact energy of the electron, but only the energy of the most probable state of the electron in a given energy level, we cannot expect that the emitted photon will always have exactly the same wavelength. Resonance of several waves of non-negligible intensities with wavelengths  $\lambda_k$  to  $\lambda_{k+i}$  can occur in a laser cavity, where  $i$  usually takes at most units. At distances  $s$

$$s = l \cdot \left( \frac{1}{2} + k \right) \text{ where } k \in \mathbb{N}_0 \quad (2)$$

destructive interference of each pair of waves with wavelengths  $\lambda_n$  and  $\lambda_{n+1}$  occurs. If the object to be measured is at one of the unsuitable distances  $s$ , measurement may not be possible [1][3][3].

## **3. DESIGN OF THE MEASUREMENT SYSTEM**

Based on the above described effects on signal quality, a measurement system was designed to measure the dependence of signal quality on these parameters at variable distance and angle of the measured surface.

Let us state upfront that the task of the measurement system is to measure the signal quality for a specific surface for different distances and angles. Based on the measured characteristics, it is then possible to assess the suitability of a given surface in a particular application.

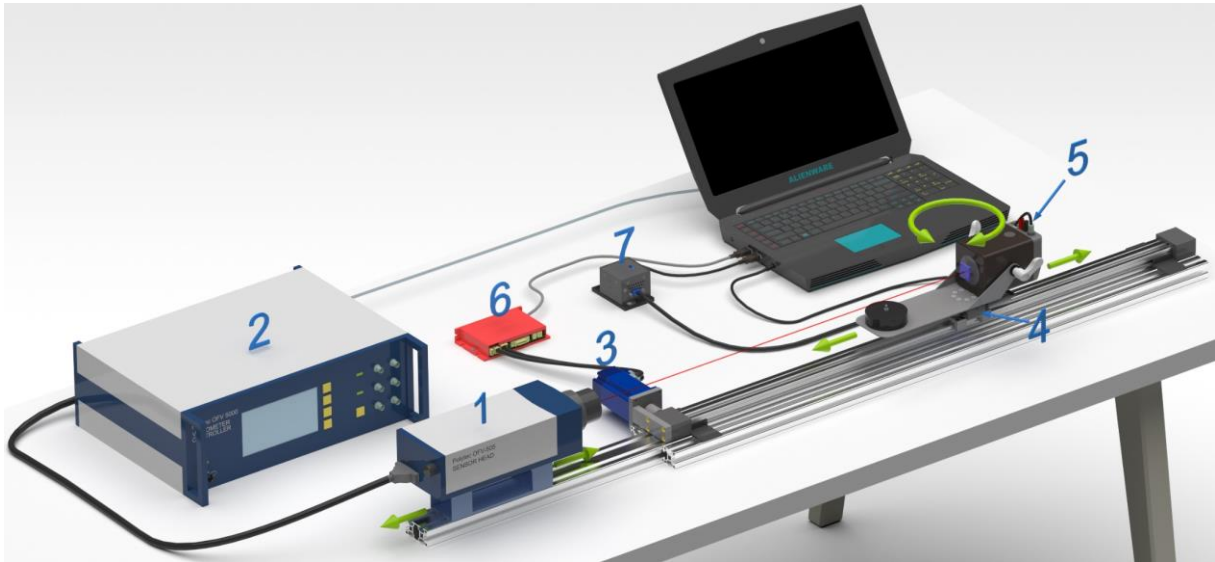
Parasitic influences affect the measurement simultaneously and degrade the signal quality. The system will not be able to evaluate these effects. However, based on the signal quality drops, we can identify the most significant of these parasites through knowledge of their characteristic behaviour. We can therefore more easily eliminate parasitic influences and take the appropriate steps for measurement suitability.

### **Hardware – description of the measuring instrumentation**

Figure 1 shows the complete measuring apparatus. The basis is a toothed belt linear actuator for changing the distance of the surface to be measured, on which a rotary table for angle control is placed. The vibrating table is mounted on the rotary table in such a way that the path of the beam remains constant as the table rotates. To the vibrating table is attached the measured surface, that can easily be replaced. The frequency and amplitude of the vibrating table does not affect the measurement. However, it is necessary to know the frequency for Fourier analysis of the vibrometer output signal.

**Table 1:** Annotations for Figure 1 and Figure 2

1	-	Laser sensor head
2	-	Vibrometer controller
3	-	Stepper motor for linear feed
4	-	Stepper motor for rotary feed
5	-	Vibration table
6	-	Stepper motor driver (linear feed)
7	-	Stepper motor driver (rotary feed)

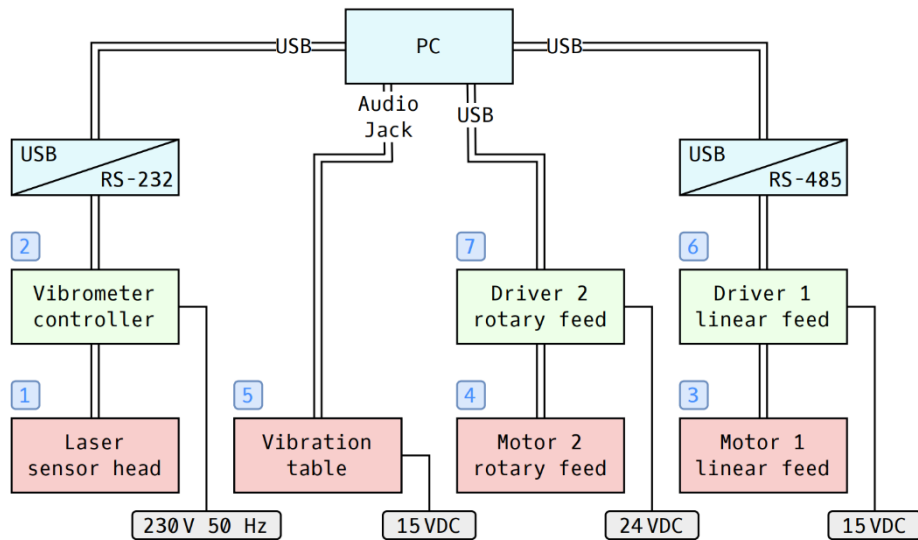


**Figure 1:** Model of measuring system - instrumentation

Current design is prepared for Polytec OFV-505 laser head and Polytec OFV-5000 controller, but in general same hardware arrangement can be used for testing of other LDVs.

The laser head (Polytec OFV-505) is mounted on an adjustable rail for a fixed change of the range of possible measurement intervals. The signal from the sensor head is processed in the Polytec OFV-5000 controller.

The range for automatic distance adjustment is 0 to 1120 mm. For range shift it is possible to manually adjust the offset by moving the rail with laser head. Angle adjustment is limited to a range  $-85^{\circ}$  to  $85^{\circ}$ . With a larger deflection, the beam may not be correctly incident on the surface to be measured.

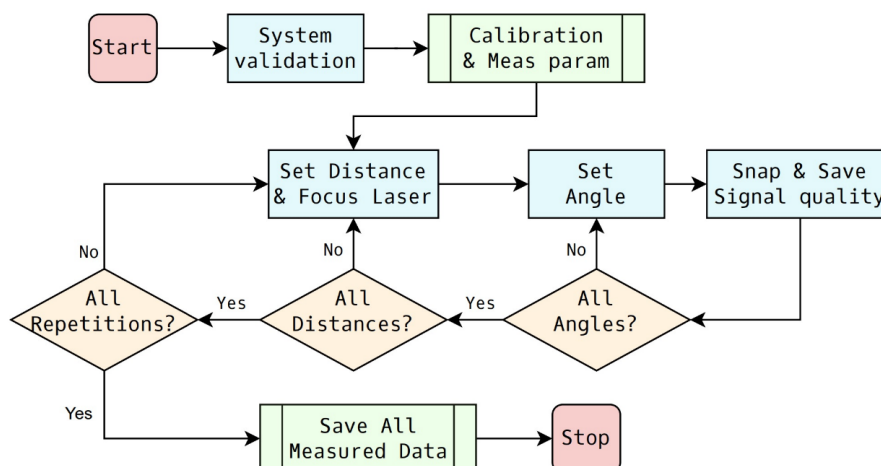


**Figure 2:** Schema of measuring system – connections of instrumentation

As shown in Figure 2, all components (linear and rotary feed, OFV-5000 controller) except the vibration table are connected to and controlled by PC via USB. The vibration table is connected and controlled via audio jack connector.

## Software – description of the system logic

The measurement program developed in LabVIEW allows the user to select a total of 7 basic measurement parameters: start position, end position, position step, start angle, end angle, angle step, number of measurement repeats. Figure 3 shows the flow chart of this software.



**Figure 3:** Flow chart of measurement system software

Before starting the automatic measurement, the user must manually set the linear feed to the stop and the rotary table to zero angle when first switching on. This is necessary as these actuators only have relative positioning. If the extension rail with the laser head is not in the fully retracted position, it is necessary to read and enter the amount of extension.

In the measurement cycle, the signal quality for each distance is measured for all permissible angles. The measurement of all combinations of angles and distances is repeated as many times as the user specifies.

The data are continuously stored during the measurement, so the results cannot be lost if the measurement is abruptly interrupted and the measurement can be continued when the measurement is resumed.

## 4. CONCLUSION

The range of applications of LDV is growing and as these are not inexpensive devices, it is necessary to have a good understanding of all the parameters that affect real-world measurements. The presented system allows to measure wide range of parameters. This can give a much greater insight into the general suitability of LDV deployment for the application in question, as well as the characteristics of the products of individual suppliers.

## REFERENCES

- [1] PAVLOŇ, Martin. *Analýza dynamických parametrů laserových vibrometrů*. Brno, 2017. Diplomová práce. VUT Brno FEKT AMT. Vedoucí práce Zdeněk Havránek.
- [2] RŮŽIČKA, Vladimír. *Kvantová elektronika*. 4. Brno: Vysoké učení technické v Brně, srpen 1987. ISBN 978-3-319-50651-7.
- [3] *Vibrometer Single Point Sensor Head OFV-505/-503: User Manual*. Waldbronn Německo: Polytec, 2004.
- [4] RENK, Karl F. *Basics of Laser Physics: For Students of Science and Engineering*. 2. Springer International Publishing, 2017. ISBN 978-3-319-50651-7. Dostupné z: doi:10.1007/978-3-319-50651-7

# IoT Weather Station

Andrej Pala, Petr Vyroubal

FEEC, Brno University of Technology, Czech Republic

E-mail: xpalaa00@vutbr.cz, vyroubal@vut.cz

**Abstract**—In this article the reader should get to know about new arising project of IoT (Internet of Things) weather station, with its existing or yet developed function, with its structure and construction.

**Keywords**—Wireless weather station, Wi-Fi connectable, Internet of Things, IoT, meteorology, weather measurement, solar powered.

## 1. INTRODUCTION

The reader of this article will certainly object that he has read and seen countless such articles on hand-made meteorological monitoring devices, and that this is just another attempt to discover the wheel.

Of course, there are already countless such devices and designs, but my own design and construction effort was not only to create a device that monitors a few types of quantities and then write data to the embedded storage medium, but to implement a device that monitors most of these meteorological quantities, and by this is as similar as possible to professional weather stations, but at the lowest possible price, or at least a more affordable price for ordinary do-it-yourselfers. It was also an idea to keep the concept as completely freely editable and freely distributable for future commercial and/or non-commercial use by the successors. The design itself also meets the other 2 criteria. First, the weather station is completely wireless. It is powered by solar panels together with batteries and can communicate through a connection to the Wi-Fi network, which secondly allows data storage not only on the embedded data carrier, but also on online storage, from which data can be further processed and evaluated.

The article itself is based on the semester thesis IoT Weather Station Design [1], which I published, where the information is much more specific.

## 2. MEASURED WEATHER QUANTITIES

The weather station, as a device equipped with sensors designed to monitor and collect data on meteorological quantity, observes a large number of these quantities, of which the manufactured equipment is to monitor the following selected: ambient temperature, ground temperature, dew point, ambient atmospheric pressure with conversion to sea level pressure, relative humidity, total rain precipitation, wind speed and direction, length of daylight, intensity of sunlight and intensity of UV radiation with conversion to UV index. In addition, professional meteorological stations monitor, for example, air quality, which includes dust and concentrations of hazardous substances, or cloud state and atmospheric visibility. However, sensors of these quantities tend to be relatively more expensive to purchase, so they were not included in the design.

The temperature is monitored by 2 sensors: by combined air pressure and temperature sensor BMP280, which measures ambient temperature, and by temperature sensor DS18B20, which measures ground temperature. The mentioned dew point can be evaluated by recalculation using Buck's equations [2] from the knowledge of the relative humidity  $RH$  in % and of the ambient temperature  $T$  in °C.

The atmospheric pressure reading is managed by the already mentioned combined air pressure and temperature sensor BMP280, which provides values in hPa. It is possible to use the barometric equation to convert the value of ambient pressure  $p$  to pressure at a height of 0 meters above sea level  $p_0$  [3].

Relative humidity is measured by sensor AM2320, the total rain precipitation by MS-WH-SP-RG, wind speed by anemometer WH-SP-WS01, wind direction by WH-SP-WD with accompanying calibration when unwanted rotation of the device occurs by digital compass HMC5883L, an intensity and length of daylight is read by sensor BH1750. However, it provides the data of the intensity  $E_{lx}$  only in units of lux, to evaluate the length of daylight the data of the intensity  $E_{W/m^2}$  in units of  $W/m^2$  is needed. There is no direct conversion between these units, as each is used to express light intensity considering a different range of the spectrum; in any case, there is an approximation between these units through a study of the relationship between the quantities [4].

The last-mentioned quantities are the intensity of UV radiation and UV index, their measurement is ensured by the ML8511 sensor. The intensity of UV radiation is obtained in units of  $mW/cm^2$ , to obtain the value of the UV index the approximation gained from the manual [5] and the technical sheet of the sensor is used [6].

ESP-12E Interfaces: SPI, UART, I2C, 1-Wire, Wi-Fi			CH340G Interfaces: UART, USB		$\mu$ SD karta Interface: SPI		6V 2W (2x)	Li-Ion 18650 2600h (2x)	DW01A + FS8205A	TP5100	LM2596
IoT			Settings		Data backup		Solar panels	Batteries	Battery protection	Charging module	Voltage converter
Communication							Power supply				
Control unit											
Sensors											
Light			Humidity	Rain precipitation	Wind			Atmospheric pressure	Temperature		
Intensity	UV index				Wind direction		Anemometer		Ambient	Groud	
BH1750 Interface: I2C	ML8511 Interface: Analog	Attiny85 Interface: I2C	AM2320 Interface: I2C	MS-WH-SP-RG Interface: Logical	HMC5883L Interface: I2C	WH-SP-WD Interface: Analog	WH-SP-WS01 Interface: Logical	BMP280 Interface: I2C		DS18B20 Interface: 1-Wire	

**Figure 1:** Block schematic of the weather station design

(Note: All these quantities should be measured at higher placement above the ground, ideally at 2 meters above the ground or higher, except the near ground temperature sensor DS18B20, it should be placed at 5 centimeters above ground.)

### 3. POWER SUPPLY

As already mentioned, the whole device is designed as wireless in the sense that it can operate without the need to wire-up to the device as such additional wires routed outside its construction. The power supply is solved by using of solar panels serving as the primary energy source, batteries serving as a backup secondary energy source. The supply voltage of the system was chosen regarding the needs of individual peripherals, namely the value of 3.3 V. The weather station should be able to withstand an estimated 9 days and 5.3 hours without the need for recharging in the proposed configuration using software techniques that limit energy consumption (estimated average current consumption is 22.9 mA and peak value of the current should not exceed the value of 1054.4 mA).

If it is more specific technical details, then the solar panels are designed 2 connected in parallel with an output voltage of 6 V and a power of 2 W each, or as one integral solar panel with the same output voltage but a power of 4 W. These are connected to a charging module with TP5100 circuit specially adapted for charging 2 lithium battery cells connected in series with a total capacity of 5,000 mAh and with charging current limitation by using of control resistors before the output of the charging module. The batteries are connected to a protection module with a DW01A circuit, which protects them against overcharging, undercharging and short circuit. The batteries with the protection and charging module are further connected to a voltage converter with the LM2596 circuit, which maintains a constant output voltage set by a resistance trimmer at 3.3 V, even in the event of significant ambient temperature fluctuations. The combination of series connection of batteries, where the voltage depending on the battery charge can range from 4.8 to 8.4 V, with a consequent reduction of the voltage to the required 3.3 V allows operability of the weather station in the full range of lithium batteries capacity. This would not be possible with a single battery, as the device also contains components that become inoperable at voltages below 3.3 V (for example, humidity sensor AM2320).

### 4. CONSTRUCTION OF THE WEATHER STATION

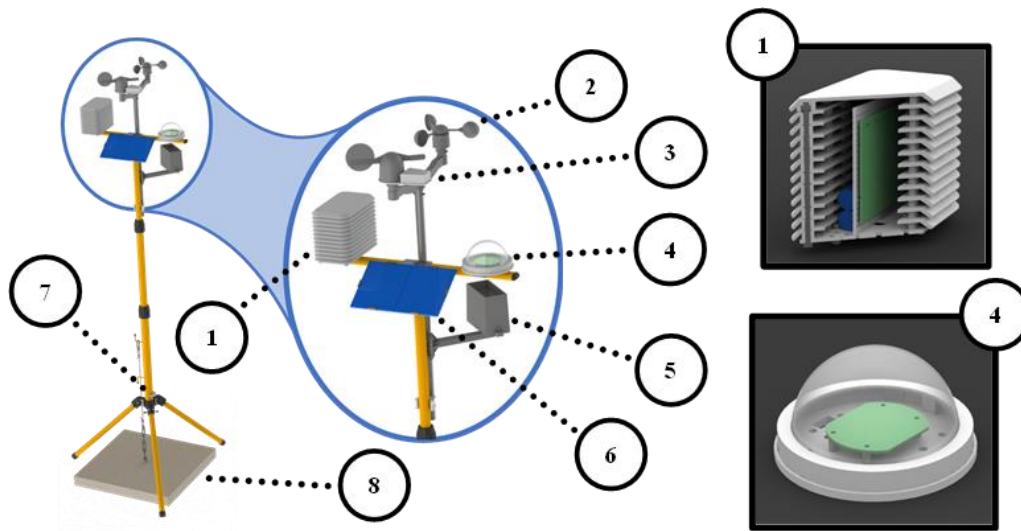
The construction of the weather station itself can be divided into 4 basic parts: the main control unit,

the optical sensor module, other sensors located outside the control unit and the purely mechanical part of the device.

In addition to the NodeMCU V3 development module with the ESP-12E control and communication module containing the Espressif ESP8266 chip, the main control unit also contains some of the sensors, a complete power supply including batteries (solar panels are located outside) and a  $\mu$ SD card slot.

The optical sensor module was designed separately regarding the requirement for the presence of a window. This was obtained by dismantling the T620A security camera dummy.

Wind direction sensors together with a calibration compass, rain gauge and anemometer are located completely separately. All these components are then to be placed on a telescopic tripod originally designed to hold 2 reflectors. However, the tripod itself is relatively light and would poorly withstand gusts of wind, for this reason it is necessary to load the tripod under its centre of gravity. For this purpose, it is possible to hang a steel rope attached to a concrete weight onto the handles originally intended to hold the winded cable for the reflectors. This load, if the rope is properly tensioned, simultaneously eliminates the clearance in the legs of the tripod and thus puts the structure into the spirit level necessary for the correct operation of the rain gauge and wind sensors.



**Figure 2:** Model of the planned weather station design (1 - Control unit case, 2 - Anemometer, 3 - Wind direction sensor with digital compass within its cover, 4 - Optical sensors within their cover, 5 - Rain gauge, 6 - Solar panels, 7 - Ground temperature sensor, 8 - Concrete load)

## 5. WEATHER STATION SOFTWARE AND USER TOOLS

The installation and setting up of the weather station begins by storing the user settings in the internal memory of the control unit. At present, this is achieved by modifying the code that controls the control unit. In the future, it is planned to make this setting through a friendlier user interface using a tool located on a computer to which the weather station control unit is connected via USB. Information, which are stored, contain for example Wi-Fi access data, server addresses or access data to IoT platform, which is a communication node available across the Internet that allows the collection and evaluation of data and their subsequent distribution to other IoT devices that make request to obtain this data. Sensitive data will not be able to be read back into the configuration tool in any way.



After completing the user configuration, the weather station is now able to operate independently. At regular intervals, it wakes up from sleep mode and tries to synchronize internal time into full hour using communication with the NTP server. So, measurement cycle repeats every hour. After synchronization, it reads the data from the sensors and then starts saving the data. If the connection to the Wi-Fi network failed or the connection to one of the servers could not be established, the weather station works in offline mode, where time synchronization is performed according to the last calibration value of the timer and saves measured data at least on the SD card in JSON format. In the next measurement cycle, it will try to re-establish the connection to store the current and previous measurements on the IoT platform. The weather station also wakes up from power saving mode to handle variables that must be measured outside the main measurement cycle.

The Thingsboard platform was chosen for the initial development phase, which is available completely free of charge and is possible to operate it on the Raspberry platform for example, even though the configuration requires very advanced knowledge in the field of information technology. The weather station uses the MQTT protocol for data storage, due to which it should not be a problem to connect the weather station with other platforms. In the future, it is also planned to expand the number of supported communication protocols, including secured ones, for even greater compatibility.

## 6. CONCLUSION

The whole project has a huge potential and can be included together with other systems in the family of IoT elements, it can help to easier and cheaper gaining of information about current and past weather, or to control of other IoT elements based on user previously set events. The whole proposal also leaves the door open for further development and optimization of its functions.

## REFERENCES

- [1] PALA, Andrej. *Design of IoT weather station*. Brno, 2021. Available from: <https://www.vutbr.cz/studenti/zav-prace/detail/137329>. Semester thesis. Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of electrical and electronic technology. Supervised by Petr Vyroubal.
- [2] BUCK, Arden L. New Equations for Computing Vapor Pressure and Enhancement Factor. *Journal of Applied Meteorology* [online]. USA: American Meteorological Society, 1981, 1981-08-03, **20**(12), 1527–1532 [cit. 2022-03-10]. ISSN 0021-8952. Available from: doi:10.1175/1520-0450(1981)020<1483:ANFDSF>2.0.CO;2
- [3] U.S. GOVERNMENT. *U.S. Standard Atmosphere* [online]. 19770009539. USA: U.S. Government Printing Office, 1976 [cit. 2022-03-10]. NASA-TM-X-74335, NOAA-S/T-76-1562. Available from: <https://ntrs.nasa.gov/api/citations/19770009539/downloads/19770009539.pdf>
- [4] JOHNSTON, Danvers E. JOHNSTON a Peter R. MORENO. A conversion guide: Solar irradiance and lux illuminance. JVE INTERNATIONAL. *Journal of Measurements in Engineering* [online]. Lithuania: JVE International, 2020, 4 December 2020, December 2020, **8**(4), 153-166 [cit. 2022-03-10]. ISSN 2424-4635. Available from: doi:<https://doi.org/10.21595/jme.2020.21667>
- [5] ROHM CO., LTD. *Ultraviolet Sensor IC with Voltage Output: ML8511 (UV Photodiode + Amp) (No.0000000016)* [online documentation]. Japan: ROHM Co. [cit. 2022-03-10]. Available from: [https://cdn.sparkfun.com/assets/learn\\_tutorials/2/0/6/ML8511\\_UV.pdf](https://cdn.sparkfun.com/assets/learn_tutorials/2/0/6/ML8511_UV.pdf)
- [6] LAPIS SEMICONDUCTOR CO., LTD. *ML8511: UV Sensor with Voltage Output* [online documentation]. FJDL8511-05. Japan: Lapis Semiconductor Co., 2008, 8. March 2013 [cit. 2022-03-10]. Available from: [https://www.laskakit.cz/user/related\\_files/ml8511\\_3-8-13.pdf](https://www.laskakit.cz/user/related_files/ml8511_3-8-13.pdf)
- [7] HORNBACH BAUMARKT CS SPOL. S R. O. Tripod for two reflectors. *Hornbach: Tripod for two reflectors* [online]. Prague: Hornbach, 1992 [cit. 2022-03-10]. Available from: [https://cdn.hornbach.cz/data/shop/D04/001/780/491/410/285/DV\\_8\\_6816107\\_01\\_4c\\_DE\\_20180809051800.jpg](https://cdn.hornbach.cz/data/shop/D04/001/780/491/410/285/DV_8_6816107_01_4c_DE_20180809051800.jpg)

# Heat Exchanger Regulation

M. Černík and O. Mihálik

Department of Control and Instrumentation, Brno University of Technology, 61600 Brno, Czech Republic.

E-mail: [xcerni12@vutbr.cz](mailto:xcerni12@vutbr.cz), [ondrej.mihalik1@vut.cz](mailto:ondrej.mihalik1@vut.cz)

**Abstract**—The work deals with the design of a laboratory task including model of heat exchanger station. It includes project documentation, with theoretical and practical part. The theoretical part is concerned with theoretical knowledge about every element used in the design. Practical part deals with the control unit of the station and the simulation of heat exchanger.

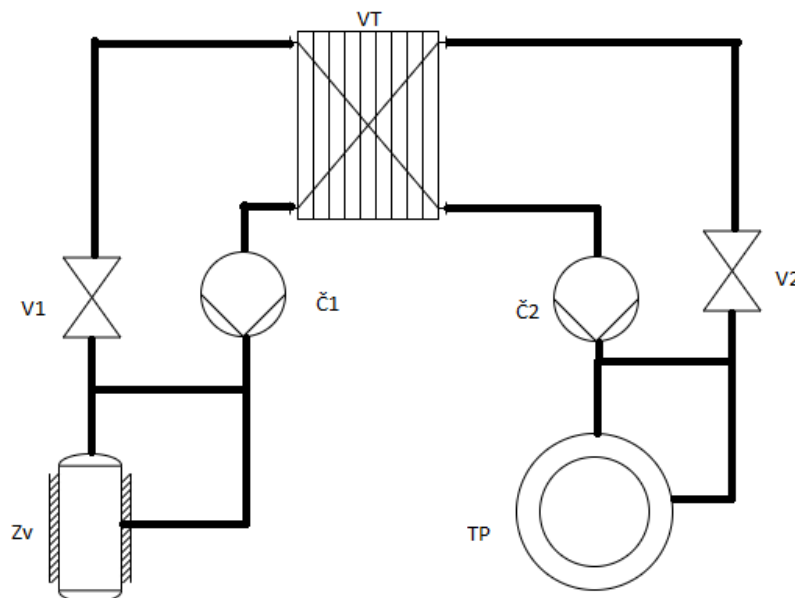
**Keywords**—Heat, heat exchanger, energy transfer, sensor, fluid

## 1. INTRODUCTION

The aim of this work is to perform the control of the heat exchanger on an existing measuring station. Not only from a practical point of view but also from a theoretical point of view using the simulation program MATLAB Simulink.

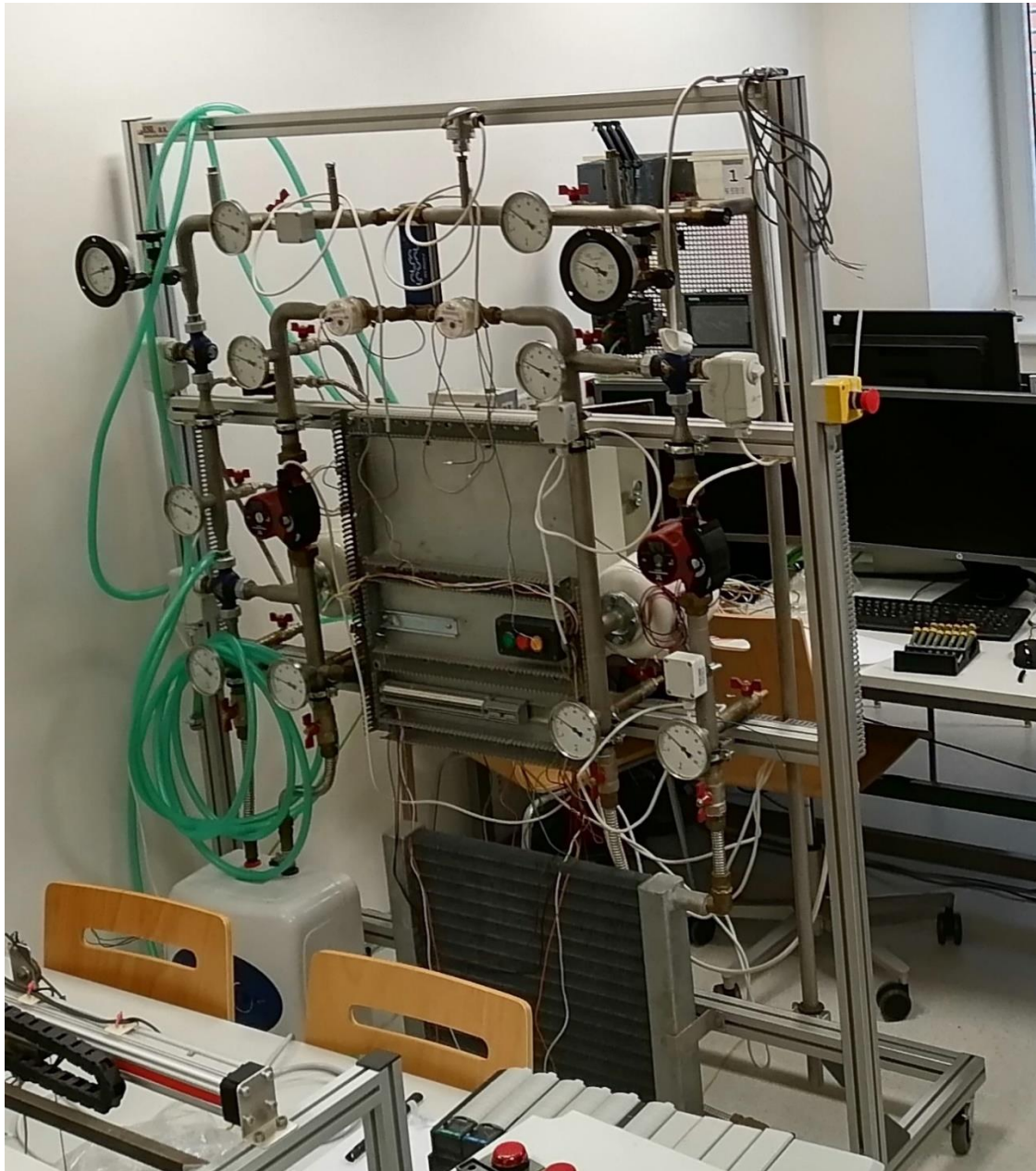
## 2. MEASURING STATION

The measuring station consists of two fluid circuits which are interconnected by a heat exchanger. In the first circuit there is a boiler, which has the task of heating the fluid. In the second circuit, on the other hand, there is a heater on which the heat is to be lost into the environment. Both circuits include pumps that drive the fluid in the pipes, measuring sensors and solenoid valves. The wiring diagram of this system can be seen in **Figure 1**.



**Figure 1** : Simplified diagram of the measuring station system

In **Figure 1** V1 and V2 denote valves, VT is heat exchanger, Č1 and Č2 are pumps, Zv is the boiler and TP is the heating element.



**Figure 2:** Measuring station with the heat exchanger

### 3. SIMULATION

I programmed a dynamic model of the system in Matlab Simulink. At the beginning, I needed to understand the function of each element in the system and replace it in the simulation in an appropriate way.

I replaced the boiler, as a constant source of hot water, with a unit jump function, which has a similar behavior. I have replaced the solenoid valves with adder blocks for the time being, as there is no need to control the temperature of the water in the system in any way. Later in the bachelor thesis there will be a controller at this point. As a substitute for the heat exchanger, I used 4 blocks that successively multiply a constant signal and are interconnected between the circuits (see **Figure 3**).

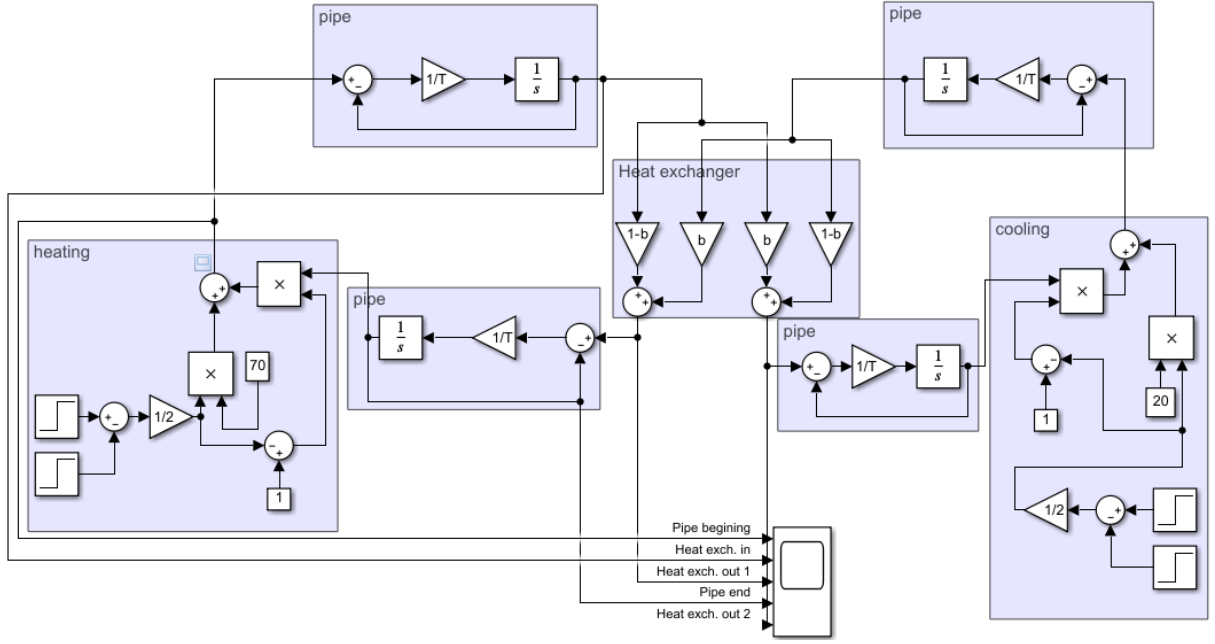
I solved the replacement of the piping itself with a 1st order feedback system. Since water travels in the pipe, it cannot be assumed that what appears at the inlet will immediately appear at the outlet. Therefore, the temperature in the system must slowly build up, which is met by the transient characteristics of a 1st order feedback system.

$$F(p) = \frac{1}{T \cdot p + 1}, \quad (1)$$

where  $T$  is time constant and the parameters of the dynamical system are

$$\begin{aligned} B &= 0,45, \\ T &= 3 \text{ s}. \end{aligned} \quad (2)$$

The parameter  $B$  determines the efficiency of the exchanger. The parameter is always in the range  $0 < B < 1/2$ . If the parameter  $B$  is close to zero, the function of the heat exchanger is almost zero, it does not transfer any heat. On the other hand, if it approaches  $1/2$ , the heat exchanger functions as an ideal exchanger, so that all the liquids which exit the exchanger are of the same temperature.

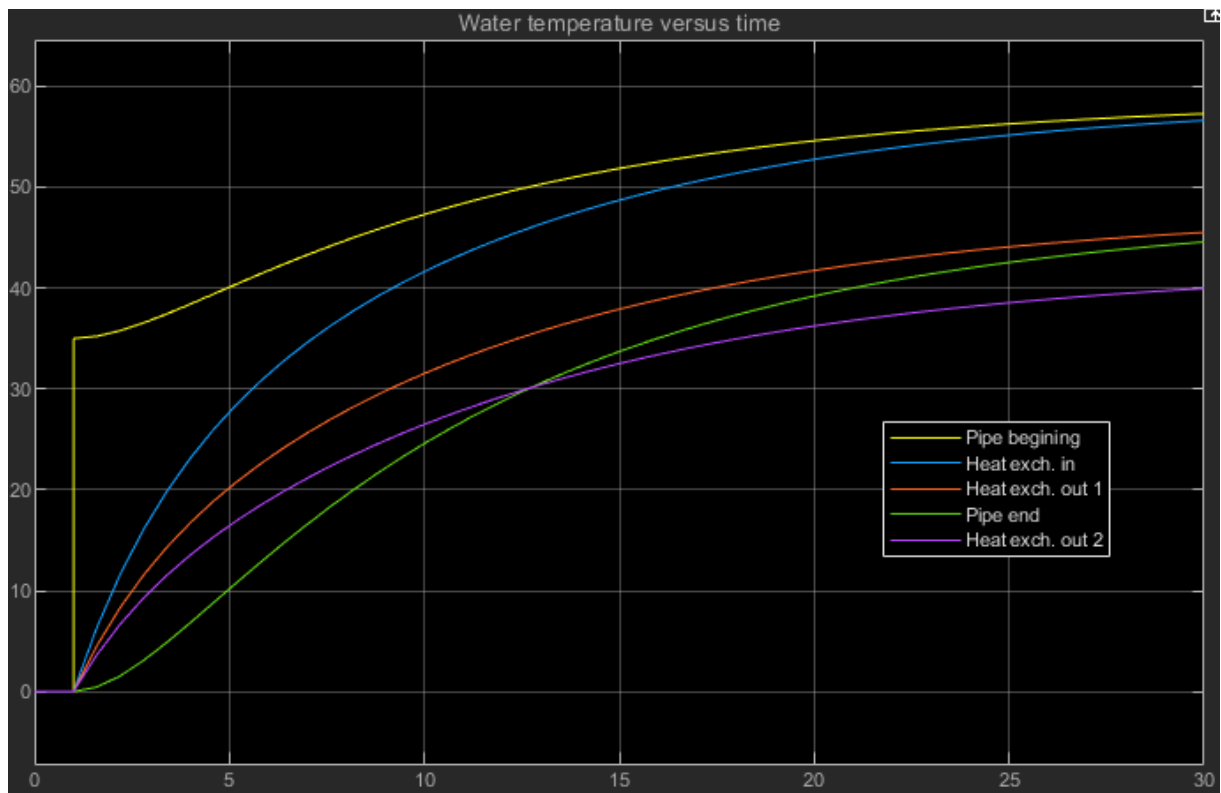


**Figure 3:** Dynamic system model in Simulink

The model which is shown in **Figure 3** will be used to design and tune a PID controller. For this purpose, standard approaches, such as zero-pole placement and optimization according to minimization of the ITAE criterion [3] will be employed. The criterion can be expressed using definite integral

$$J = \int_0^{\infty} t|e(t)| dt = \int_0^{\infty} t|w(t) - y(t)| dt, \quad (3)$$

where the quantity  $e(t)$  denotes the error of the controller, which is being optimized,  $w(t)$  is the required temperature and  $y(t)$  stands for the temperature in the secondary circuit of the exchanger. This signal is displayed using magenta line in the following figure.



**Figure 4:** Simulink step responses of the modelled non-linear dynamical system

#### 4. CONCLUSION

In this paper we discussed a method of constructing a non-linear dynamical model of a heat exchanger. The simulation result, which can be found in **Figure 4**, appears to be in accordance with the physical theory. The behaviour of the whole system depends on the magnitude of the input signal, and the initial temperatures of different parts of the system. Once the boiler is turned on, the system gradually increases its temperature. This enables us to model the system for the purpose of offline tuning of a PID controller, which will be the next aim of our work. So far, we have focused on theory and programming, without the opportunity to test the program on the exchanger station.

#### ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-20-6205 — “Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0” financially supported by the Internal science fund of Brno University of Technology

#### REFERENCES

- [1] Yuvraj Bhushan Khare, „PID Control of Heat Exchanger System,“ International Journal of Computer Applications, 2010, vol. 8, no. 6, p. 22–27.
- [2] Pájené výměníky: SCHÉMA PROUDĚNÍ - Typ 0. In: *G-MAR* [online]. Karlovy Vary: G-MAR PLUS, 2014 [cit. 2021-01-02]. Dostupné z: <https://www.g-mar.cz/pajene-vymeniky/>
- [3] P. BLAHA, P. VAVŘÍN. *Řízení a regulace I*. VUT v Brně, 2005.

# DESIGN OF 3RD-ORDER ELECTRONICALLY RECONFIGURABLE FILTER BY INVERSE CHEBYSHEV APPROXIMATION

T. Kříčka, L. Langhammer

Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Radio Electronic, Technická 12, 616 00 Brno, Czech Republic

E-mail: [xkrick00@vutbr.cz](mailto:xkrick00@vutbr.cz), [langhammer@vut.cz](mailto:langhammer@vut.cz)

**Abstract**—This paper deals with the proposal of an analog filter of the 3rd-order using Inverse-Follow-the-Leader-Feedback (IFLF) topology, which uses approximation of inverse Chebyshev. The filter is working in the current mode. The main advantage of this filter is an ability to electrically reconfigure its available transfer functions. The work focuses on the use of modern active elements. For design the filters was used transconductance amplifiers, second-generation current conveyors, current followers and current amplifiers. The function of the filter is verified by simulations in the OrCad PSpice software, with the ideal and behavior models of used active elements.

**Keywords**—frequency filter, electronic reconfiguration, active element, iFLF topology, current conveyor, operational transconductance amplifier, transfer function, current mode, output summation

## 1. INTRODUCTION

The ability of the electronic reconfiguration is, that it is possible to change the resulting output response transfer functions without any reconnection of electrical circuit. Therefore, it is possible to make SISO (Single-Input-Single-Output) filter structures with more than one transfer function. Electronically controllable parameters of suitably designed active elements are utilized to control the transfer function. The main advantage of these structures is that the parameters can be controlled by external current sources fastly and continuously without any transient. The functions can be fine-tuned. Required function is achieved by the setting active elements, which eliminate specific term of the transfer function [1] [2].

Many design methods topologies could be used to design higher order structures, for example KHN (Kerwin, Huelsman, Newcomb), MLS (Multi Loop Structure) [5] [4]. One of them is IFLF (inverse-follow-the-leader-feedback) topology. Higher order filters reach a steeper transfer function steeper transition between the pass-band and stop-band area. The frequency response has high slope, when is going to stopband. The steepness of transfer function is defined by  $n \cdot 20$  dB/dec, where  $n$  is number of filter order.

## 2. FILTER DESIGN

The structures based on IFLF topology have more component requirements in opposite of noninverting FLF or another high order structures. The most stressed component in IFLF structures is the output component (e.g. electric integrator or derivator), because it requires multiple outputs to create feedback into individual nodes of the topology. In the theory, if I have a 3<sup>rd</sup>-order filter, the latest OTA must have 4 outputs [1]. The connection of OTA-C blocks in a cascade is used to design IFLF filters in this particular case.

The output function is carried out by the method OS (output summation) [3] [4]. OS has advantage, because this method doesn't need add additional active component. This solution is taking a current response from every OTA-C node and sums them in one node. The controlling of transfer function is made by current amplifiers, which are connected behind every OTA-C structure. At the output the currents are summed up and Kirchhoff's law is used to get required function.

To determine the transfer function of the filter with help of software SMAP 3V02. The program NAF calculated 4 coefficients of the transfer function  $b_{0-3}$ , which are determined by selected tolerance field.



Tolerance field settings for 3-order low-pass filter:

- approximation inverse Chebyshev,
- characteristic frequency  $f_0 = 63,661$  kHz,
- minimal possible attenuation in band-pass = -3 dB,
- frequency of stop-band = 63,661 MHz,
- minimal possible attenuation in stop-band = 60 dB.

$$b_3 = 1 \quad ; \quad b_2 = 8,032 \cdot 10^5 \quad ; \quad b_1 = 3,225 \cdot 10^{11} \quad ; \quad b_0 = 6,537 \cdot 10^{16} \quad (1)$$

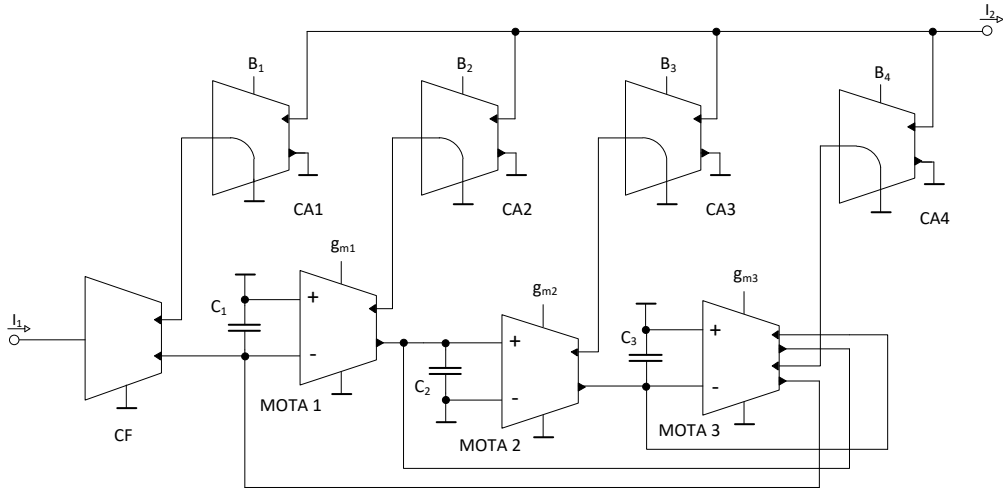
$$\begin{aligned} N(p) = & -p^3 C_3 C_2 C_1 B_1 \\ & + p^2 (C_3 C_2 g_{m1} B_2 - C_2 C_1 g_{m3} B_1) \\ & + p (C_3 g_{m2} g_{m1} B_2 + C_1 g_{m3} g_{m2} B_1 - C_2 g_{m3} g_{m1} B_2) \\ & - g_{m3} g_{m2} g_{m1} B_4 - g_{m3} g_{m2} g_{m1} B_3 + g_{m3} g_{m2} g_{m1} B_2 - g_{m3} g_{m2} g_{m1} B_1 \end{aligned} \quad (2)$$

$$K_{I(p)} = \frac{N_p}{p^3 C_3 C_2 C_1 + p^2 C_2 C_1 g_{m3} + p C_1 g_{m3} g_{m2} + g_{m3} g_{m2} g_{m1}}. \quad (3)$$

Structure of the filter is designed for the 3<sup>rd</sup> -order, the cascade is composed of a 3 OTA-C blocks. Every block represents 1<sup>st</sup>-order filter. Values of capacitors are chosen as  $C_1 = C_2 = C_3 = C = 2$  nF. Based on capacitors values and b parameters, it is possible to calculate the values of transconductance of each OTA.

### 3. FILTER INVOLVENT AND CONTROLLING

The structure of a filter consists of three MOTA (Multiple output OTA). Other used active elements a CF (current follower), which copies currents to individual nodes. The other four are ACA (Adjustable Current Amplifiers) used to control the resulting transfer function.



**Figure 1:** Proposed electronically reconfigurable filter

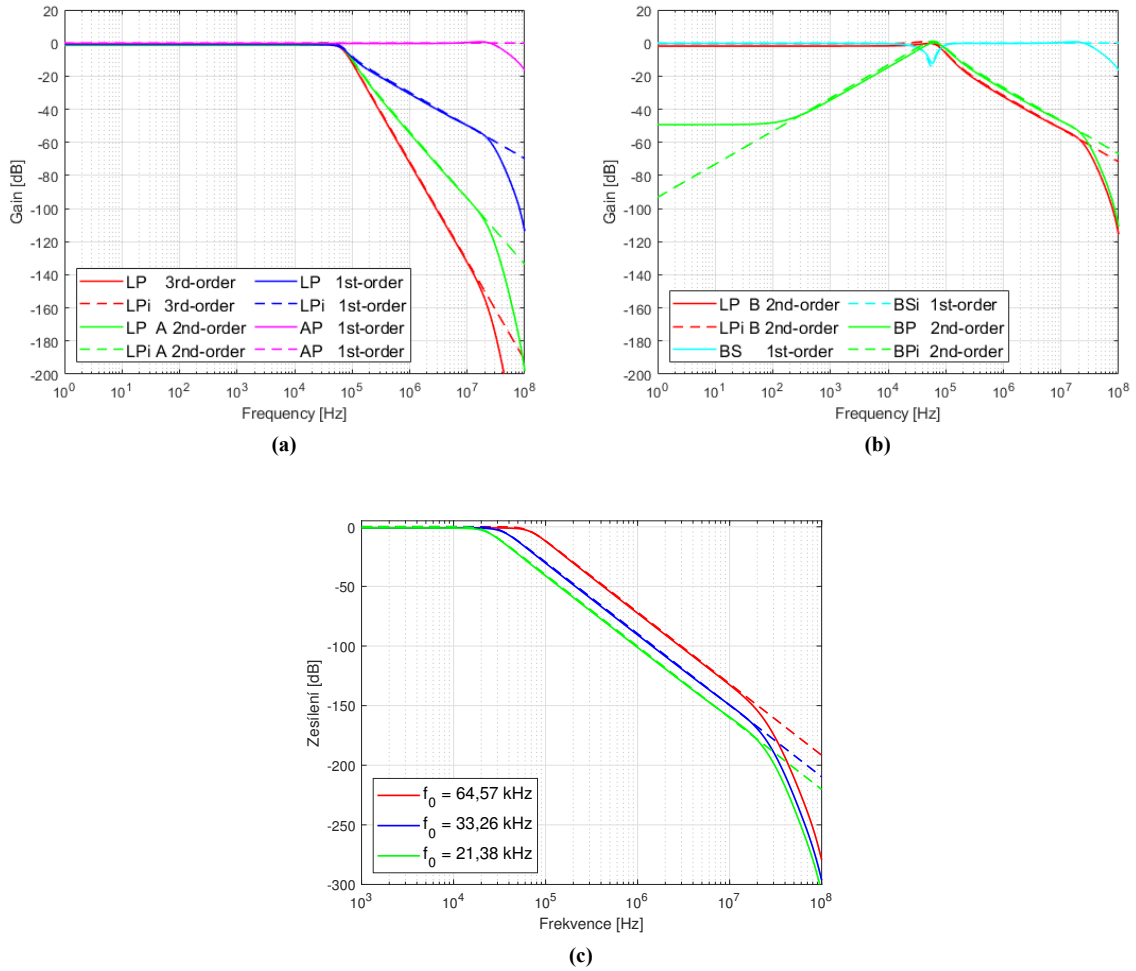
The transfer function is controlled by external sources of voltage, which are used to set current amplification of four amplifiers ACA<sub>1-4</sub>. By setting the amplifiers according to the table I, the appropriate term in the transfer equation is eliminated. Furthermore, it is possible to control the characteristic frequency  $f_0$ . In order to achieve the change of  $f_0$  it is necessary to change all  $g_m$  in the same ratio (multiply or divide  $g_m$  parameter).

**Table I:** Available electrical control transfer functions

Filter	B1 [-]	B2 [-]	B3 [-]	B4 [-]
Low-pass, 3 <sup>rd</sup> -order	0	0	0	1
Low-pass, 2 <sup>nd</sup> -order	0	0	1	0
Low-pass A, 1 <sup>st</sup> -order	0	1	0	0
Band-pass, 2 <sup>nd</sup> -order	0	1,4	0	1,4
Low-pass B, 2 <sup>nd</sup> -order	0	0,8	0,95	0,8
All-pass, 1 <sup>st</sup> -order	1	0	0	0
Band-stop, 1 <sup>st</sup> -order	1	1	0	1

#### 4. SIMULATION

Setting MOTA transconductance to  $g_{m1} = 405,39 \mu\text{S}$ ,  $g_{m2} = 803,09 \mu\text{S}$ ,  $g_{m3} = 1,6 \text{ mS}$ . These values are calculated from the equation 3, based on the values of reaching parameters  $b$  and  $C$  [1] stated in Chapter 2. All simulation simulations of designed IFLF filter are carried out in program OrCad PSpice. For CF and MOTAs I used the UCC (Universal Current Conveyor), made by  $0,35 \mu\text{m}$  technology, to simulate filter on behavioural level. For ACA I used EL2082, where is current gain control with external voltage source.

**Figure 2:** (a) (b) Simulation of available module characteristic (c) Demonstration of changing  $f_0$

The rendered transfer functions are in the graphs 2. Solid lines are displayed in simulations at behavioral level. Dotted line represent ideal parts. From graph A is recognizable, that the filter is designed for low-pass 3<sup>rd</sup> order. These functions are very similar to the ideal ones. In the picture B minor functions are displayed. Some of them were needed to be tuned to the level of amplifiers I. The imperfections in transfer functions are caused by non-ideal elimination members in the denominator 2. Inaccuracy from the ideal transfer is about 20 MHz, which is caused by frequency limitations of real components. In picture C, transfer functions represent change of characteristic frequency  $f_0$  by multiplying the  $g_m$  parameter of each MOTA.

## 5. CONCLUSION

In this paper a fully electronically reconfigurable 3<sup>rd</sup>-order filter with IFLF topology is designed. The transfer functions are simulated with ideal and behavioral simulation models. The structure is designed using the SISO and the control of the transfer function is done by summing the currents at the output. This structure is suitable for PCB.

## REFERENCES

- [1] L. LANGHAMMER, R. ŠOTNER, Third-Order Reconnection-Less Electronically Reconfigurable Multifunctional Filter,” 2020 IEEE International Symposium on Circuits and Systems (ISCAS), 2020, pp. 1-5, doi: 10.1109/ISCAS45731.2020.9180562.
- [2] L. LANGHAMMER, J. DVOŘÁK, R. ŠOTNER, J. JEŘÁBEK, Reconnection-less reconfigurable low-pass filtering topology suitable for higher-order fractional-order design. J Adv Res. 2022 Jul 4;25;257-274. doi: 10.1016/j.jare.2020.06.022. PMID: 32922992; PMC7474249.
- [3] T. DOSTÁL, R. ŠOTNER, J. SLEZÁK, Current-Mode Circuit Base on SIMO OTA :: Review and NEW Applications in Filters. Hikari [online]. 2009, 480 – 495 [cit. 2022-03-07].
- [4] T. DOSTAL, High-Order Filters with Multi-Loop Structure in Current Mode. Radioengineering, 2003, Vol. 12, No. 3, pp. 6–11.
- [5] A.M. KUMAR, B.M.C SRIVASTA, C.U. KUMAR, Voltage mode OTA based KHN filter, 2012 IEEE 55th International Midwestern Symposium on Circuits and Systems, 2012, pp. 698-700, doi:10.1109/MWSCAS.2012.6292116.

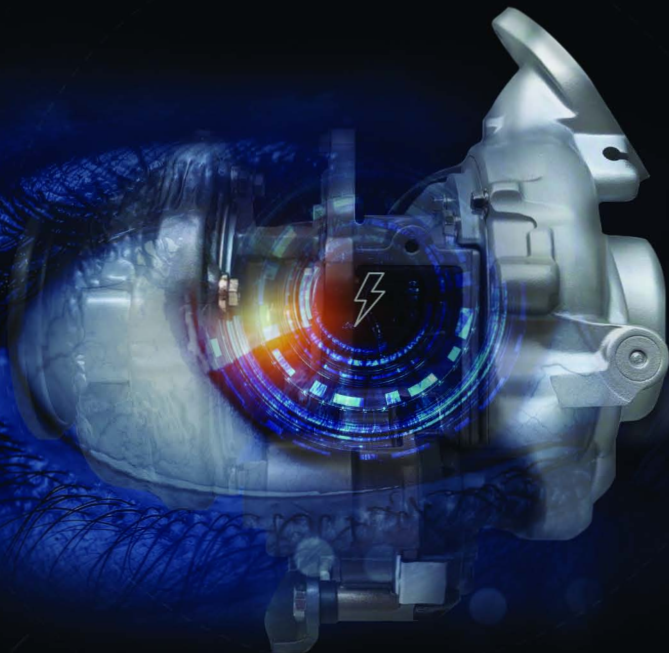
# ELECTRIFICATION IS IN OUR NATURE

## Start your career with us

Garrett develops exciting cutting-edge technologies that drive vehicle hybridization and electrification. From the first hydrogen fuel cell compressor for passenger cars to the industry's first electric turbo, Garrett technologies are shaping a cleaner future of mobility.



Learn more about  
our hybridization  
and electrification  
solutions



**Garrett**  
ADVANCING MOTION

  
Turbo  
Technology

  
Electric &  
Hybrid

  
Connected  
Vehicle

# Universal asynchronous receiver/transmitter implementation in VHDL

Pavel Prášil<sup>1</sup>, Petr Petyovský<sup>2</sup>

<sup>1</sup>Bachelor Degree Programme (3), FEEC Brno University of Technology, Czech Republic

<sup>2</sup>Department of Control and Instrumentation, FEEC Brno University of Technology, Czech Republic

E-mail: [xprasi07@vut.cz](mailto:xprasi07@vut.cz), [petyovsky@vut.cz](mailto:petyovsky@vut.cz)

**Abstract**—The article deals with the design of an asynchronous serial receiver/transmitter and its implementation into the FPGA. The design will be used as a laboratory exercise in the course "Logical circuits and systems". This paper contains the basic design of UART and the following features which will be added. UART design will be used as a communication interface between PC and an existing programmable multichannel sound generator (PSG) design, which is already implemented in FPGA.

**Keywords**—UART, FPGA, VHDL, programmable multichannel sound generator

## 1. INTRODUCTION

UART (Universal asynchronous receiver/transmitter) is an extended communication interface used in embedded systems and microcontrollers. The main advantages are UART is only three-wire connection and the possibility of full-duplex operation. For implementation into FPGA (Field Programmable Gate Array) was used VHDL (Very High Speed Integrated Circuit Hardware Description Language). Description of designed UART implementation is shown in section 2. UART is the basic principle of RS232, RS485, etc. communication standards.

UART design will be used for laboratory exercise purposes, as is shown in section 3. Connection between PSG and UART, and its requirements are also described in that section.

## 2. UART DESIGN

The UART design is compiled from components, which contain previous laboratory exercises of the Bachelor study programme course called "Logical circuits and systems". The architecture of the UART is entirely synchronous by 100 MHz clock signal produced by an oscillator, which is part of the development board NEXYS3. The development board NEXYS3 contains Xilinx FPGA chip SPARTAN-6 LX16. As an inspiration to UART design were used SCI (Serial Communication Interface) of Freescale microcontroller M68HC11 and book "Data, čipy, procesory" written by Czech author Martin Malý [1] [3].

### 2.1. Transmitter

The designed Transmitter is compounded from three components, as is shown in figure 1. The main component is Transmitter FSM (Finite State Machine) and its VHDL description corresponds with standard FSM pattern containing three processes. Data transmission provides the Shift register component. The Shift register is controlled by FSM output signals *load* (transfer data from the buffer to the shift register) and *se* (shift enable). Baudrate of data shift determine signal *ce* (clock enable), that is generated by Clock divider component. Baudrate is set by the generic parameter of the clock divider component. Transmission starts by activation of input signal *tx\_send*. If the Transmitter isn't transmitting data, FSM sets up output signal *tx\_ready*. There is an input signal called *tx\_use\_parity* specifying whether parity is used or not. Therefore the Shift register and FSM have *use\_par* inputs. If this input is set, FSM component increases number of transmitted bits and the Shift register computes and transmits the parity bit after data bits. The Shift register has another input *par\_opt* (parity option) used to select kind of computed parity. Available are even, odd, space and mark parity. Thus mark parity can be used as another stop bit.

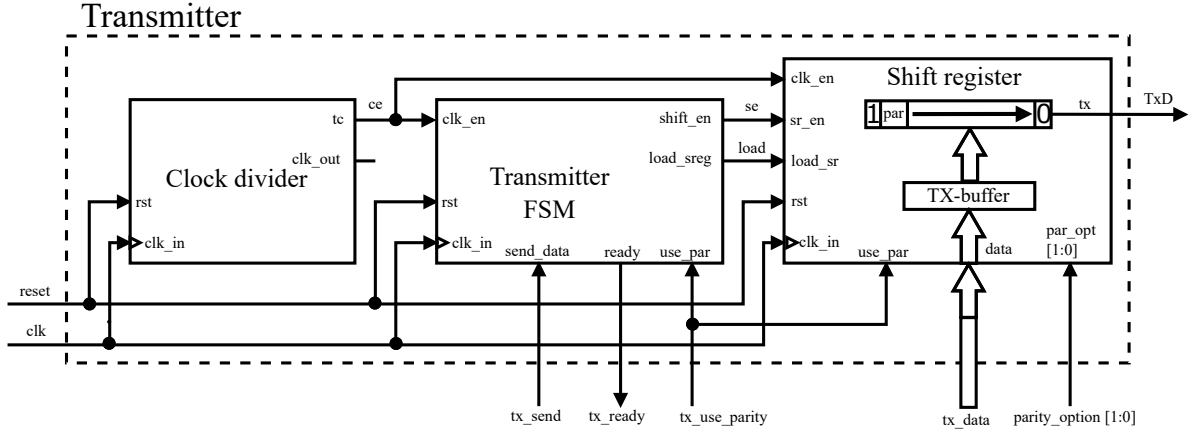


Figure 1: Transmitter functional block diagram.

## 2.2. Receiver

The receiver design is very similar to a transmitter. There are two other components, the Filter and the Start detector. The Filter component works as a double flip-flop synchronizer. The consequence of using this synchronizer is that the metastable state cannot enter further into the circuit. The Start detector provides only start bit detection, therefore it indicates falling edge of filtered received data input signal *rxflt*. The Clock divider component of the receiver works similarly as the Clock divider in the Transmitter, but if no data are incoming to the receiver, the Clock divider remains off. After the start bit is detected, the Clock divider starts to generate *ce* signal. The first generated period is half time due to sampling incoming bits in the middle of theirs. The main FSM component provides its output signals *run* (start and run clock divider) and *load* (transfer received data from the shift register to the buffer). Another output signal of the FSM block called *rx\_data\_status* indicates validity or impropriety of received start, stop and parity bits. The Shift register works similarly as the Shift register in the Transmitter, but the data byte is transferred from the Shift register to the receive buffer at the appropriate moment. There is *par\_chk* (parity check) output signal, which indicates correct or incorrect received parity bit. To enable parity and set expected parity bit value are used similar signals as in the design of the Transmitter. The block diagram is shown in figure 2.

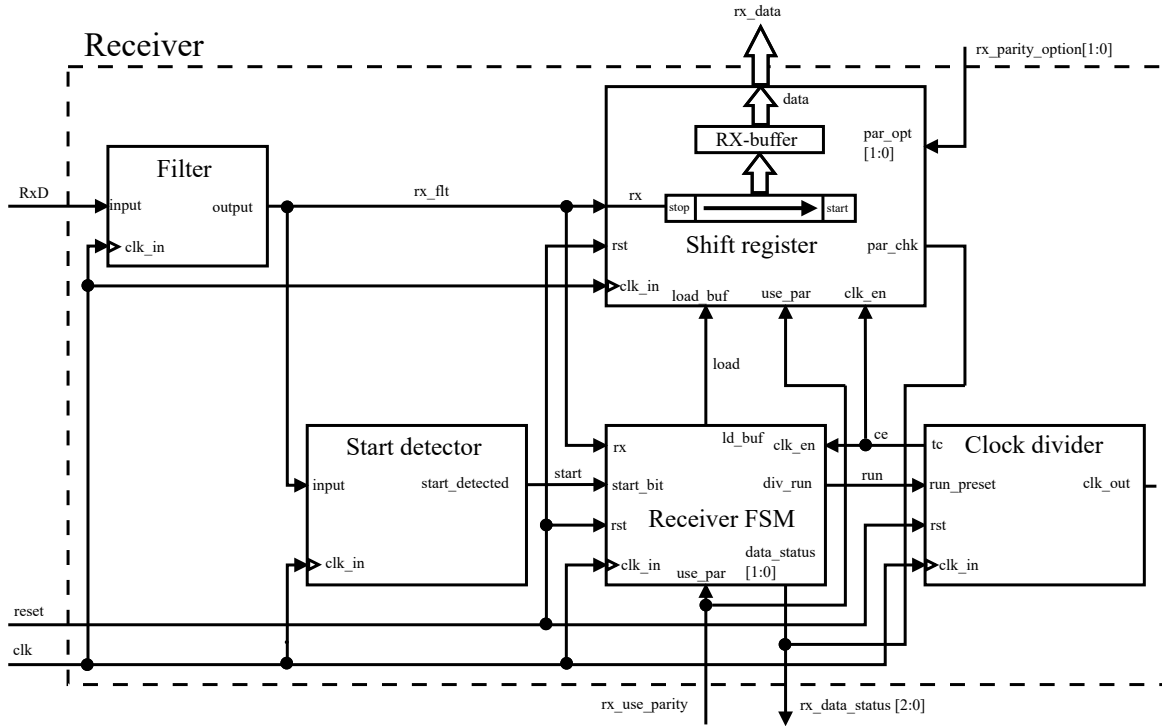


Figure 2: Receiver functional block diagram.



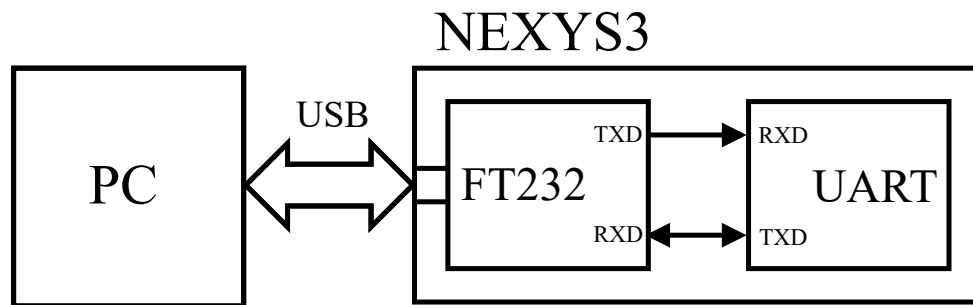
### 3. LABORATORY EXERCISE

Laboratory exercise has been devised as the connection between the PC and NEXYS3 development board, as is shown in figure 3. Development board NEXYS3 contains UART to USB convertor FTDI FT232 [3] [4]. Students can connect *tx\_data* input vector to 8 switches for setting up sending data byte and signal *tx\_send* to one of the buttons. Received data can be shown as a binary code on 8 leds or as a character on a 7-seg display. As PC terminal was chosen Putty client, because it supports serial communication.

The second variant is to connect the *rx\_data* and *tx\_data* signals. Received data are immediately sent back to the PC and shown in the putty terminal. Therefore written characters are immediately shown again in the terminal.

There is a plan that UART VHDL design will be available for students without architectures of some components and students' objective will be to complete those architectures. Hidden architectures will be chosen after consultation with the guarantor of the course Logical circuits and systems.

The main objective is to use UART implementation to connect PC and an existing implementation of programmable multichannel sound generator, which was created as a bachelor thesis last year. Currently, sound data for PSG are implemented directly into FPGA as an array of values. This allows playing only short part of a song, approximately 30 seconds. With UART communication will be possible to play whole songs. This requires connecting received data to registers of PSG and properly controlling addressing and data writing to PSG registers by PSG control signals. There are requirements to indicate whether received data has already been processed by PSG, and to generate confirmation messages and checksum. Construction of these messages has been suggested in the thesis, which deals with PSG [2].



**Figure 3:** PC - NEXYS3 connection block diagram.

### 4. FPGA UTILIZATION

The current design uses 96 D type flip-flops (DFF) out of 18224 and 133 Look-up tables (LUT) out of 9112 (1 %). There are 151 LUT Flip-Flop pairs and 78 of these are fully used, this represents 51 % of them. Inputs and outputs utilization is 36 IOBs out of 232 (15 %). Maximum clock frequency is 100 MHz generated by CMOS oscillator.

### 5. CONCLUSION

The Transmitter is transmitting data properly and computing parity bit correctly. The Receiver also receives data correctly. This was verified by showing ASCII code of each received character on leds. 7-seg display shows start, stop or parity framing errors.

Current design isn't able to change baudrate without another implementation into FPGA, consequently the improvement of Clock dividers is needed. There are two suggestions to compute divider constant. The first one is usage of prescaler and selection bits, according to M68HC11 chip [3]. The second suggestion of baud rate generation is to invent accurate baud rate generation using the Fraction baudrate generator, as MAX3108 UART chip does [5]. The second one seems to be better, because arbitrary and precise baud rate can be generated. The receiver data sampling can be enhanced to triple sampling in the middle of incoming data bits and evaluate bit value as a majority of these three samples, as M68HC11 SCI does [3]. For better data transmission should be designed and implemented software flow control (XON/XOFF) mechanism. Another enhancement is to increase the depth of data buffers, it would work as a FIFO(First-In First-Out) memory.

## REFERENCES

- [1] MALÝ, Martin. *Data, čipy, procesory: vlastní integrované obvody na koleni*. Praha: CZ.NIC, z.s.p.o., 2020. CZ.NIC. ISBN 978-80-88168-53-9.
- [2] HOMZOVÁ, Eliška. *Realizace výukového vícekanálového zvukového obvodu* [online]. Brno, 2021 [cit. 2021-12-13]. Dostupné z: <http://hdl.handle.net/11012/198025>. Bakalářská práce. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav automatizace a měřicí techniky. Vedoucí práce Petr Petyovský.
- [3] *M68HC11 Reference Manual* [online]. 2007 [cit. 2021-11-28]. Dostupné z: <https://www.nxp.com/docs/en/reference-manual/M68HC11RM.pdf>
- [4] *FTDI FT232R: Datasheet* [online]. [cit. 2021-12-22]. Dostupné z: [https://ftdichip.com/wp-content/uploads/2020/08/DS\\_FT232R.pdf](https://ftdichip.com/wp-content/uploads/2020/08/DS_FT232R.pdf)
- [5] *MAX3108: Datasheet* [online]. 160 Rio Robles, San Jose, CA 95134 USA: Maxim Integrated, 2013 [cit. 2022-03-24]. Dostupné z: <https://datasheets.maximintegrated.com/en/ds/MAX3108.pdf>

# Electrochemical etching of graphite tips for potential use in nanoscience

Adam Jonáš

Department of Physics, Faculty of Electrical Engineering and Communication, Brno University of Technology,  
Technická 2848/8, 61600 Brno, Czech Republic

E-mail: [adamjonas86@gmail.com](mailto:adamjonas86@gmail.com)

**Abstract**—The main goal is to find an easy, and more importantly, cheap process of producing sharp tips for the potential use in an atomic force microscope (AFM) or as a source of electrons utilizing cold field emission. This objective is being achieved by the method named electrochemical etching, and the used material is cheap and easily accessible. For this experiment, graphite leads of different hardnesses were used, etched with two types of etching solutions. The first solution contained potassium hydroxide (KOH) dissolved in water and the second contained sodium hydroxide (NaOH), also dissolved in water. Different ratios of chemicals were tried for the manufacturing process. Hardnesses of leads were 2B, B, and HB. After the process of etching, products were examined with a scanning electron microscope (SEM). The radius of tips was measured, and results were compared and evaluated. Stating on the output data of tips, it can be said, that goal was achieved, and tips can be used for their potential purpose. There would be needed another research about how well they fit to work.

**Keywords**—graphite, electrochemical etching, tips, SEM, conditions

## 1. INTRODUCTION

The initiation of this experiment is based on searching for cheap materials which can be used as tips for atomic force microscopes (AFM). Usual tips are made from highly expensive metals like tungsten, platinum, iridium, and others. For example, the monthly average price of platinum in February 2022 was 1049\$ for the troy ounce (=31,1 grams) [1]. So, this is the reason, why this work looks for an alternative material, which would not be so expensive and yet will satisfy the needs and requirements for this task. In this work, the experimental material was graphite, graphite leads, commonly used for writing/drawing. This material is cheap and very easily accessible on the internet [2] or it can be bought in usual stationery. Another advantage of graphite is electrical conductivity [3], which makes it suitable for the method of electrochemical etching.

### 1.1. Electrochemical etching

It is a method, where etching is carried out by electricity. Energy is flowing through the circuit, formed by the metal arm (acting as a cathode), etching solution (medium), and graphite lead (acting like anode). Contact of solution and graphite is an area of reaction. The medium used for the process must be an electrolyte because we need electricity to flow through it. As it is generally known, graphite is made from carbon atoms arranged in a hexagonal structure and its charge is positive. So, when the electricity is turned on, carbon atoms want to move towards the cathode, and they are torn off from the structure. The shape of the metal circle secures even consumption of material. This action will manufacture around and sharp tip. When the tip is created and the lower part of the lead falls away, it is essential to stop electricity as soon as possible, because the top of the tip is still in contact with the solution, and electricity is flowing. Subsequent etching would cause harm to the tip quality [4].

## 2. MATERIALS AND METHODS

### 2.1. Conditions and data

The idea of researching is to find out the most effective combination of conditions, which will create a perfectly sharp tip. Each hardness met both types of solutions, and all used concentrations of these

solutions, at least three times, so it can be checked if the same result was done repeatedly. During all the etching tries, time was measured by the watches. A device used for this work is an etching station for probe production (NT-MDT Spectrum Instruments, Moscow Russia). In total, two types of chemicals, three different concentrations of these chemicals, three types of graphite leads were used. On one pencil lead falls six different conditions, each condition repeated at least three times. Their quality was measured in form of tip radius. The smaller the radius, the sharper the tip. Measuring was implemented with a scanning electron microscope (SEM) Lyra3 (Tescan, Brno, Czech Republic). Settings during measurements were immutable, measuring with a view field of 200 $\mu$ m and accelerating voltage of 10kV.

## 2.2. Pencil leads

As it was said, commercially accessible graphite leads were used, three hardnesses from the same producer (Pilot, Tokyo, Japan). The marking of hardnesses is HB, B, and 2B. Each hardness depends on the content of clay [5]. Clay is usually additive to graphite leads, more specifically, clay is a term for the mineral named kaolinite. With the content of clay in pencil, hardness can be affected. The relation between hardness and volume of clay is directly proportional. In this case, HB hardness is the same as F-marking, F – stands for the firm and contains the biggest amount of clay, B – stands for black, which has a smaller volume of clay, and 2B has the smallest amount of clay.

## 2.3. Etching substances

During this research, two different chemical materials were used: potassium hydroxide (KOH) and sodium hydroxide (NaOH). They were used in diverse quantities for dissolving in water. Amounts used for creating multiple solutions were 2g/40ml, 4g/40ml, and 6g/40ml. These chemicals were used because they are quite easily accessible and cheap indeed. They can be bought in the usual drugstore. When they are dissolved in water, chemicals will split into two groups, K or Na and OH group. OH, group is needed and important, because of its negative charge and subsequent carbon attraction.

## 2.4. Method of tip production

Lead is firmly placed into the holder and goes through the circle of the metal arm. This metal arm is submerged into solution, round space of circular part gets filled by solution and returns to its primary position. Now graphite lead is again in the middle of round space, soaked in electrolyte. Then electricity is turned on and electrochemical etching will start. It will continue until the lead is separated into two parts and the tip loses contact with the solution. At this moment, the circuit is broken, electricity is not flowing anymore, and the process is done.

## 3. RESULTS

Results are summarized into two tables. **Table 1** represents the results of the KOH solution, and **Table 2** represents the NaOH solution. In the top cell of the column is listed concentration of etching solution. Then each cell underneath represents one hardness with two main information, the arithmetic average of etching time, calculated from all tries in the same conditions, and best-achieved radius from all tries in the same conditions. The format of time in cells is in seconds, and radius sizes are in micrometers ( $\mu$ m).

**Table 1: Variations of KOH solution**

Concentration of 6g/40ml	Concentration of 4g/40ml	Concentration of 2g/40ml
2B Average etching time: 113s Best achieved radius: 52.07 $\mu$ m	2B Average etching time: 102s Best achieved radius: 27.47 $\mu$ m	2B Average etching time: 169s Best achieved radius: 11.02 $\mu$ m
HB Average etching time: 140s Best achieved radius: 15.33 $\mu$ m	HB Average etching time: 381s Best achieved radius: 27.74 $\mu$ m	HB Average etching time: 442s Best achieved radius: 6.76 $\mu$ m
B Average etching time: 181s Best achieved radius: 12.81 $\mu$ m	B Average etching time: 201s Best achieved radius: 10.63 $\mu$ m	B Average etching time: 317s Best achieved radius: 7.67 $\mu$ m

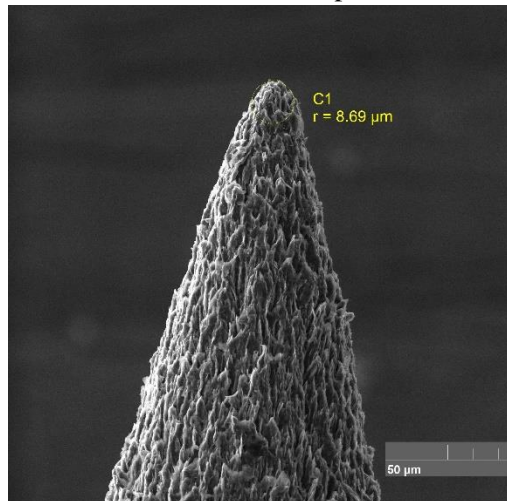
**Table 2: Variations of NaOH solution**

Concentration of 6g/40ml		Concentration of 4g/40ml		Concentration of 2g/40ml	
2B	Average etching time: 153s Best achieved radius: 18.47 $\mu$ m	2B	Average etching time: 82s Best achieved radius: 9.6 $\mu$ m	2B	Average etching time: 119s Best achieved radius: 17.3 $\mu$ m
HB	Average etching time: 246s Best achieved radius: 28.04 $\mu$ m	HB	Average etching time: 236s Best achieved radius: 10.47 $\mu$ m	HB	Average etching time: 209s Best achieved radius: 8.69 $\mu$ m
B	Average etching time: 176s Best achieved radius: 13.37 $\mu$ m	B	Average etching time: 91s Best achieved radius: 16.71 $\mu$ m	B	Average etching time: 181s Best achieved radius: 7.74 $\mu$ m

Few observations can be described. The first observation is logical, the hardest leads take a longer time for creation than the softest ones. This is happening repeatedly in all cases except one. An interesting case is when hardness B during etching with 6g/40ml KOH solution, has the longest average time of production, although it does not contain the biggest amount of clay. It is not known why this action happens, and there would be needed another research for finding out the reason. This phenomenon is not repeated in different conditions. So, it can be stated that almost in every case, producing time is moving directly proportional to the content of clay.

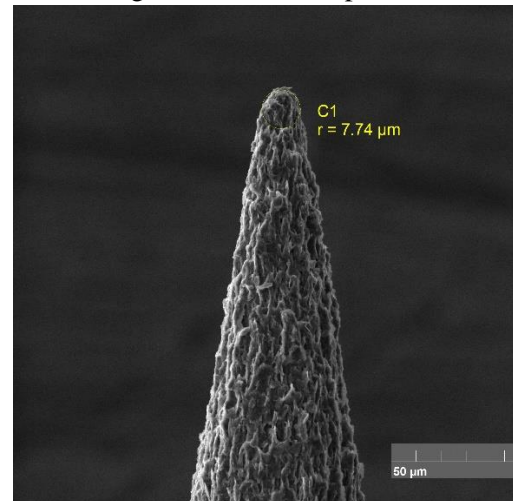
## 2.5. Photos of the best tips

These are the photos of the best-produced tips. They were taken by the electron microscope during the examining and measuring process. The yellow inscription is an expression for radius, r stands for radius, and it is measured in micrometers. In the right lower part of the photo is a grey inscription, it is a scale measured in micrometers. Each photo has a title beneath according to conditions of production.



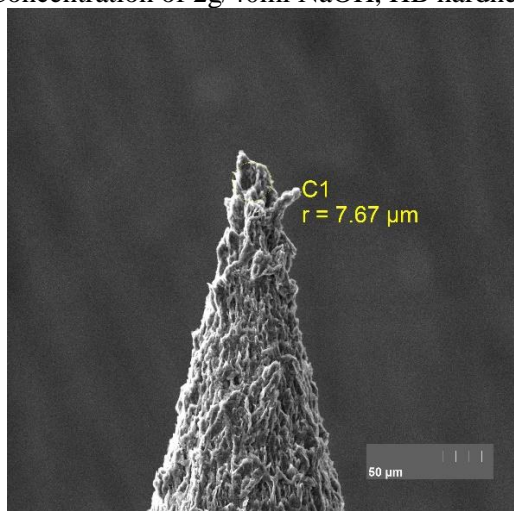
**Figure 1:**

Concentration of 2g/40ml NaOH, HB hardness



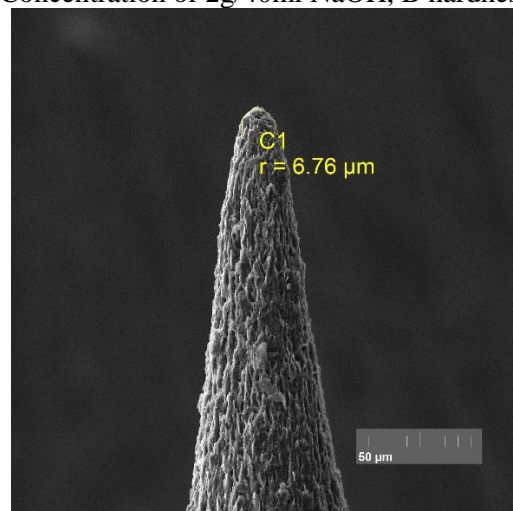
**Figure 2:**

Concentration of 2g/40ml NaOH, B hardness



**Figure 3:**

Concentration of 2g/40ml KOH, B hardness



**Figure 4:**

Concentration of 2g/40ml KOH, HB hardness

#### 4. CONCLUSION

After monitoring the results, several statements can be made. The best results were achieved with the weakest solutions and the worst with the strongest solutions. It cannot be stated that each hardness acts the same in equal conditions. Three times repeated conditions of 6g/40ml KOH and hardness HB, gave very different radii (1. 58.39 $\mu$ m, 2. 15.33 $\mu$ m, 3. 34.65 $\mu$ m). From this example is clear, that equal conditions which can be set, do not guarantee the same result. This phenomenon occurs with other conditions too. What can be stated is, that the average radius of tips decreases with the concentration of the solution. The average radius in 6g/40ml KOH is 26.73 $\mu$ m, in 4g/40ml KOH 21.94 $\mu$ m, and in 2g/40ml average radius is only 8.43 $\mu$ m. The same decreasing process happens with the second type of chemical. The best tip was produced with conditions of 2g/40ml KOH and with the strongest hardness HB. On the contrary, the worst tip was produced with the strongest concentration of 6g/40ml KOH and with the softest hardness 2B. In total, it can be stated that this method of producing is capable of creating tips almost sharp enough to be used in SEM or AFM. Almost sharp means that some additional sharpening must be done. For example, with a focused ion beam (FIB) but this process could be another research in the future. The life expectancy of this type of tip is certainly lower than usual metal tips because graphite is a much softer material than tungsten or platinum. But the time and price of manufacturing are a lot lower, so it is profitable to choose this method and save both, time and money. Even if this research is realized on small scale with limited options, quite interesting and good-looking results are created. It is important to invest in this type of research in the future because it opens new possibilities for cheap tips fabrication and usability.

#### ACKNOWLEDGMENT

I would like to thank a lot to the advisor of my work, Ing. Nikola Papež, Ph.D. for his patience, explanation, valuable comments, and guidance. Also, for giving me the opportunity to participate in the EEICT competition and for the opportunity to feel the world of researching. Without him, I would not be able to create this work.

#### REFERENCES

- [1] "World Bank Commodities Price Data (The Pink Sheet)." Worldbank.org <https://thedocs.worldbank.org/en/doc/5d903e848db1d1b83e0ec8f744e55570-0350012021/related/CMO-Pink-Sheet-March-2022.pdf> (accessed March 2, 2022)
- [2] "Leads" Koh-i-Noor.eu <https://eshop.koh-i-noor.eu/category/leads>
- [3] Marina Cristina Tanzi, Silvia Farè, Gabriel Candiani. "Graphite" in Foundations of Biomaterial Engineering. Cambridge, MA, USA: Academic Press, p.55, 2019 [Online]. Available: <https://www.sciencedirect.com/book/9780081010341/foundations-of-biomaterials-engineering#book-info>
- [4] Rakesh K. Prasad, Dilip K. Singh "Experimental details" in Low-cost electrical probe station using etched Tungsten nanoprobe: Role of cathode geometry. Department of Physics, Birla Institute of Technology Mesra, Ranchi-835215, p.2, 2020 [Online]. Available: <https://iopscience.iop.org/article/10.1088/2632-959X/abb6c4/pdf>
- [5] Allie Gluya. "Graphite pencil degrees of hardness explained." FaberCastell.com <https://www.fabercastell.com/blogs/creativity-for-life/graphite-pencil-lead-degree-hardness> (accessed April 20, 2021)



# Automated firesport stopwatch

Jan Drška

Brno University of Technology – Faculty of Electrical Engineering and Communication, The Czech Republic

E-mail: xdrska@vutbr.cz

**Abstract**—The main object of this article is the design and implementation of an automatic timing system for fire sports, which includes time measurement, automatic target discharging and a large-format display to show the results. The system is based on Arduino and uses modules of this platform. Due to the distances involved, the use of wireless communication is also described. The possibility of extending the system with additional features and functions is also discussed.

**Keywords**—Arduino, firesport stopwatch, firesport, large format display, process automation

## 1. INTRODUCTION

This article deals with the design of an automatic timer for fire sport. Fire sport is nowadays a very popular sport, which is mainly participated by members of volunteer fire departments (hereinafter referred to as VFDs). This sport is most widespread in villages and smaller towns, where VFD units are active. Firefighting sport is still evolving and the requirements not only for the competitors but also for the equipment are increasing. Until now, the older methods of timekeeping have been used in competitions using stopwatches, which required an operator. As the sport and the units themselves evolved, better performances began to be achieved, with units of seconds or hundredths of a second. Manual stopwatch timekeeping ceased to meet the requirements and is gradually being replaced by automatic stopwatches, which improve accuracy and eliminate operator error.

Nowadays, there are timekeepers that can be freely purchased from online stores. They can also be purchased with additional accessories at the appropriate price depending on the design and size of the set. The problem with today's timekeepers is the size and legibility of the display, which plays a big role in competitions. It is essential that the display is readable in direct sunlight or in reduced visibility such as fog, rain or gloom. Due to the length of the fire sports courses, it is necessary that the digits on the display are large enough and legible. Another problem with most solutions is the need to manually operate the targets and discharge water, which usually requires at least one person.

The racecourse consists of several basic parts, which are two water targets, a base for the fire engine, water tank and starting line. During the race itself, the competition team and their equipment are on the course. This equipment is owned by each unit and is not part of the timing design. The competition scheme is shown in Figure 1. [1]

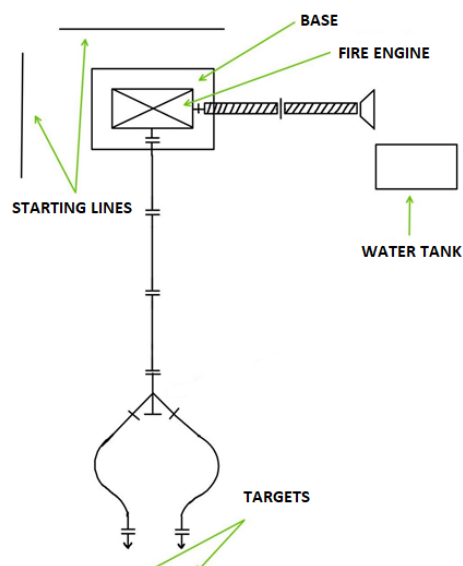


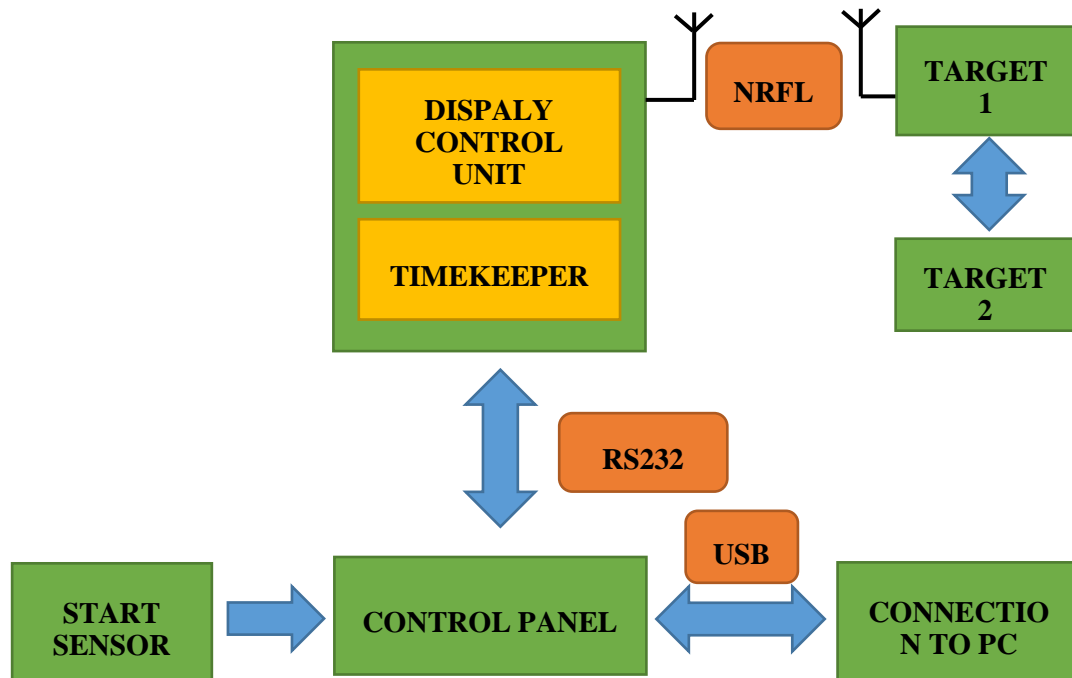
Figure 1: Competition scheme

## 2. NEW STOPWATCH DESIGN

There are plenty of ready-made commercial solutions on the market that can be used, it just depends on the customer's requirements. The available sets are very similar and often differ only in small details.

Most solutions have automatic timekeeping, light or sound signaling. The better sets have built-in wireless communication and have large format displays to show actual times. The new timekeeper combines the benefits and capabilities of available timekeepers and adds some brand-new features.

The system is based on Arduino, which is nowadays readily available and fully sufficient for timekeeping needs. Yes, I can be based on Raspberry or ESP, but we need to keep this simple and easy for service and operating. The timer is based on 3 Arduino boards that communicate with each other. These are the units of the main control panel, the control panel and the water targets. The communication and interconnection of the individual parts of the setup is shown in Figure 2.



**Figure 2:** Scheme of communication between each module

The target control unit is a single circuit board that processes and controls both targets and communicates with the main unit. The targets include water level detection and a fill light. None of the commercially available timers include the ability to automatically discharge the targets when they are full. This new timer will have electric valves that will drain the water themselves when filled. This feature will save at least one operator.

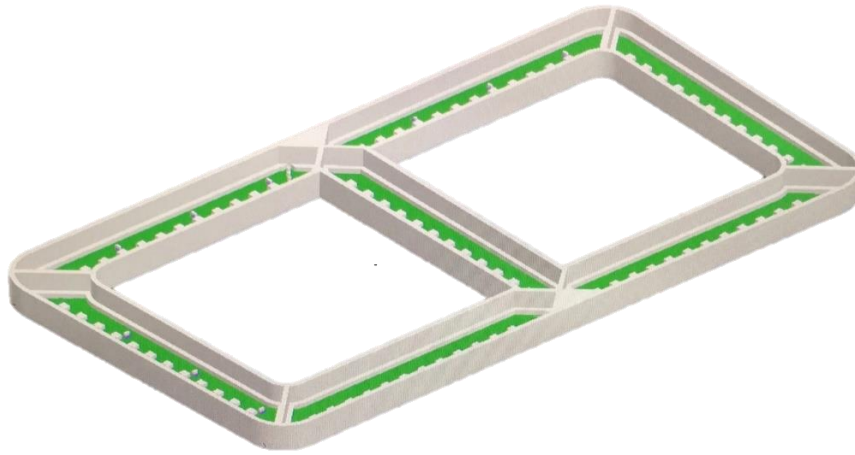
The control panel is used for the referee who controls the whole competition. On this unit there is an LCD display with a rotary encoder module for easy orientation in the menu, buttons and there is also the possibility to connect an external keyboard according to customer requirements. With the help of this control unit, it is possible to start, stop or restart the stopwatch. The start can be carried out using a start sensor (IR, start gun, etc.).

The main control unit is used to measure and display the time. The unit communicates with the targets and the operator control panel. This unit includes a large format display of its own design on which the times of two streams can be displayed. The resolution of the display is twice 4 digits. This configuration is fully sufficient to display minutes and seconds. Milliseconds are displayed only to the referee on the small LCD display. [2]

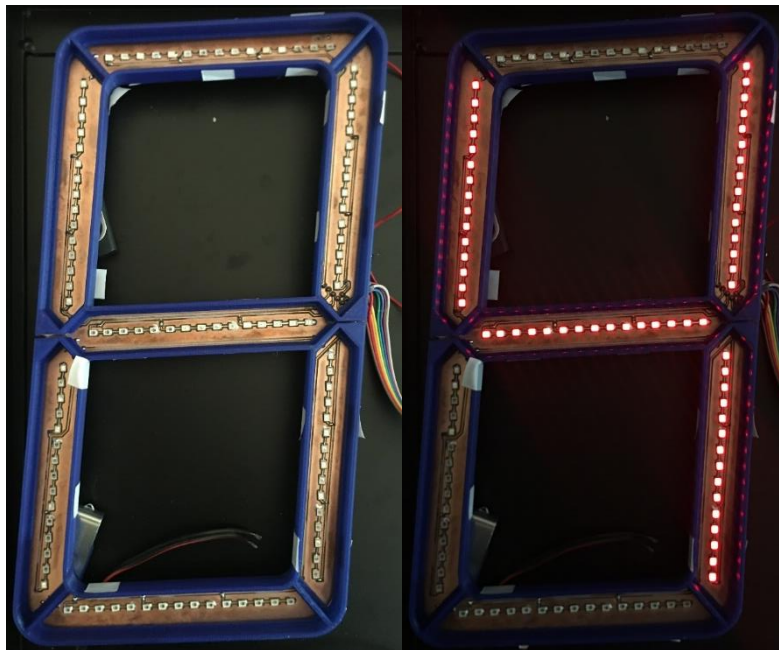
### 3. PRODUCTION OF THE DESIGNED SOLUTION

Several PCBs have been designed for the timer to drive sub-components. These boards serve as expansion boards for Arduino modules that are ready for easy removal or installation. Everything is prepared with connectors to make any maintenance or assembly easier. The boards were manufactured by JLCPCB and were hand fitted with components. The essential part of the system is the large format display, which is divided into 8 digits, where each digit is divided into 5 parts for ease of installation.

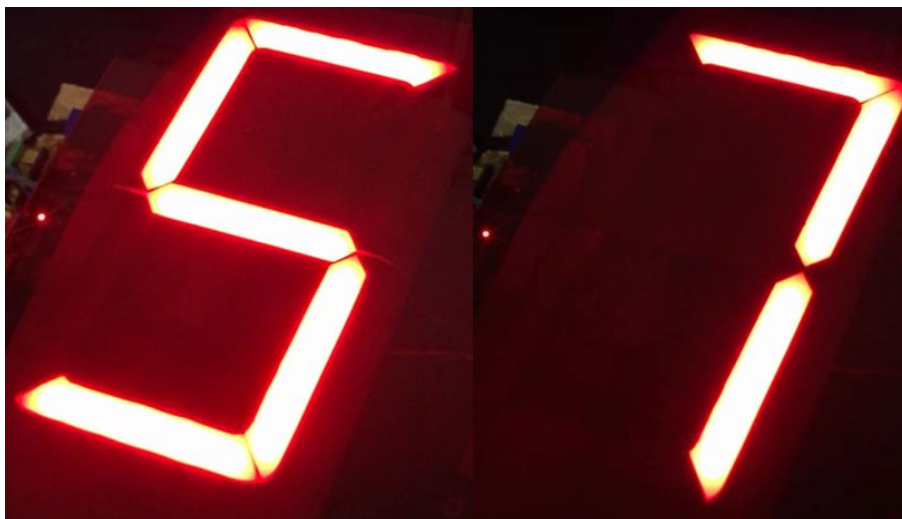
For the final assembly of the display, 3D printed structural parts are designed into which the individual PCBs are built and attached. These assembled parts are installed on a solid board which will be the base of the display screen. Each digit is controlled separately by an I2C bus, over which communication between the display and the main board takes place. Thanks to this topology, we can control each segment separately and there is no need to multiplex the individual digits, which often leads to a reduction in brightness. Each display includes a unit with an I2C communicator and a register to store the current value. The value on the display only changes when it is needed. [2]



**Figure 3:** PCBs of one digit of display with 3D printed body



**Figure 4:** Assembled prototype of one digit (reduced brightness of digit is necessary to take a photo)



**Figure 4:** Assembled prototype of one digit with diffuser from baking paper

#### 4. IMPORTANT TECHNICAL DATA

Max distance of wireless communication between targets and main unit:

-up to 800 m at open space (the distance between the targets and the control unit is about 100 m)

Display data:

One digit (w x h): approximately 300 x 1750 mm

The whole display (2 x 4 digits in two rows): approximately 900 x 700 mm

Power consumption of whole display is about 8 amps (only if the number 8 is displayed on all digits).

#### 5. CONCLUSION

This project is still active, and its development and production of prototypes is underway. All PCBs are now being installed and tested. Thanks to the combination of commercially available timepieces and the addition of new functions, the timekeeper will be developed that will be multifunctional. All commercial timekeepers have their own hardware and software, which in most cases is locked and cannot be modified. Here, thanks to the Arduino platform, it is very easy to edit the code and use the timer, for example, as a score counter, clock, countdown or whatever you can think of. We are limited only by the performance of the Arduino processor, which, however, is quite sufficient for these functions. After its completion in the bachelor's thesis, this timeline will be released on the GitHub portal and will be run as an open-source project that will hopefully contribute to the community around VFDs. If you are more interested in this topic, there is a link below to the semester work on which this article is based.

#### ACKNOWLEDGMENT

This is a non-commercial opensource project that will be published on Github after the completion of the bachelor thesis. The project is being created with the help of my bachelor thesis advisors. I would like to thank ing. Jan Dušek and doc. Jan Mikulka for their help in the design and for their exemplary guidance of my thesis.

#### REFERENCES

- [1] *Pravidla požárního sportu* [online]. Available from: <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjv4Nmuj5b0AhXN0qQKHVgmCdsQFnoECBUQAQ&url=https%3A%2F%2Fwww.hzscr.cz%2Fsubor%2Fpravidla-pozarniho-sportu-pdf.aspx&usg=AOvVaw2qQTgZb0iTgEBEnSn5qFhb>
- [2] DRŠKA, Jan. *Časomíra pro požární sport*. Brno, 2022. Semester thesis

# Sodium-ion batteries: research of carbon-based negative electrode materials

Tomáš Homolka, Jiří Libich

FEEC, Brno University of Technology, Czech Republic

E-mail: [xhomol26@vutbr.cz](mailto:xhomol26@vutbr.cz), [libich@vutbr.cz](mailto:libich@vutbr.cz)

**Abstract** – This paper is mainly focused on research of hard carbon as promising negative electrode material for sodium-ion batteries. The first part deals with batteries and their usage in general. The next section deals with sodium-ion batteries, their development and differences compared to lithium-ion, as they seem to be one of the promising post-lithium technologies with the potential of quick commercialization. The emphasis in this work is put on the negative electrode materials for sodium-ion batteries based on hard carbon as an active electrode material.

**Keywords** – Sodium-ion battery, hard carbon, sodium, carbon, Na-ion, anode, capacity

## 1. INTRODUCTION

Batteries are becoming an essential part of many devices and technologies we use daily. This has been only accelerated by the advent of affordable portable electronics, increasing usage of renewable power sources and electric vehicles. Currently, there are three main technologies. Lead-acid, used mostly in cars and other vehicles to start their internal combustion engine. Nickel-metal hydride, which replaced the older nickel-cadmium, is generally used in some low power devices in form of AA or AAA batteries or in some hybrid electric vehicles. The third one is lithium-ion technology. It is used in a wide range of applications thanks to its high specific capacity and power. It is used in smartphones, drones, laptops, battery-powered electric vehicles and also large-scale grid storage units. However, their biggest problems are their price, limited availability and uneven distribution of lithium throughout the earth's crust, which makes them susceptible to supply disruptions.

## 2. SODIUM-ION BATTERIES

Sodium-ion batteries are based on a similar working principle as lithium-ion ones and have been developed alongside them during the 1970s and 1980s. However, due to the success of graphite as a negative electrode material and the great performance of lithium cobalt oxide, their development was put aside in favor of lithium counterparts. Other reasons for almost abandoning the sodium-ion technology were its lower capacity compared to lithium-ion and the inability to use the same materials, as sodium ions are larger than lithium and graphite shows much lower capacity with them. Researchers started to focus on them again during the 2000s and have continued ever since [1].

## 3. BENEFITS AND DRAWBACKS

As mentioned previously, the main advantage of sodium-ion technology is the price and availability of sodium. Another benefit is the similarity of those technologies, as lithium-ion batteries are well understood and their manufacturing has been very reliable. Most processes can be reused or adapted to be used with sodium-ion. That is except the used materials. Due to the larger size of sodium ions new electrolytes and electrode materials are required. The main drawback is their lower theoretical capacity and incompatibility with most of the previously used electrode materials from lithium-ions [1].

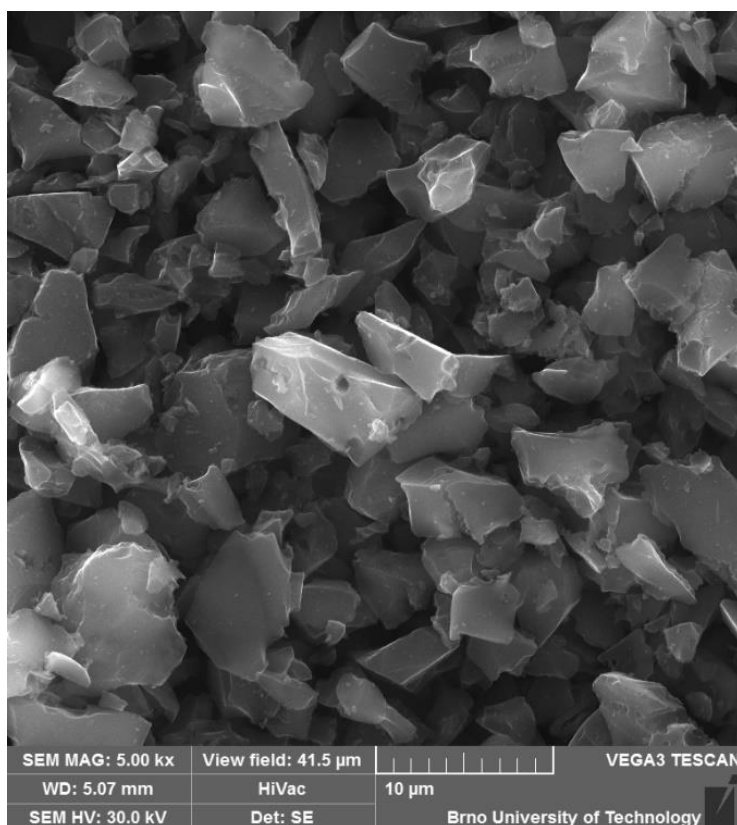
## 4. NEGATIVE ELECTRODE MATERIALS

As previously mentioned for negative electrode materials, the obvious candidate graphite, used in most commercially available lithium-ion batteries, cannot be used with sodium ions due to their bigger size. This led researchers to search for alternatives.

One of them are conversion materials. They exhibit high theoretical capacities, but due to very large volume changes during charge and discharge, the capacity fades rapidly. Alloying-based materials are also being considered, but similarly, they exhibit poor reversibility due to the same large volume changes causing pulverization of the material.

Another option, already successfully used in some commercial lithium-ion batteries, is titanium-based oxides, which use insertion reactions to store sodium ions. They show small volume changes, very good cycling ability, and high rate capability. Their main problem is high working voltage and low specific capacity, which result in low energy density of the full battery.

Carbon-based materials are another alternative, mainly hard carbon and graphene. Graphene is still very expensive to produce, which makes it impractical for commercial use. Hard carbon on the other hand is relatively cheap and shows some impressive results. Its structure is comprised of small graphite-like regions, randomly orientated and disorganized. This results in an uneven surface with many nanopores and places suitable for the insertion of sodium ions, as can be seen in Fig. 1 [2, 3].



**Figure 1:** Structure of used hard carbon under electron microscope

## 5. PREPARATION OF ELECTRODES

A total mass of 0.4 g of the electrode material was prepared, consisting of 10 wt. % binder, in this case, PVDF (polyvinylidene fluoride) was used, and the remaining 90 wt. % hard carbon. 1300 µl of NMP (1-methyl-2-pyrrolidinone) was used as a solvent to dissolve the PVDF and allow it to mix with the hard carbon. The mixing process took about 2 days, to ensure, that the material was uniform enough and ready to spread.

The next step was applying an 80 µm layer of this electrode mass to a sheet of copper foil. This was done using a special K-Hand Coater rod to spread the material evenly. Then it was placed in a dryer at 60 °C for 3 days to dry out all the NMP.

After that, the electrodes were punched out of the foil using a 16 mm diameter punch. Then half of the obtained samples were pressed using pressure of 20 kN for about 6 seconds. This step was skipped for the other half because of concerns that the hard carbon structure might collapse under the pressure and



lower the useable capacity. Subsequently all the electrodes were placed in a vacuum dryer at 55 °C to ensure no residual water was left.

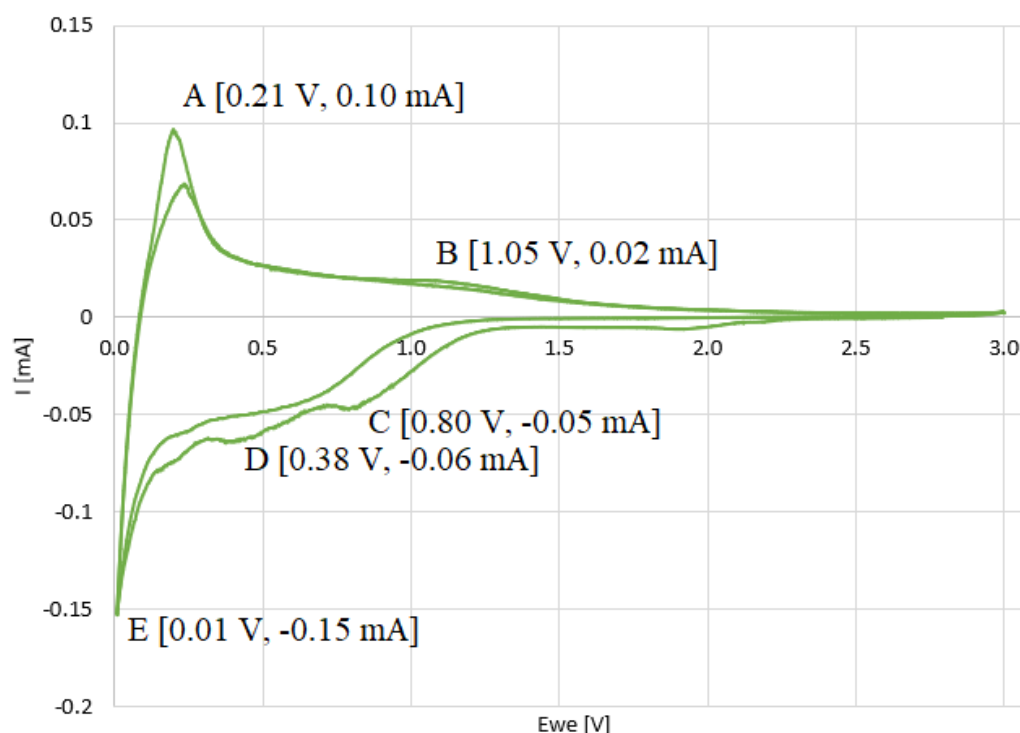
As an electrolyte, a mixture of 50 wt. % EC (ethylene carbonate) and 50 wt. % PC (propylene carbonate) was prepared. A total amount of 5 ml of 1 molar solution with the chosen sodium salt  $\text{NaPF}_6$  (sodium hexafluorophosphate) was made.

## 6. ASSEMBLING THE TESTING CELL

The assembly of the testing cell was carried out under a protective atmosphere of argon to prevent contamination and oxidation of the electrolyte and sodium. First layer was the 16 mm diameter circle of metallic sodium, then 18 mm circular sheet of separator and electrolyte. Specifically 170  $\mu\text{l}$  of the EC:PC mixture previously mentioned. On top was placed one of the prepared electrodes. The weight of the copper current collector was determined beforehand to calculate the amount of the active material. The completed electrode was then weighed and the mass of the hard carbon was determined to be 2.61 mg. The cell was the hermetically closed using a lid, to allow it to be taken out of the argon atmosphere.

## 7. MEASUREMENTS

Before the measurements, the cell was left for about 24 hours to stabilize. After that, it reached a voltage of about 2.75 V. The measurements were performed using a Biologic VMP 16-channel potentiostat. It was set to cyclic voltammetry (CV), with voltage limits of 5 mV and 3 V. Several peaks were observed on the resulting data, as can be seen in Fig. 2. They correspond with the expected values based on other research. The peaks A and E represent the reversible insertion of  $\text{Na}^+$  ions into the nanopores of the material. Peaks C and D appear only on the first cycle and are attributed to the formation of a solid electrolyte interface (SEI) layer. Peak B might be attributed to intercalation of sodium ions into the carbon layers. [4]



**Figure 2:** Results of the cyclic voltammetry with marked peaks

## 8. CONCLUSION

The goal of this work was to check the properties of hard carbon as negative electrode material, as it seems to be one of the promising candidates for commercial sodium-ion batteries. Only a few of the first measurements were completed, but even they show some of the basic properties of hard

carbon. The plan is to also test a mixture of hard carbon and super P to make the resulting material more conductive and to try and compare the properties of the pressed electrodes and those that weren't. Testing of hard carbon from another manufacturer is also planned to compare the results, as the properties of hard carbon depend on the process of its making and its precursor.

## ACKNOWLEDGMENTS

This work was supported by specific graduate research project of the Brno University of Technology: Materialy a technologie pro elektrotechniku IV., reg. no. FEKT-S-20-6206. Thanks for the supervision of Ing. Jiri Libich, Ph.D.

## REFERENCES

- [1] A. Eftekhari and D. -W. Kim, "Sodium-ion batteries: New opportunities beyond energy storage by lithium", *Journal of Power Sources*, vol. 395, pp. 336-348, 2018.
- [2] L. -F. Zhao, Z. Hu, W. -H. Lai, Y. Tao, J. Peng, Z. -C. Miao, Y. -X. Wang, S. -L. Chou, H. -K. Liu, and S. -X. Dou, "Hard Carbon Anodes: Fundamental Understanding and Commercial Perspectives for Na-Ion Batteries beyond Li-Ion and K-Ion Counterparts", *Advanced Energy Materials*, vol. 11, no. 1, 2021.
- [3] D. A. Stevens and J. R. Dahn, "High Capacity Anode Materials for Rechargeable Sodium-Ion Batteries", *Journal of The Electrochemical Society*, vol. 147, no. 4, 2000.
- [4] M. Wahid, Y. Gawli, D. Puthusseri, A. Kumar, M. V. Shelke, and S. Ogale, "Nubby Carbon: Morphology Replicating Hard Carbon from Walnut Shell for Na Ion Battery Anode", *ACS Omega*, vol. 2, no. 7, pp. 3601-3609, Jul. 2017.

# Transparent and conductive titanium nitride thin films for implantable bioelectronics deposited at low temperature using two Kaufman ion-beam sources

J. Jarušek<sup>1</sup>, I. Gablech<sup>1,2</sup>

<sup>1</sup>Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Microelectronics, Technická 3058/10, 61600 Brno, Czech Republic

<sup>2</sup>Central European Institute of Technology, Brno University of Technology, Purkyňova 123, 612 00 Brno, Czech Republic

E-mail: [Jaromir.Jarusek@vut.cz](mailto:Jaromir.Jarusek@vut.cz), [imrich.gablech@ceitec.vutbr.cz](mailto:imrich.gablech@ceitec.vutbr.cz)

**Abstract**— This work proposes ultra-thin films of titanium nitride as a material for the realization of transparent and conductive layers suitable for implantable bioelectronics. Results further show that good results of crystallography, transmittance, and sheet resistance can be achieved for thin films with thicknesses of 15 nm and 20 nm deposited at the beam energy of the primary ion source of 400 eV and temperature below 100 °C.

**Keywords**—titanium nitride, transparent conducting film, sputtering, bioelectronics implants, thin films, visible transmittance

## 1. INTRODUCTION

Implantable bioelectronic devices place special requirements on the materials used. The main requirements are biocompatibility and long-term stability for chronic implants. Several ways for the preparation of transparent conductive thin films are proposed. One of the conventionally used biocompatible materials is gold. However, gold electrodes, compared to titanium nitride electrodes, are prone to anodic dissolution and protective layers are required to prevent this [1]. In addition, the deposition of continuous conductive ultrathin films ( $\approx 20$  nm) for transparent layers is problematic on many substrates because electrically insulated islands are formed during deposition, and their elimination requires non-conventional substrate materials, higher deposition temperatures, or subsequent thermal annealing which is not possible for most of organic substrates/underlayers [2] [3].

Other ways include the use of polymeric nanocomposites filled with conductive materials such as metals, carbon nanotubes, or graphene oxide. However, many of the biopolymers studied show good degradability, which makes them suitable for short-term, not long-term applications [4]. In addition, the materials prepared in this way require careful control of the filler dispersion to ensure homogeneous properties. Other polymeric materials, such as poly(3,4-ethylenedioxythiophene) (PEDOT), exhibit various problematic properties such as oxidation and/or reduction in a physiological environment, acidity, hygroscopicity, low durability, etc. [5], thus they are not suitable for chronic implants.

On the other side, titanium nitride possesses very good *in-vivo* long-term stability [6], potentially low sheet resistance ( $R_{\square}$ ) and high transmittance are expected, and lower costs than indium tin oxide. Therefore, this paper proposes ultrathin, transparent, and conductive titanium nitride thin films for implantable bioelectronic devices.

## 2. EXPERIMENTAL DETAILS

The sputtering system consists of two Kaufman ion-beam sources (IBS). The IBS aimed at the target of the sputtered material is noted as the primary IBS. The second IBS is aimed at the substrate, this source is noted as secondary IBS for ion-beam assisted deposition (IBAD). Additionally, secondary IBS was used for in-situ precleaning using Ar ions. All samples were cleaned using an ultrasonic bath with acetone, followed by isopropyl alcohol, and finally dried by nitrogen and then treated in oxygen plasma. No additional heating was used during the depositions, thus temperatures were below 100 °C.

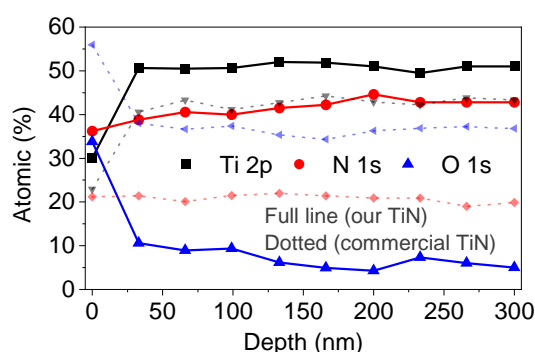
The deposition of titanium nitride thin films was performed in three steps. In the first step, films with an estimated thickness of 300 nm were deposited on 1-inch silicon wafers at different ion-beam energies of the primary source in the range from 400 eV to 1200 eV with the step of 200 eV. For the depositions at given energies, the values of the argon and nitrogen flow rates of the primary ion source were adjusted in a 5.5:4.5 ratio according to previous experiments [7], while the values of the deposition times were roughly estimated. Values of flow rates Ar and N<sub>2</sub>, beam current, deposition rates, and deposition times are in Table . These deposited films were then partially masked using drop-casting of AZ-5214E photoresist and then baked. Several etching methods were then tried. Wet etching mixtures such as buffered oxide etch (BOE) or HF with HNO<sub>3</sub> were not able to etch titanium nitride. It is assumed that this is caused by the low content of oxygen compared to commercially available TiN, as is shown in Figure 1. It can be also assumed that films prepared by IBAD are more chemically resistant compared to evaporation or magnetron sputtering methods. Better etching results were obtained using reactive ion etching (RIE) with a 1:5 ratio of BCl<sub>3</sub> and Cl<sub>2</sub>. The thickness of the deposited films was then measured using a Bruker Dektak XT mechanical profilometer with a 2  $\mu$ m radius tip. Average deposition rates were determined from measured values.

In the second step, the deposition times were set according to known deposition rates to reach the desired thickness of  $\approx$  200 nm which is sufficient for materials analyses such as X-ray-diffractometry (XRD). These thin films were then also characterized by a four-point probe method using a Keithley 4200-SCS Parameter Analyzer and a Rigaku SmartLab XRD. For verification, these films were also etched using RIE, and their thicknesses were measured using a profilometer.

In the last step of the experiment, thin films with thicknesses ranging from 5 nm to 50 nm were deposited on microscope slides using ion-beam energy of 400 eV. Deposition times were set based on the verified deposition rate from the second part of the experiment. These samples were then characterized by the four-point probe method, XRD and VIS/NIR Ocean optics NIRQuest 512 optical spectrometer.

**Table I:** Deposition parameters for set primary ion-beam energy

Primary ion-beam energy (eV)	400	600	800	1000	1200
Ar flow (sccm)	3.0	3.6	4.1	5.0	5.5
N <sub>2</sub> flow (sccm)	2.5	3.0	3.4	4.0	4.5
Beam current (mA)	30	44	55	65	80
1st experiment deposition time (s)	30000	24801	20000	10000	5000
1st experiment deposition rate (nm/s)	0.0149	0.0161	0.0181	0.0250	0.0242
2nd experiment deposition time (s)	13452	12400	11050	8000	3472
2nd experiment deposition rate (nm/s)	0.0074	0.0114	0.0181	0.0250	0.0432



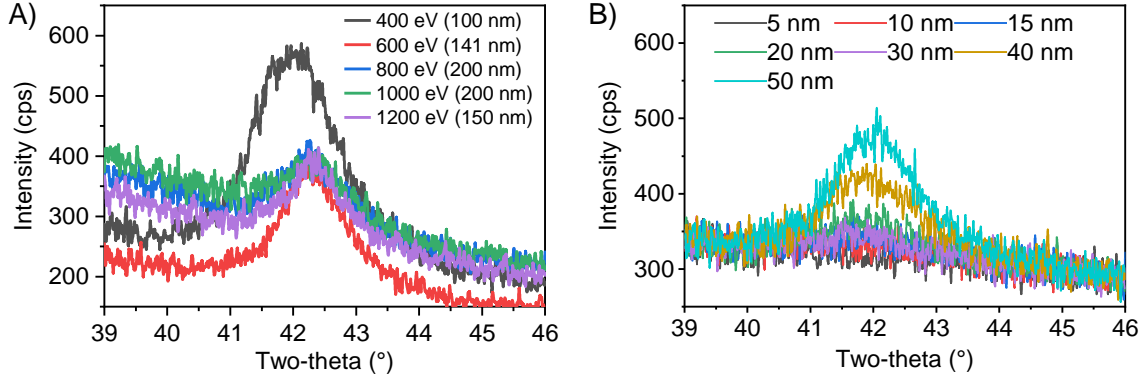
**Figure 1:** Comparison of atomic percentages of prepared TiN and commercially available TiN using X-ray Photoelectron Spectroscopy.

### 3. RESULTS AND DISCUSSION

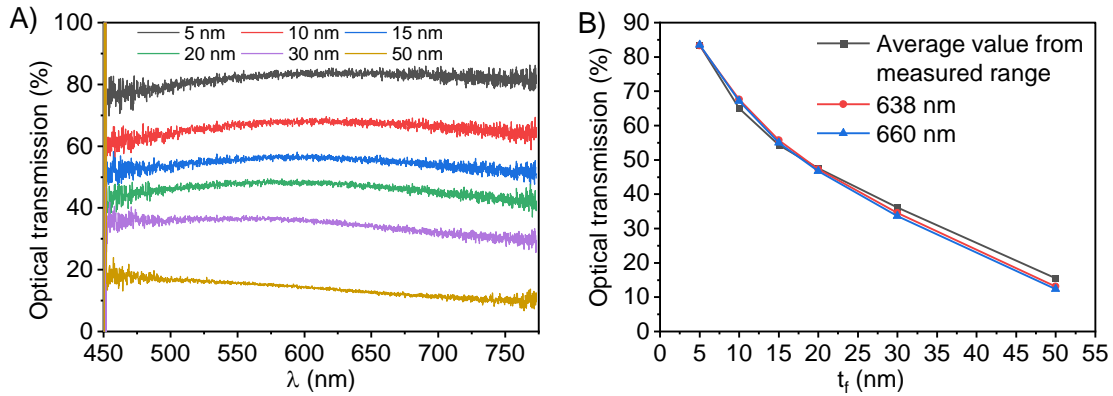
The results from XRD measurements of the film deposited on silicon wafers in Figure 2A show a significant dependence between the energy of primary IBS and the crystallinity of films. TiN (200) peaks are clearer for films deposited at lower energies. Deviation of film thickness at energies 400 eV and 600 eV were caused by difficulties in the first etching attempt using wet etching mixtures, which caused incorrect results from the profilometer and consequently incorrectly determined

deposition rate. Despite this issue, a process with 400 eV beam energy was chosen for the following experiments due to a clearly better degree of crystallinity and expected better electrical and optical properties.

Figure 2B shows XRD results of thin films with various thicknesses in the range from 5 nm to 50 nm, deposited on microscope slides. Although, thin films with thicknesses of 40 nm and 50 nm have clearly visible peaks. It can be estimated that all films have a crystal orientation (200), with only a slight deviation from films deposited on silicon wafers. The low diffraction intensity is caused due to very low thickness of prepared layers and lattice mismatch between substrate and thin film.



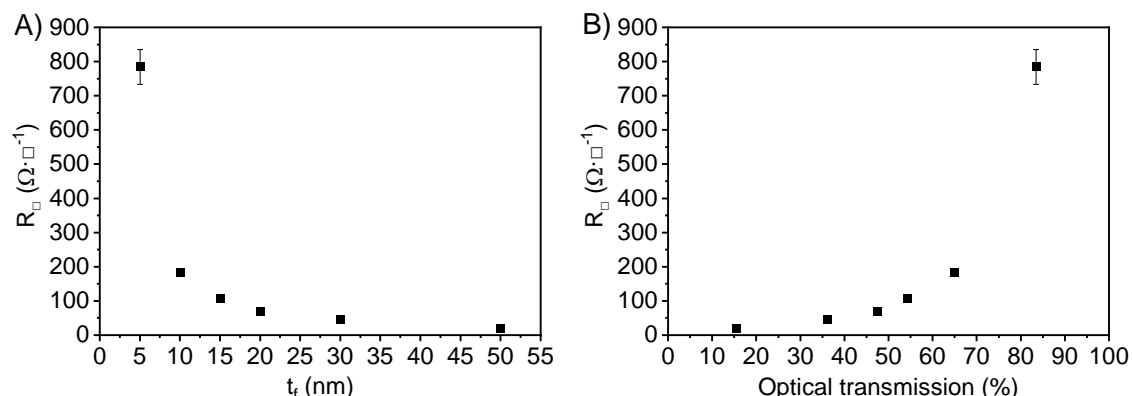
**Figure 2:** A) TiN (200) XRD peaks for the films deposited at beam energies between 400 eV and 1200 eV. B) TiN (200) XRD peaks for the films deposited on microscope slides at a beam energy of 400 eV.



**Figure 3:** A) Optical transmission of films with thickness in the range of 5 nm and 50 nm deposited at microscope slides. B) Corresponding relation between average optical transmission, transmission at wavelengths of 638 nm and 660 nm, and film thickness.

Results of optical transmission in the visible spectrum in Figure 3A show interesting results for film thicknesses in the range from 5 nm to 20 nm. These films show roughly even transmittance. Films of thicknesses 30 nm and 50 nm show a slight tendency to block longer wavelengths. It is estimated that this is due to applying reflectance. Comparing optical transmittance of the film thickness 50 nm at shorter wavelengths and longer wavelengths, it can be evaluated that film of this thickness is not suitable for infrared and near-infrared applications. Figure 3B shows average optical transmission and transmission at wavelengths of 638 nm and 660 nm. These wavelengths were chosen as the nominal wavelengths of LEDs used in related research [8]. It can be evaluated that for film thickness above 20 nm transmittance decreases approximately linearly. For film thicknesses below 20 nm, optical transmission is higher than could be extrapolated from greater thicknesses. According to these results, films with thickness 20 nm and lesser are promising for practical use. Figure 4A shows that sheet resistance for films of thicknesses below 20 nm is highly disproportional to thickness due to changing electrical resistivity. Because of this significant dependency of sheet resistance on the film thickness, the use of films with thickness 10 nm and lesser is not practical and not even beneficial considering the only small increase in optical transmission compared to the increase in sheet resistance shown in Figure 4B.

By considering the results mentioned above it can be evaluated that films with thicknesses of 15 nm and 20 nm show interesting results, and these thicknesses deposited at beam energy 400 eV have promising properties for practical use.



**Figure 4:** A) Sheet resistance as a function of film thickness. B) Sheet resistance as a function of optical transmission.

#### 4. CONCLUSION

From the above results, it can be concluded that films with thicknesses of 15 nm and 20 nm deposited at beam energy 400 eV have promising properties for practical use. For further experiments, the research will be focused on the optimization of process parameters, such as ion-beam optics and deposition temperature, to achieve lower values of  $R_{\square}$ .

#### ACKNOWLEDGMENT

CzechNanoLab project LM2018110 funded by MEYS CR is gratefully acknowledged for the financial support of the measurements/sample fabrication at CEITEC Nano Research Infrastructure.


#### REFERENCES

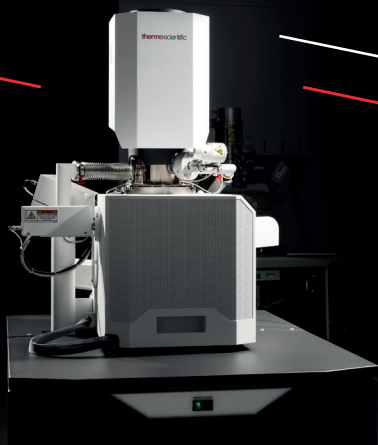
- [1] S. Szunerits, Y. Coffinier, S. Janel and R. Boukherroub, "Stability of the Gold/Silica Thin Film Interface: Electrochemical and Surface Plasmon Resonance Studies", *Langmuir*, vol. 22, no. 25, pp. 10716-10722, 2006.
- [2] A. Kossoy, V. Merk, D. Simakov, K. Leosson, S. Kéna-Cohen and S. Maier, "Optical and Structural Properties of Ultra-thin Gold Films", *Advanced Optical Materials*, vol. 3, no. 1, pp. 71-77, 2015.
- [3] K. Leosson, A. Ingason, B. Agnarsson, A. Kossoy, S. Olafsson and M. Gather, "Ultra-thin gold films on transparent polymers", *Nanophotonics*, vol. 2, no. 1, pp. 3-11, 2013.
- [4] C. Wang, T. Yokota and T. Someya, "Natural Biopolymer-Based Biocompatible Conductors for Stretchable Bioelectronics", *Chemical Reviews*, vol. 121, no. 4, pp. 2109-2146, 2021.
- [5] J. Luceño Sánchez, R. Peña Capilla and A. Díez-Pascual, "High-Performance PEDOT: PSS/Hexamethylene Diisocyanate-Functionalized Graphene Oxide Nanocomposites", *Polymers*, vol. 10, no. 10, 2018.
- [6] H. Hämmerle, K. Kobuch, K. Kohler, W. Nisch, H. Sachs and M. Stelzle, "Biostability of micro-photodiode arrays for subretinal implantation", *Biomaterials*, vol. 23, no. 3, pp. 797-804, 2002.
- [7] I. Gablech, V. Svatoš, O. Caha, A. Dubroka, J. Pekárek, J. Klempa, P. Neužil, M. Schneider and T. Šíkola, "Preparation of high-quality stress-free (001) aluminum nitride thin film using a dual Kaufman ion-beam source setup", *Thin Solid Films*, vol. 670, pp. 105-112, 2019.
- [8] M. Silverå Ejneby, M. Jakešová, J. Ferrero, L. Migliaccio, I. Sahalianov, Z. Zhao, M. Berggren, D. Khodagholy, V. Derek, J. Gelinas and E. Głowacki, "Chronic electrical stimulation of peripheral nerves via deep-red light transduced by an implanted organic photocapacitor", *Nature Biomedical Engineering*.



# Naše práce je věda

V našem brněnském technologickém centru vznikají špičkové elektronové mikroskopy a spektrometry, které dodáváme do celého světa. Studují se jimi viry, vznikají díky nim vakcíny, vyvíjí se lepší materiály i elektronika. Pracujeme se špičkovými technologiemi, které posouvají lidské poznání. Najdeš mezi námi odborníky na fyziku, elektroniku, software, mechanickou konstrukci nebo logistiku. Chceš toho být součástí?

 [thermofisher.jobs.cz](https://thermofisher.jobs.cz)



Brno, Vlastimila Pecha 12

# Batch process in connection with industry 4.0

J. Boch<sup>1</sup>, M. Husák<sup>2</sup>

Brno University of Technology, Department of Control and Instrumentation, Czech Republic

E-mail: [xbochj00@vutbr.cz](mailto:xbochj00@vutbr.cz)<sup>1</sup>, [xhusak08@vut.cz](mailto:xhusak08@vut.cz)<sup>2</sup>

**Abstract**—This project deals with the creation of autonomous liquid storage cell with a cooling system. The cell is part of a larger project called the Self-Acting Barman. Project Barman serves to demonstrate the principles of Industry 4.0 and is gradually made up of student works. The aim of the project is to create a cell that can dispense the required amount of fluid. The cell is cooled by an external cooling cell. The entire liquid batch process is controlled by a PLC located at the cell interface. By default, the process would be controlled by a superior batch system that complies with the ISA-S88 standard. The paper aims to clarify the fusion of the S88 standard with the new 4.0 approach.

**Keywords**— Industry 4.0, Self-Acting Barman, Testbed, Autonomous Cell, Batch control, Asset Administration Shell

## 1. THEORETICAL INTRODUCTION

This chapter deals with a brief description of the term industry 4.0 and its basic principles. The Self-Acting Bartender project is also described here.

### 1.1. Industry 4.0

The term Industry 4.0 is a name for another industrial revolution that is currently underway. The main idea of this revolution is to create a global network for the entire factory, which will include all machines, storage facilities, production systems and security systems, as one large cyber-physical system.

These factories are called Smart Factories. Smart factories produce products that are easily identifiable, can be located in every part of the production process and know their history, current state and the way to achieve the desired state.[1] The entire journey of the product is documented by its assignment to production to distribution logistics.

### 1.2. Principles of industry 4.0

Industry 4.0 uses several basic principles such as: Virtualization, Decentralization, Modularity, Reconfigurability and Interoperability[2].

**Virtualization** means converting the entire production process (in our case an autonomous cell) into digital form. The whole digital model can be created to work exactly like the real model. This model is called the digital twin. The functionality of the entire proposed solution can be tested on a digital model before the start of the physical creation. This allows us to quickly detect errors and shortages of the proposed solution and thus avoid the production of non-functional prototypes. The advantage of digital twin is the simple replacement of components, actuators and all other elements if we find that they do not suit us. Cell virtualization is created in the Siemens NX design environment.

Each component of Industry 4.0 is in a kind of envelope called an Asset administration shell (AAS). AAS covers the entire component and creates an interface between the physical and software part. AAS is basically a standardized digital representation of a component (digital twin). AAS is the core of Interoperability. The body of the AAS contains certain structured information in the form of submodels for the specification of properties, parameters, functions given component[3].

**Decentralization** means a situation where each part of the production process requests the necessary information and materials. The product itself carries information about what it needs for its production

and passes it on in individual parts of the production process. The individual parts of the production process communicate with each other without the help of any higher layer. We call such a network IoT.

**Modularity** means that the individual parts of the production process are physically independent of each other. These individual modular parts are interconnected by robotic arms, conveyors and other manipulators.

**Interoperability** means the ability to connect all parts of the production process (machines and workers) into one system, where all parts can communicate with each other. In our case, this is provided by the AAS.

### 1.3. Self-Acting Barman

The testbed Barman is divided into several cells according to the principles of Decentralization and Modularity. The cells are a glass storage, a soda maker, a shaker, an ice dispenser, a storage of alcoholic liquids and a coolant cell. All cells except the storage of alcoholic liquids are the same size, have same connectors and the user interface according to the principle of reconfigurability.

In the middle of the testbed is a robotic manipulator that moves the glass between the individual cells. The last part of the testbed is a conveyor belt, which is used to take already made drinks.

## 2. PROJECT CONCEPT

The hearth of the cell is a stainless steel tank, which consists of four separate containers for liquids. The containers are not closed from the top and there is an outlet at the bottom. Around these containers is a hollow tank housing that serves as a space for cooling liquid. The whole design of the cell is based on this tank.

### 2.1. Cell concept

The key idea of the liquid storage cell concept is pumping the liquid from tank containers. The liquid flows down into the hose to the pump which drives the liquid back up over the glass into which it falls. A peristaltic pump was chosen for this purpose. The main advantage of peristaltic pumps is that they are relatively easy to manufacture using parts printed on a 3D printer, bearings and hoses. The peristaltic pump is driven by a stepper motor. This gives us the possibility to change the direction of liquid pumping in case of leakage of redundant liquid.

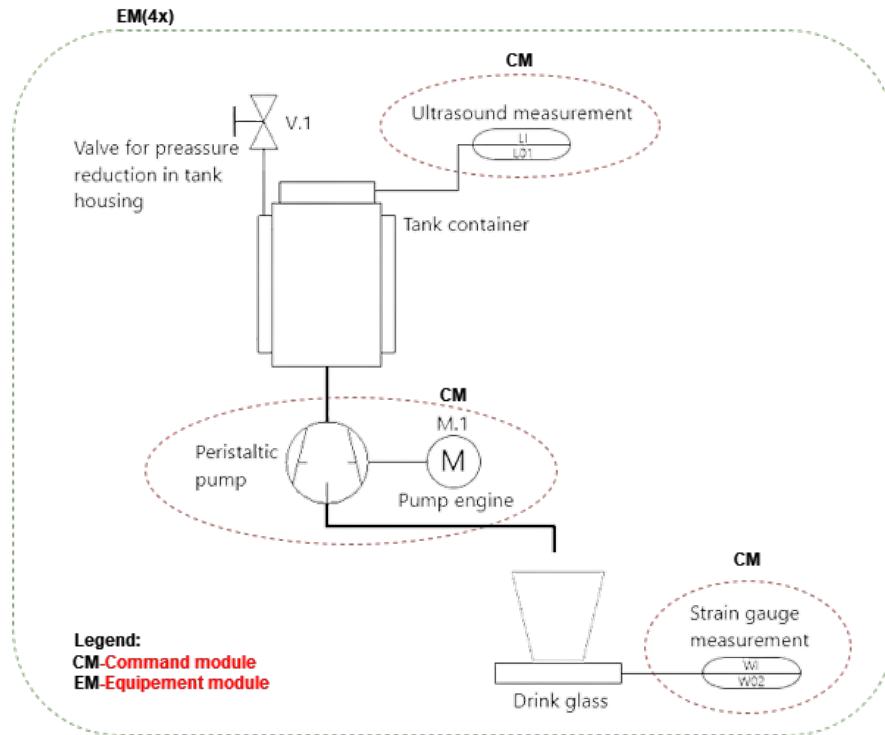
### 2.2. Cell instrumentation

There are two types of sensors in the cell for measuring quantities. The first are ultrasonic sensors which are used to obtain information about the height of liquids in the containers of the tank. This sensor is located in each cover that closes the top of the tank container. The second type of sensor is a strain gauge. Sensor is located under the glass holder and is used to determine the weight of the liquid already pumped in the glass. The amount of fluid pumped can be also calculated with knowledge of number of stepper motor steps. All sensors are connected to a universal PCB board called SKUseCon. It is used to connect the sensor, process the signal from the sensor and then send the signal to the PLC using output pins.

Two different pump models were purchased and a third model was manually created on the 3D printer to select the correct peristaltic pump. These three models were then compared by the fluid pumping speed and the fastest of them was selected for use. The pumps are driven by a NEMA17 stepper motor. These are connected to the TB6600 4A drivers. Drivers can be used to set the amount of current supplying the motor and the number of microsteps of the stepper motor. The driver is then connected to the PLC.

Every element in the cell is connected to the Siemens PLC S7-1200. PLC can then be commanded from HMI Panel KTP400 Basic. These are connected together in switch which is located under the HMI.

At the top of the cell is a three-colored beacon that shows the current state of the cell. There is also a emergency stop button.



**Figure 1: P&ID diagram**

### 3. BATCH PROCESS

The batch process is standardized according to the standard ANSI/ISA-S88. This standard divides our technology into certain parts by three models.

The first part is a physical model. The physical model is used to describe the equipment needed to make the product. Specifically, our physical model begins with a unit (our cell). The unit is further divided into individual device modules. There are four of them in our case. It is a one tank container together with a pump, motor, ultrasonic sensor and a common strain gauge. The lowest part of the structure is the control module. This is, for example, a separate sensor or a pump.[4]

The second part is a procedural model. The procedural model combines the physical model and process model and describes the hierarchy of functions that needs to be done during batch production. The most important part for us in this model is phase. Phase is connected to the device module. In our case, it will be, for example, the liquid pumping phase or tank cleaning.[4]

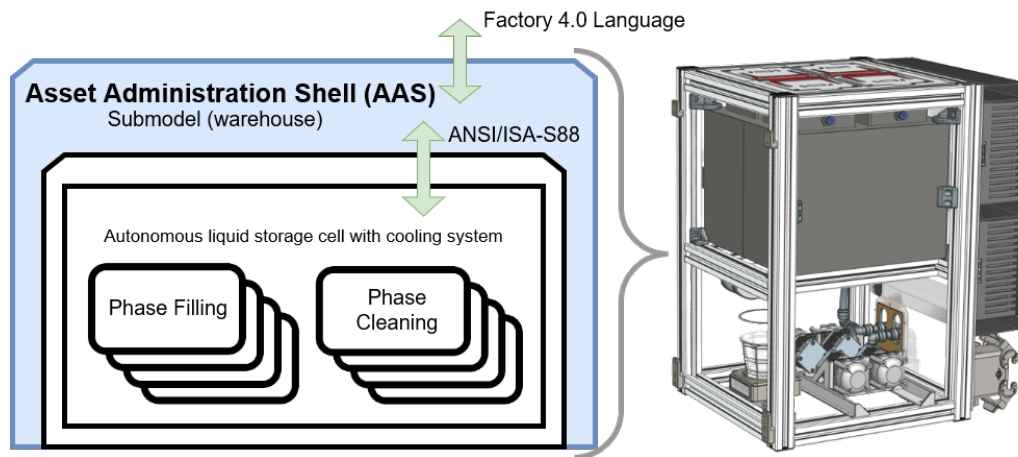
The third part is a process model. The process model represents the product production process.[4]

#### 3.1. PackML standard

Another similar standard to the batch process is a PackML standard. PackML standard was created for better implementation of the packing robots into the industry. PackML standard was based on the ISA-S88 standard. That's why the packaging line hierarchy has a lot of similarities. The batch process standard is more suitable in our case.

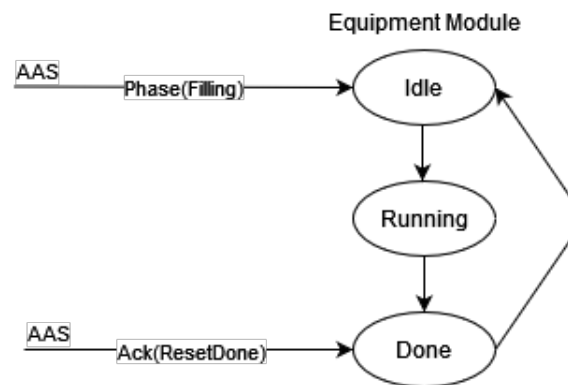
### 4. AAS TO BATCH

The physical model part of the batch control standard is cut between the unit(our cell) and the process cell. On the procedural level, the highest part is the phase. Unit is then covered in the AAS which makes the connection between the Batch control standard and the Industry 4.0 system. These shells communicate with each other without other superior system. The AAS compares their capabilities to the required service and decides to perform the phase.



**Figure 2:** Aset administration shell

The phase is connected to the equipment module which contains all the necessary control modules for successful completion of the phase. AAS gives the first signal to start the phase. Signal from the AAS also contains all parameters about the volume of batch. State machines then go from state idle to state working. After completion of the phase, state machine moves to state done. AAS then obtains the information about the state and resets the state machine to state idle.



**Figure 3:** State Machine of the Phase

The main difference between the batch system and the factory 4.0 system is that the batch system is controlled by a superior layer (process cell), which has the whole system under control. The Factory 4.0 system is not controlled centrally from the superior unit, but all cells are able to communicate with each other using AAS envelope (Interoperability). In case of need, the AAS envelope can be changed with process cell to control the whole system very easily.

## 5. CONCLUSION

The autonomous cell created meets all the requirements to demonstrate the principles of industry 4.0. This work shows a real connection between the batch process and elements of industry 4.0. AAS makes the connection between the batch process system and the factory 4.0 ecosystem. Batch process system ends on the unit level in the physical model and on the phase level in the procedural model. The phase is then being launched by the AAS and the state machine of the process works as shown in figure 3. The AAS serves as a translator between the recipe procedure, which is included in the product image, and the phases of the production unit.

The aim of this project was to create an autonomous cell that will operate according to the principles of Industry 4.0. The software ends at the phase level, which includes the state machine as shown in Figure 3. Assigning an AAS to a cell is part of another project.

## ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-20-6205 –”Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0” financially supported by the Internal science fund of Brno University of Technology.

## REFERENCES

- [1] M. Hermann, T. Pentek and B. Otto, “Design Principles for Industrie 4.0 Scenarios”, A Literature Review, 2015, DOI:[10.13140/RG.2.2.29269.22248](https://doi.org/10.13140/RG.2.2.29269.22248).
- [2] V. Kaczmarczyk, O. Baštán, Z. Bradáč and J. Arm, ”An Industry 4.0 Testbed (Self-Acting Barman): Principles and Design,” IFAC-PapersOnLine [online]. 2018, 51(6), 263-270 [cit. 2021-03-07], ISSN: 24058963.
- [3] Z. Bradáč, P. Marcon, F. Zezulka, J. Arm and T. Benešl “Digital Twin and AAS in the Industry 4.0 Framework”, IOP Conference Series: Materials Science and Engineering. 618. 012001 DOI:[10.1088/1757-899X/618/1/012001](https://doi.org/10.1088/1757-899X/618/1/012001).
- [4] M. Husák, “Návrh a realizace laboratorní úlohy na téma ”Standard ISAS88”, Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2018. 51 s. Vedoucí bakalářské práce Ing. Václav Kaczmarczyk, Ph.D.. [Bachelor thesis](#).



# Testbed for virtual commissioning of PLC system

**Jakub Nosek, and Jakub Arm**

Brno University of Technology, Czech Republic

E-mail: [221007@vut.cz](mailto:221007@vut.cz), [jakub.arm@vut.cz](mailto:jakub.arm@vut.cz)

**Abstract**— The work deals with the creation of a testbed for the virtual commissioning of a PLC system. The testbed consists of Unity, a virtual PLC, a Rest API server and a C # application. The production process is simulated at Unity using a Siemens virtual PLC. The data is sent from Unity in JSON format to the Rest API server, where it is read by a C # application and the program is modified if necessary. After editing, the simulation restarts. The testbed is intended to optimize the PLC application, i.e. crisis treatment, collision avoidance and line tact reduction.

**Keywords**—PLC, C#, Unity, JSON, Simulation

## 1. INTRODUCTION

The testbed containing a virtual commissioning PLC system is to be used for automated testing and even debugging of PLC applications. The data for PLC program repairs are taken from a dynamic simulation of real technology, which is connected to a virtual PLC. At this point, we assume that we will optimize the following aspects of the PLC program:

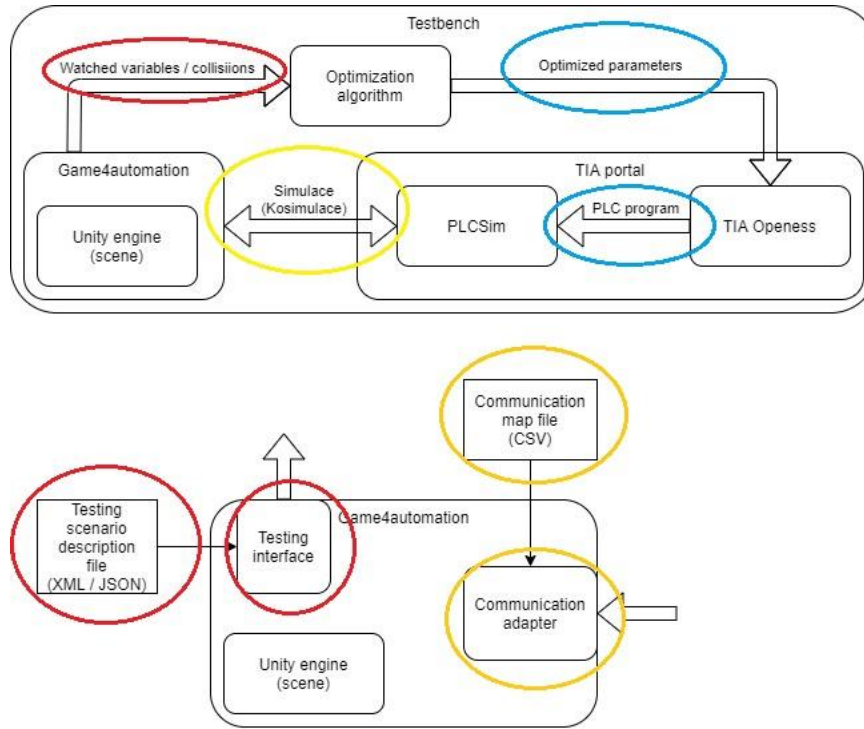
- timer values
- collision treatment of technological parts
- line cycle time.

The work focuses on the reworking of the project from the subject BPC-PPA by Dr. Arm. The original intention was to program a virtual PLC using Codesys software, because it is free and contains a virtual PLC. The aim of the simulation was to use a robot to place six boxes on the main conveyor, where the boxes were filled with cans and then placed on a shelf according to the student's ID. The project was available to students as an exe file.

In [1], the process of optimizing the production process based on video analysis is presented. In our case, however, we proceed by first creating a credible virtual model of the production process (the so-called passive digital twin) and modifying the PLC application according to the operation in the virtual environment.

Mathematical technology models can be used to verify the correctness of the PLC application. However, this approach requires ensuring the credibility of such a model. In addition, it also suffers from problems arising from nondeterministic communication between the virtual PLC and the model, as stated [2]. Our approach uses a dynamic physical model, which also takes into account the physical properties of objects (mass, friction, inertia).

## 2. TESTBED DESIGN



**Figure 1: Schema of testbed**

The task of the test interface will be to transfer the monitored parameters to the newly created application, where the data will be processed and then applied to the PLC.

AutomationML will be used to describe the testing scenario, according to which the individual industrial components, their behavior and interconnections will be described.

The Rest API server working with the JSON data format was chosen for the communication of the test application with the simulation in Unity for the purpose of data sharing. The principle of data collection works by storing the sample time, all inputs and outputs between the PLC and the simulation, the number of collisions and information about all collisions in the given frame (collision position and names of collided objects). This data is saved as a new element in the List. After the selected time in seconds, data is uploaded to the API server, which contains information that a new, possible restart of the simulation from the test application, the number of elements in the List and the List itself, which has the previously mentioned information about individual frames.

The task of the C # application is to process data from the simulation in Unity and subsequently rework the program and its parameters using an optimization algorithm. Then it starts the simulation again with the new parameters.

According to tests, Modbus TCP will be used as a communication adapter, for better response. In case of problems, it will be a backup variant S7 with a real PLC.

## 3. IMPLEMENTATION

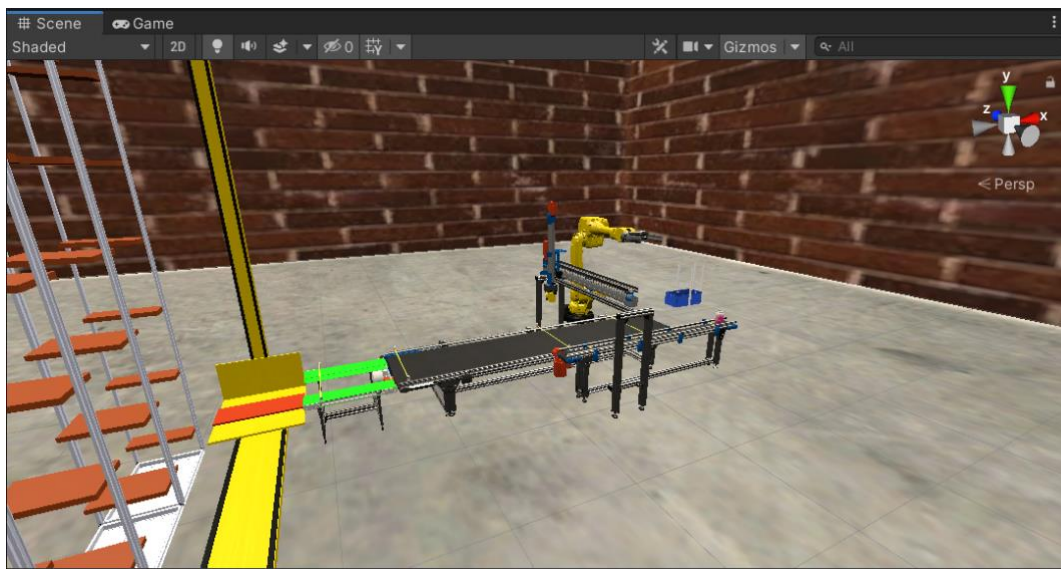
Codesys software has been replaced by TIA Portal V16. Codesys SoftPLC communicated with Unity via Modbus TCP at IP address 127.0.0.1, which is the computer's local address on localhost. When programming in the TIA Portal, it was not possible to set the virtual PLC in the PLCSIM V16 software to the IP address 127.0.0.1, because Siemens does not allow this address, as Codesys. In general, it is not possible for the basic PLCSIM V16 program to communicate with the Modbus TCP communication protocol because it does not have the required virtual adapter. Therefore, PLCSIM Advanced, which includes this feature, was used. A project with a virtual PLC SIMATIC S7 / 1512C-1 PN was created for this solution in the TIA Portal. Subsequently, the IP address in Unity changed. Unfortunately, this was not enough and the code for Modbus TCP had to be modified. The original version of Modbus TCP read from the Holding registers and the Input registers. Codesys can distinguish this, but TIA Portal cannot and uses everything as Holding registers. The Modbus TCP code has changed that Unity will

process everything as Holding registers, only the first 10 Words will be taken as Unity output and the next 10 Words as Unity input.

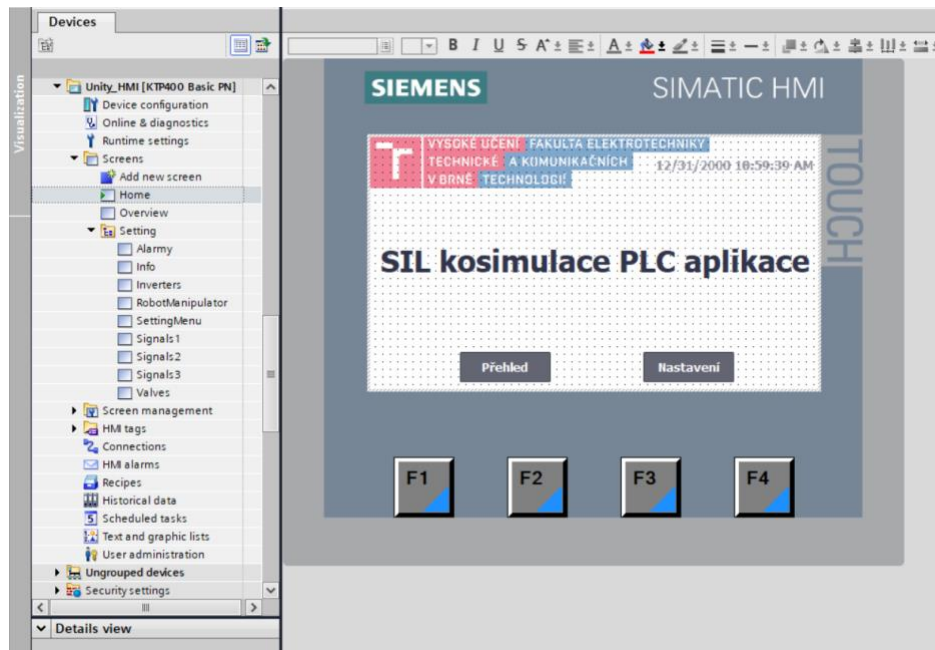
Two communication protocols were considered for the testbed. With virtual PLC Modbus TCP and with real PLC S7. The response was measured in Unity for both communication protocols. The code was modified in the script to save the results to a text file. A string variable was created in the InterfaceThreadedBaseClass script, which stores the response time from the CommCycleMs variable as a new line each time the program iterates. This recording takes place only after the tenth iteration due to the fact that the first measured values may not be valid due to the loading of the environment. The measurement starts at the beginning of the simulation. In 30,000 iterations, the measurement results are saved in a text file. As a result, there are 29,899 stored measurement results. At the end of the text file, the minimum and maximum of the measured values, the average value and the variance are found. The minimum and maximum values found were found during the program run. The individual measurement values in each iteration were stored in the temporary field, and the variance and average value were calculated in the last iteration. For Modbus TCP on the virtual PLC, the minimum value was 1 ms, the maximum value was 12 ms, the arithmetic mean was 1,91 ms, the variance was 0,39 ms, and the measurement uncertainty was 0,007206 ms. For S7 on the real PLC, the minimum value was 50 ms, the maximum value was 143 ms, the arithmetic mean was 98,57 ms, the variance was 9,42 ms, and the measurement uncertainty was 0,035548 ms.

#### 4. WORK PROGRESS

The entire program in Unity is currently being redesigned so that it can share data in a suitable format with the Rest API server and work with Siemens PLCs via Modbus TCP. A PLC program with HMI visualization was created for the simulation. The Rest API server was created to have the same data structure as the simulation. The main goal of the Rest API server is that after receiving new data, it stores it in a global variable until it is overwritten again by simulation. The C # application is currently being worked on, but it is already communicating with the Rest API server and collecting data. At the same time, PLC communication with simulation in Unity was measured.



**Figure 2:** Simulation in Unity



**Figure 3:** HMI Main screen

## 5. CONCLUSION

A conclusion was drawn from the measurement results of the communication protocols. For Modbus TCP, the minimum value was 1 ms, the maximum value was 12 ms, and the arithmetic mean was 1,91 ms. For S7, the minimum value was 50 ms, the maximum value was 143 ms, the arithmetic mean was 98,57 ms, the variance was 9,42 ms, and the measurement uncertainty was 0,035548 ms. It follows that Modbus TCP with a virtual PLC is clearly the fastest. It is clear that a real PLC that uses its own processor can never compare to the response speed of a virtual PLC on the same device where the simulation is taking place. It is necessary to realize that communication is not real-time, so in the simulation there is always a small delay.

I have currently completed one of the last important tasks, which is the communication between the simulation in Unity and the test application. I have successfully tested this and will continue to make minor improvements. Even though this is not the main goal of the work, I will try to create at least a basic optimization algorithm.

## ACKNOWLEDGMENT

This work was supported by Brno University of Technology and was carried out with the support of the project FEKT-S-20-6205 Research in automation, cybernetics and artificial intelligence for Industry 4.0.

## REFERENCES

- [1] R. Beloiu, "Virtual Commissioning of Wheel Robot Processing," *2021 12th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, 2021, pp. 1-4, doi: 10.1109/ATEE52255.2021.9425077.
- [2] H. Carlsson, B. Svensson, F. Danielsson and B. Lennartson, "Methods for Reliable Simulation-Based PLC Code Verification," in *IEEE Transactions on Industrial Informatics*, vol. 8, no. 2, pp. 267-278, May 2012, doi: 10.1109/TII.2011.2182653.

# Asset Administration Shell for the operator

V. Houdek<sup>1</sup>, O. Baštán<sup>1</sup>

<sup>1</sup>Faculty of electrical Engineering and Communication, Brno University of Technology, Brno, Czech Republic

E-mail: [xhoude07@vutbr.cz](mailto:xhoude07@vutbr.cz), [Ondrej.Bastan@vut.cz](mailto:Ondrej.Bastan@vut.cz)

**Abstract**—The aim of this article is to acquaint the reader with the implemented project, which dealt with the design of its own Asset Administration Shell (AAS) for the operator. In this article, we focus on the definition of important terms related to AAS in the environment of Industry 4.0 (I4.0). Subsequently, we design our own metamodel, communication and state machine of our AAS in accordance with standards and I4.0 compatibility.

**Keywords**—I4.0, AAS, OPC UA, TCP/IP, IoS, DT, Android

## 1. INTRODUCTION

With the implementation of I4.0 and its elements such as the digital twin (DT), Internet Services (IoS), or artificial intelligence (AI) algorithms, there is an increasing effort to create a fully self-contained automated production decentralized system.[1]

One of the newer concepts of I4.0 is the Asset Administration Shell (AAS) technology, which brings us closer to an effective decentralized system. The development of this I4.0 element has not yet been completed and therefore there is no standard directly specifying and describing AAS. There are only experimental proposals and attempts to unify approaches to the design of the AAS concept. For these reasons, we decided to try to design our own AAS concept.[1] [3]

During the implementation of the AAS, we were limited by the feasibility of the entire project at other devices within the already completed school testbed. Due to the need for compatibility of the entire system with equipment such as PLC, and also due to time constraints, we decided to implement in our AAS design only the basic elements of the extensive AAS issues.

Our implementation of AAS for the operator aims to effectively integrate the human factor into the decentralized I4.0 system. We chose a mobile device with the android operating system as the interface between the person and his digital representation. We will deal with the topic of implementation platforms in the chapter 2.

### 1.1. AAS DEFINITION

In essence, AAS virtually represents our chosen object and at the same time mediates all activities of the object in connection with the interaction with its environment, mainly in a decentralized industrial system.[3][5]

AAS can be divided into active and passive parts. The passive part contains all data, whether publicly accessible or internal AAS data. Each piece of data can be divided into type and instance. Type represents a data template based on standardized metamodels, and an instance is an already created data object with specific data assigned.[3][4][5]

The active part includes communication within the IoS network, implementation of internal AAS methods and communication with the asset itself. When communicating within an IoS network, we ideally talk about communication between two AASs.[5]

### 1.2. AAS IN I4.0

AAS can be viewed in I4.0 from several perspectives, the inclusion of AAS in the RAM I4.0 model is probably a view that allows visualization of AAS in a broader context and context to other elements in the implementation of I4.0. When implementing AAS to I4.0, different views on this issue can be found from the professional community. Examples include a view of the deployment of AAS in I4.0 from Standardization council Industrie 4.0 (SCI 4.0) and a view of SAP.[1][6]



## 2. PLATFORMS FOR REALIZATION AN AAS

AAS can be implemented on almost any hardware device (HW) and in any programming language. It only has to be placed on the standardization and interoperability of the resulting device.[5]

We chose to develop our own AAS from the ground up with our own SW development. In this case, it seems like an ideal choice of programming language from a menu of higher programming languages such as Python, C ++, go or Java.

In our case, it will be a combination of C ++, in which the main communication and the AAS state machine will be implemented, and we will combine the C++ code parts with the Java programming language in the android studio environment for human interface implementation. The resulting combination of partial parts of the code gives a complete AAS.

As for the HW implementation, in our implementation case we are limited to devices with the android operating system (OS). So smartphones and tablets. These devices appear to be the most practical for implementing AAS for the human operator, for several reasons such as: device availability, compactness, reliability, robustness, open source OS. In other implementation cases, we are almost not limited here and everything depends on compatibility with other devices in IoT.

## 3. METAMODEL OF AAS

Our proposed AAS metamodel see: Figure 1 is based on the structure of AAS according to ZVEI. We designed our own metamodel AS due to the practical feasibility of the resulting AAS not only at our device, but also at other devices at BUT UAMT. When creating it, we tried to stick to the already established properties. [3] The structure of the metamodel consists of the following key parts:

- Root node - this is the AAS itself determined according to DIN 91345.
- Header - contains the minimum amount of information publicly available, in order to identify the AAS.
- Body - This part is a passive container according to the IEC 62832 standard. Additional elements are added to the definition in the AAS, namely services with the possibility of invoking services.[3]

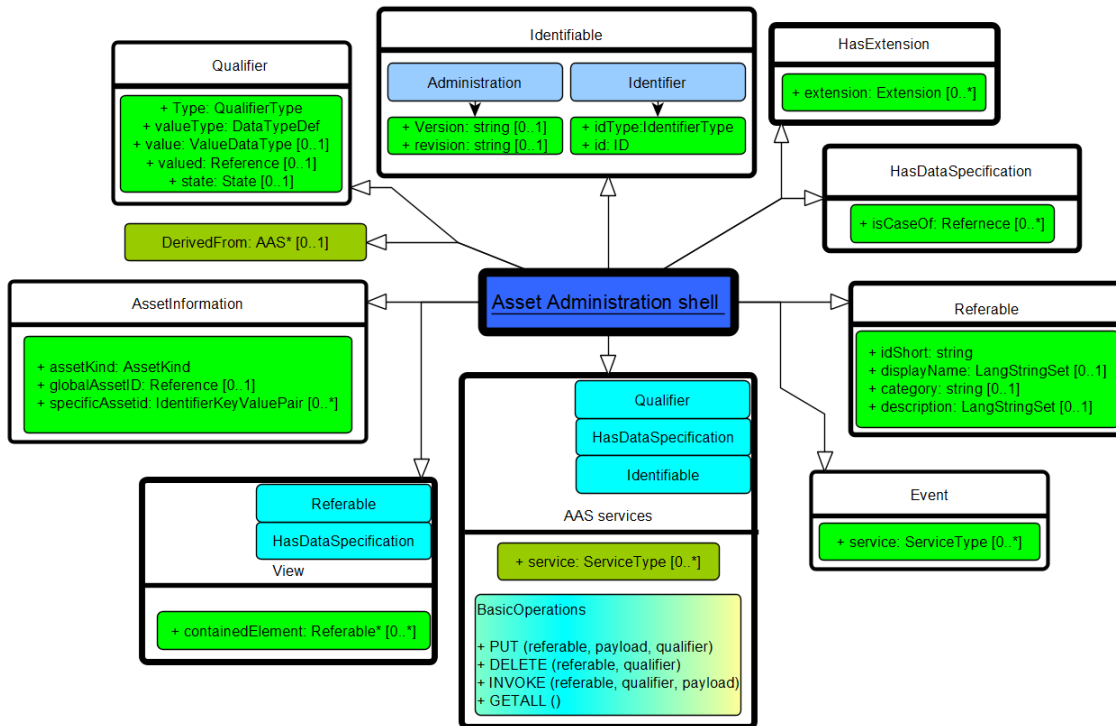


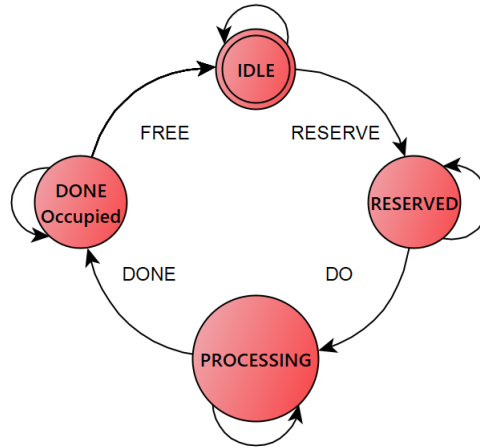
Figure 1: AAS metamodel design diagram



#### 4. STATE MACHINE OF AAS

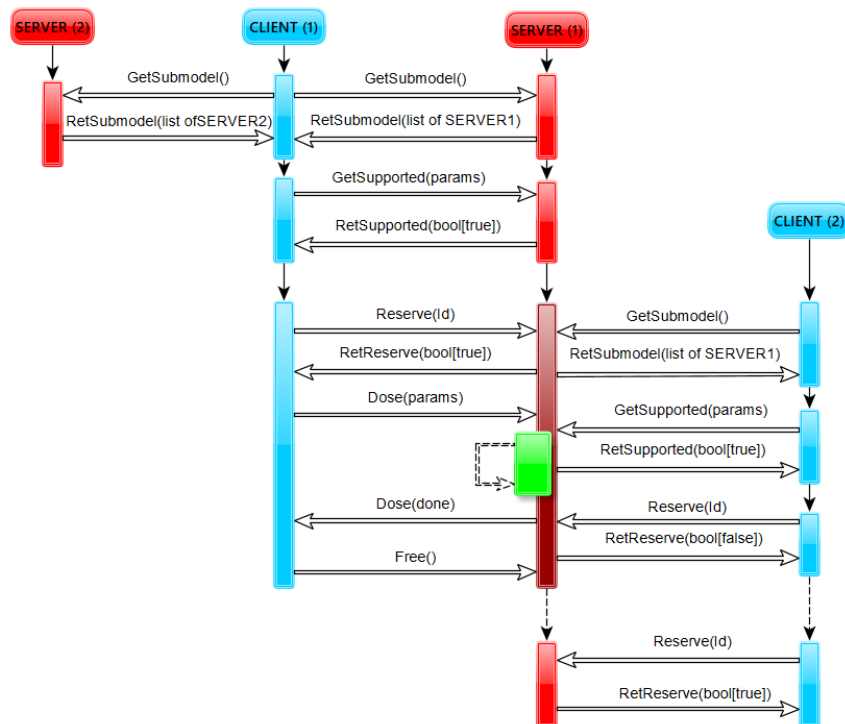
As such, the State Machine of AAS has not yet been specified or standardized in any detail by any institution. In the German ZVEI publication Details of the Asset Administration Shell. Part 2, general specifications of methods and functions can be found on an abstract basis. This document also discusses the breakdown of the various interfaces in AAS and the different views of the various APIs that AAS encounters in its life cycle.[4]

For our AAS implementation, we designed a simple and functional state machine that can be easily implemented on other platforms and in other programming languages than our chosen C ++.



**Figure 2:** AAS state machine for specific implementation

State machine designed by us, see: Figure 2, consists of 4 states: IDLE, RESERVED, PROCESSING, DONE Occupied. The IDLE state is the initialization and base state. The RESERVED state is a state when it is no longer possible to book a service from another requester. After a direct request for action, the PROCESSING state occurs, when AAS performs the requested action. After its end, it moves to the DONE Occupied state, waiting for the reservation to be released from the original service applicant.



**Figure 3:** Communication diagram between AASs, between product (Client) and production cell (Server)

## 5. COMUNICATION BETWEEN AASS

We chose OPC UA with TCP / IP protocol as the most suitable communication architecture for AAS. The form of this communication is visualized by a flow diagram, see: figure 3. When implementing a communication architecture, there is an experimentally frequently used MQTT architecture. However, this architecture includes a central element of the broker. This element is often redundant in these implementations and therefore centralization is lost.[2]

In a smaller number of connected devices, the MQTT achieves a lower latency of messages, but there is a problem (after exceeding a certain number of devices in the network) in congestion of such a network, until the critical moment of collapse and blockage of the communication network AASSs.

The diagram (see: Figure 3) shows an example of visualization of communication AASSs servers and clients via the already mentioned OPC UA. The methods that are attached to our state machine are shown and implemented here (see: Figure 2). [3][4][5]

## 6. CONCLUSION

When designing our own AAS model, we had to consider compatibility with other devices in our network and the implementation of world properties already created for AAS. To meet both conditions, we were forced to take only the basics of AAS world properties.

The state machine designed by us was designed with the aim of functionality and simple implementation. However, it has minor shortcomings (deadlock situation), which may be addressed in future versions of our AAS. Despite the room for improvement, this AAS proposal can be considered successful, as we were able to design and partially implement the basic functionality of the extensive and rich issues of AAS technology according to existing standards in I4.0.

## ACKNOWLEDGMENT

The completion of this paper was made possible by the grants:

No. FEKT-S-20-6205 - "Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0";

## REFERENCES

- [1] Marcon, P.; Arm, J.; Benesl, T.; Zezulka, F.; Diedrich, C.; Schröder, T.; Belyaev, A.; Dohnal, P.; Kriz, T.; Bradac, Z. New Approaches to Implementing the SmartJacket into Industry 4.0 ‡. *Sensors* 2019, 19, 1592. <https://doi.org/10.3390/s19071592>
- [2] Arm, J.; Benesl, T.; Marcon, P.; Bradac, Z.; Schröder, T.; Belyaev, A.; Werner, T.; Braun, V.; Kamensky, P.; Zezulka, F.; Diedrich, C.; Dohnal, P. Automated Design and Integration of Asset Administration Shells in Components of Industry 4.0. *Sensors* 2021, 21, 2004. <https://doi.org/10.3390/s21062004>
- [3] Barnstedt, E and Bedenbender, H and Billmann, M and Boss, B and Clauer, E and Fritsche, M and Garrels, K and Hankel, M and Hillermeier, O and Hoffmeister, M and others. Details of the Asset Administration Shell Part 1. Technical report, Federal Ministry for Economic Affairs, 2020.
- [4] Bader, S., Berres, B., Boss, B.& Ziesche, C. Details of the Asset Administration Shell. Part 2 -Interoperability at Runtime - Exchanging Information via Application Programming Interfaces (Version 1.0RC01). (2020,11).
- [5] Pribiš, R. Beňo, L. Drahoš, P. Asset Administration Shell Design Methodology Using Embedded OPC Unified Architecture Server. *Electronics* 2021, 10, 2520. Dostupné z URL:<<https://doi.org/10.3390/electronics10202520>>.
- [6] Industrie 4.0 component and the concept of the Asset Administration Shell. Standartization council Industrie 4.0 [online]. Německo: -, 2017 [cit. 2021-11-20]. Dostupné z URL:<<https://www.sci40.com/english/asset-administration-shell/>>.

# Demonstrative Manufacturing System Controlled by MES Utilizing AAS

P. Černocký<sup>1</sup>, and J. Arm<sup>2</sup>

<sup>1</sup>Bachelor Degree Programme (3.), FEEC, Brno University of Technology, Czech Republic

<sup>2</sup>Department of Control and Instrumentation, FEEC, Brno University of Technology, Czech Republic

E-mail: [xcerno29@vut.cz](mailto:xcerno29@vut.cz), [arm@vut.cz](mailto:arm@vut.cz)

**Abstract**—This work examines today's modern possibilities for production management. More specifically, we focus on MES (Manufacturing Execution Systems) and its integration within the concept of industry 4.0 using AAS (Asset Administration Shell). It also describes a specific integration for a simple virtual line designed in ABB RobotStudio, which is first controlled using MES and then the production is encapsulated using AAS. The AAS is then supposed to interact with ERP (vertical integration) and also with suppliers and other manufacturing units (horizontal integration).

**Keywords**—MES, AAS, IMES, RobotStudio, virtual factory, OPC UA

## 1. INTRODUCTION

This article deals with the possibilities of using advanced production management using MES. For demonstration purposes, we will manage a virtual line created with the help of ABB RobotStudio. As MES, we used one of the open source applications available on the Github server.

All communication takes place using the OPC UA protocol, both with the database and the virtual line and in the next phase with the AAS and the virtual line. This communication is mediated through the NodeRed tool, thanks to which we have relatively easy access to the Firebase realtime database.

A similar topic was dealt with by colleagues in the article [1] in their case, however, it was the use of AAS for MES and its superior system - ie ERP (Enterprise resource planning). Our work is more focused on communication of MES with a lower level - ie with line or PLCs. In the article, however, they used MES from a different creator than us.

## 2. AAS – ASSET ADMINISTRATION SHELL

The Industry 4.0 concept uses their AAS - Asset Administration Shell digital envelope to standardize equipment descriptions. The purpose of these envelopes is to ensure the exchange of information between the facilities, between them and the production coordination system and the engineering tools. [2]

The figure 1 shows the description and connection between the physical device and the AAS. The device envelope (AAS) consists of two parts. Header, which lists unique device identifiers. A body, in which other information about the device, its properties and other important information such as the production process is given. [3]

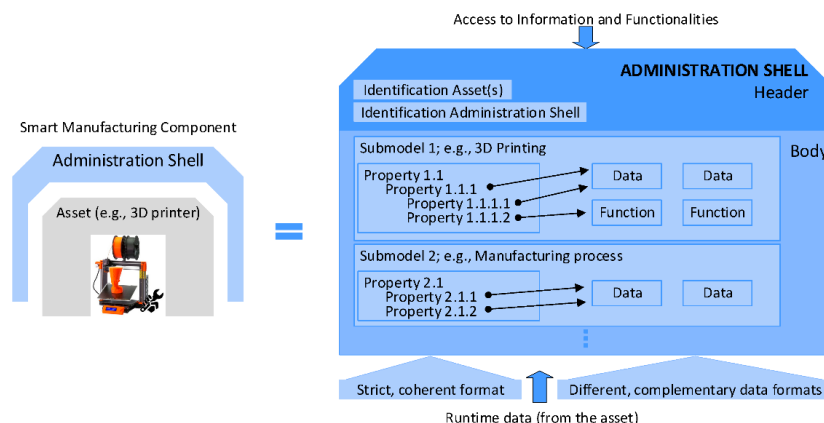
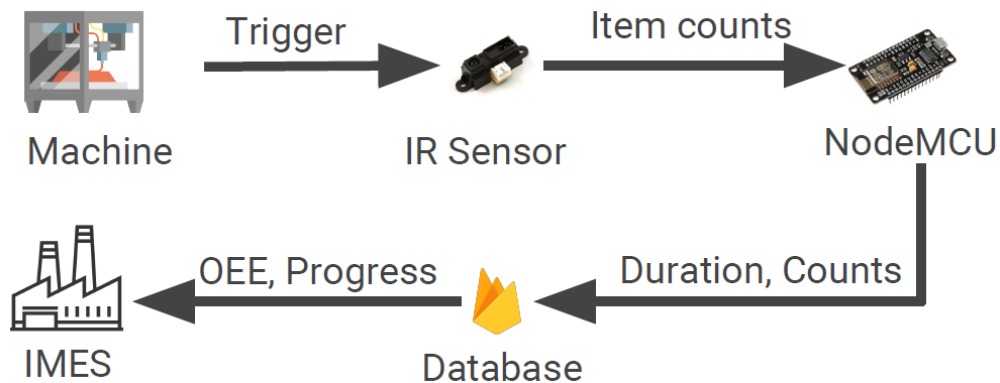


Figure 1: Structure of AAS [3]

### 3. IMES

We used the available open source IMES application to manage our virtual line. This application is relatively simple and should be fully suitable for our demonstration purposes. The advantage of this application is that it already has a module ready for possible sensors that can record the progress of order processing in the company.



**Figure 2:** Structure of AAS [4]

### 4. COMMUNICATION BETWEEN MES AND FACTORY

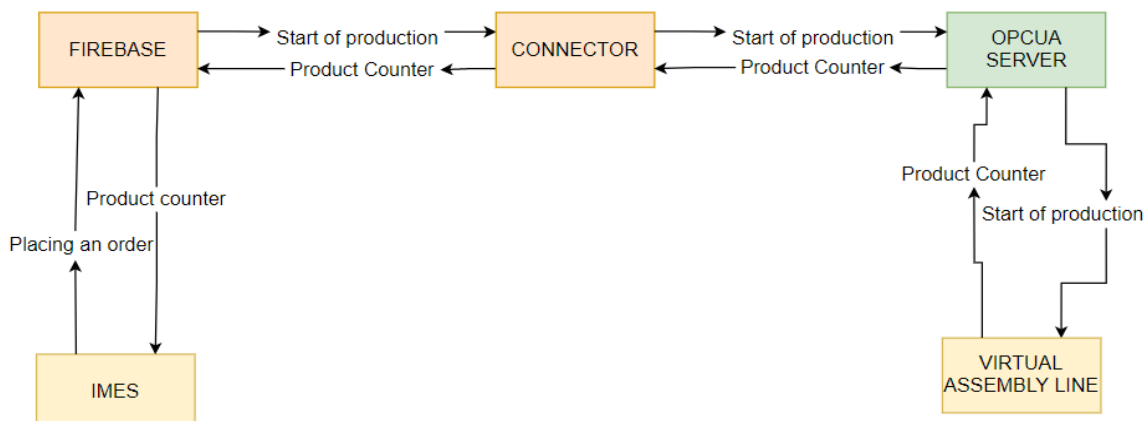
We will use the protocol for industrial communication for communication between individual components. OPC communication is generally used for the exchange of data between different industrial systems. In automation, it is a universal communication platform that can connect to the data of hundreds of different types of devices from different manufacturers and convert this data into a single OPC communication, understandable to many superior applications such as ERP, SCADA, or in our case MES.

Communication between the client and the server takes place exclusively through calls and processing of services (Services), which deal with the control of individual parts of the OPC UA server functions. Both queries and answers have their common headers, where the client has, for example, the ability to set the required information to be returned by the server for all queries. [5]

### 5. MANAGEMENT STRUCTURE WITHOUT AAS

In this case, it is practically a classic pyramid control, where data is exchanged between the virtual line and the database of our application. RobotStudio creates the OPC Server and sends the simulation data to it. We read the production progress data from the server using the NodeRed tool. We then send the data to our Firebase real-time database. In the same way, communication takes place in the opposite direction, where we read information from the database and send it to the server.

This is illustrated in the block diagram in Figure 3



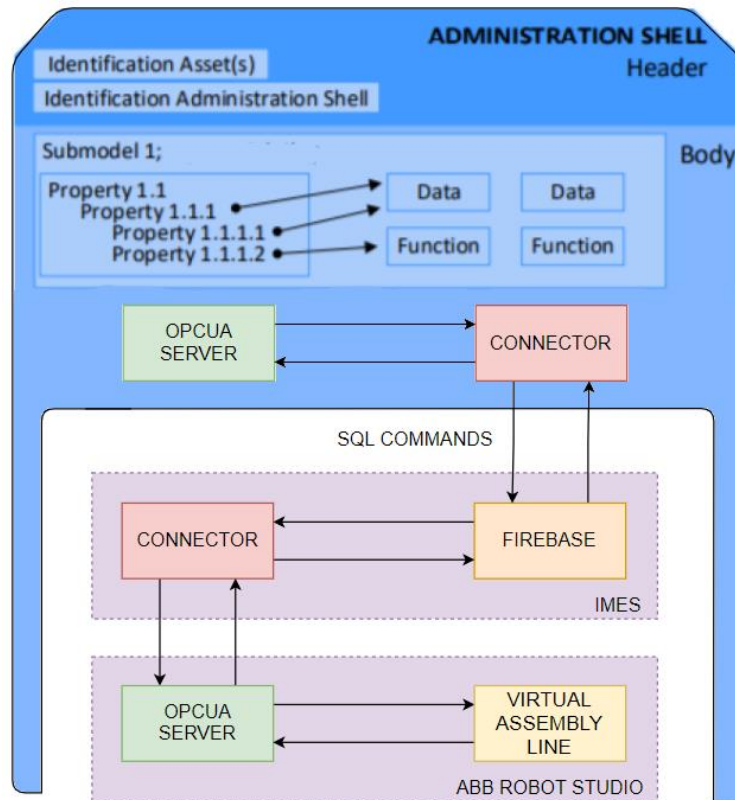
**Figure 3:** Block diagram classic management

## 6. MANAGEMENT STRUCTURE WITH AAS

When managing with the help of AAS, we will create an envelope, which we will cover our entire application and we will communicate only with the header of our asset. Eventually, the entire AAS will have modules in place for both communication with the enterprise management system (ERP) and communication with the lower tier. The ISA-95 standard tells us what information should be passed on. In our case, however, this would mean that we would have to modify the database of our MES application. Therefore, we will prepare only the given submodels in our AAS and I will use only those that will be beneficial for our application.

Communication between the asset and AAS blocks takes place on the basis of SQL statements. Based on them, the data will be written directly to the Firebase database of our MES application. The configuration then takes place on the basis of our selected submodels "communication settings" and "definition of variables and methods". However, these submodels can be extended by others.

The block diagram here shows the possible structure of the project



**Figure 4:** Block diagram management with AAS

## 7. DEMONSTRATION OF MES AND VIRTUAL LINE DEPLOYMENT

Figure 5 shows one of the possible deployment methods. This is line control without the use of AAS.

Here we see a virtual line created in RobotStudio and part of the IMES application. More specifically, the production monitoring section, to which we receive data from the simulation.



**Figure 5:** example of IMES deployment

## 8. CONCLUSION

This work deals with the possibilities of production management on a demonstration virtual line. Both the classical methods of control using the MES itself and the possibilities of control using the AAS are discussed here. This means for us that we will pack our entire application in asset and create submodels according to the ISA-95 standard. This standard tells us which data and information are to be sent one level up (to the ERP) and also one level down, e.g. to the control PLCs.

In our work, we first had to run the IMES application and pair it with the Firebase Real time database. After that, we used the NodeRed tool to connect our database with the OPC UA server, which generates simulations in ABB RobotStudio. In this simulation, pulses are generated for the simulated sensor, we then calculate these pulses and thanks to that we can monitor the production process after that. We try to ensure all data transmission through communication via the OPC UA protocol, which is one of today's standard industry protocols.

## 9. ACKNOWLEDGMENT

This work was supported by Brno University of Technology and was carried out with the support of the project FEKT-S-20-6205 Research in automation, cybernetics, and artificial intelligence for Industry 4.0.

## REFERENCES

- [1] X. Ye, M. Yu, W. S. Song and S. H. Hong, "An Asset Administration Shell Method for Data Exchange Between Manufacturing Software Applications," in *IEEE Access*, vol. 9, pp. 144171-144178, 2021, doi: 10.1109/ACCESS.2021.3122175.
- [2] Industrie 4.0 Asset Administration Shell: koncepce a využití v praxi. AUTOMA: časopis pro automatizaci [online]. [2019], 2019(8-9) [ref. 2021-11-20]. Available at: [https://automa.cz/cz/casopis-clanky/industrie-4-0-asset-administration-shell-koncepce-a-vyuziti-v-praxi-2019\\_08\\_0\\_12444/](https://automa.cz/cz/casopis-clanky/industrie-4-0-asset-administration-shell-koncepce-a-vyuziti-v-praxi-2019_08_0_12444/)
- [3] Arm, J.; Benesl, T.; Marcon, P.; Bradac, Z.; Schröder, T.; Belyaev, A.; Werner, T.; Braun, V.; Kamensky, P.; Zezulka, F.; Diedrich, C.; Dohnal, P. Automated Design and Integration of Asset Administration Shells in Components of Industry 4.0. *Sensors* [online] [2004]. [ref. 2021-12-10]. Available at: <https://doi.org/10.3390/s21062004>
- [4] Product poster of IMES application. [online][2017] [ref.2021-12-22]. Available at: <https://github.com/jukbot/smart-industry/blob/master/About%20Project.pdf>
- [5] OPC UA VÁM ZJEDNODUŠŠÍ ŽIVOT, ALE... FOXON [online]. Liberec: FOXON, 2021, 4. září 2019 [ref. 2021-5-19]. Available at: <https://foxon.cz/blog/ostatni/454-opc-uavam-zjednodusi-zivot-ale>



# Ice cube feeder

**M. Mařas, T. Sýkora**

Brno University of Technology

Faculty of Electrical Engineering and Communication, Department of Control and Instrumentation,  
Czech Republic

E-mail: [xmaras01@vutbr.cz](mailto:xmaras01@vutbr.cz), [xsykor23@vutbr.cz](mailto:xsykor23@vutbr.cz)

**Abstract**—This article deals with the description of the proposal and implementation of autonomous cell called Ice cube feeder. The following lines contain the description of autonomous cell in terms of design, electrical and programming. The concept represents the ideas of Industry 4.0 and the entire production process follows the ISA-S88 standard, which describes the batch control process.

**Keywords**— Industry 4.0, Testbed, Self-acting bartender, autonomous cell, Siemens NX, PLC, HMI. ice

## 1. INTRODUCTION

This article describes one of the autonomous cells that make up the testbed. This testbed was built as a demonstrator of the principles of Industry 4.0.

The first part of this article describes the testbed with the working title Self-acting bartender. The second part focus on the mechanical and electrical parts of an autonomous cell called the Ice cube feeder.

The main goal of this article is the description of an autonomous cell that can feed ice cubes from different manufacturers. That means it can dose different cube sizes.

## 2. SELF-ACTING BARTENDER

The testbed named self-acting barman represents the process of making an alcoholic drink. The process of drink production is realized using several autonomous cells that interact with each other. The main emphasis was placed on the possibility of exchanging individual production cells on the testbed. This method demonstrates the mutual independence of production units.

The product, in this case, a glass, has an NFC chip in which all important information about the required drink is stored. This feature allows the glass to order operations on cells that can do it. The most important thing stored in the NFC chip is the drink preparation recipe itself, which determines the production process. Each cell can read the information about the required operation using a reader located inside the stand on which the glass is placed. [1]

For example, the SCARA robot allows the glass to be moved between individual production units. Some cells may offer more features that may be beneficial in the event of a failure of another manufacturing cell. With this system, the individual cells work together to achieve the desired result in the form of a mixed drink. [2]

The testbed also demonstrates the batch process according to the ISA-S88 Batch control standard. [3]

### 2.1. Testbed and Industry 4.0

The testbed corresponds to certain ideas of Industry 4.0, such as virtualization, decentralization, and modularity. In this case, virtualization means 3D models in Siemens NX thanks to which it can then carry out research or testing without the product having to be made or during its production. It is also possible to create a cybernetic twin. The cyber twin is a mechatronic model that implements the functionality

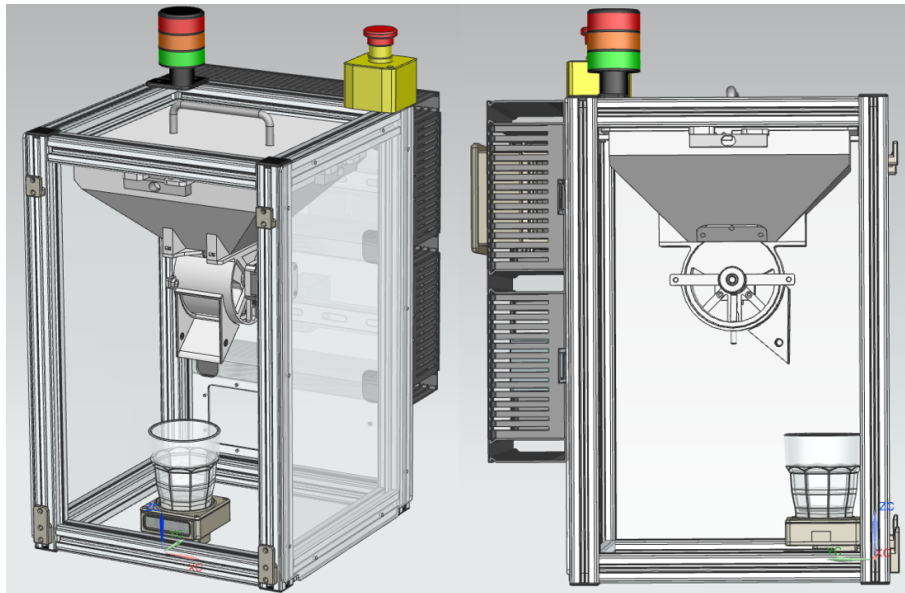
of cells in virtual form. This allows testing of critical system conditions or conditions that would cause damage to the equipment.

Decentralization is a way of passing instructions between a product and cells. In this case, the recipe data is stored "right in the glass" and this allows the glass to specify commands for individual cells. Modularity and reconfigurability in the testbed are realized except for exceptions cells can be identified as structures composed of aluminum profiles of the same dimensions. Thanks to this, we can change the individual positions of the cells and then them configure as needed.[4]

### 3. ICE CUBE FEEDER

The previous concept of this cell was to crush ice cubes into crushed ice. However, after testing, it has been found that great demands are placed on structural strength because the parts of the crusher have irreversibly deformed.

The new ice cube feeder concept is to feed whole ice cubes. This process is mediated by a rotating cylinder with inner space for an ice cube. The radial movement will allow the ice cube to be transported into the cylinder and further rotation will ensure that the cube is transported into the glass. Figure 1 shows the construction of an ice dispenser. The picture shows a hopper with an attached mechanism for dosing cubes.



**Figure 1:** 3D model of new Ice cube feeder

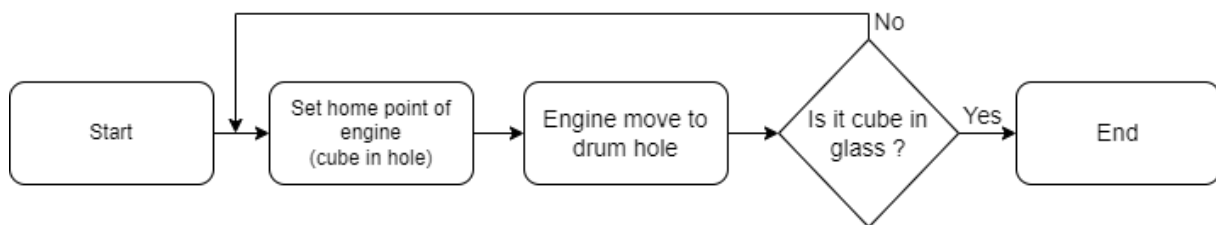
#### 3.1. Project benefits

- modularity of the cylinder - possibility of exchanging cylinders with different hole sizes for the cube (allows feeding of cubes from different manufacturers)
- determining the weight of the cubes = determining the number of cubes in the hopper
- drainage of melted water from the cubes
- verification of the transport of the cubes into the glass using a strain gauge
- verification of lid closure using end switches
- industrial design for longer life and reliability

### 3.2. Feeding process

This section describes the sequence of the feeding process. This sequence is also shown in the flowchart on Figure 2.

1. Locking the motor in the starting position, the hole in the cylinder is located above the hopper opening.
2. Transport of the ice cube into the hole in the cylinder by gravity.
3. Turning the cylinder to a position where the hole in the cylinder is parallel to the opening of the drum => the cube falls into the glass.
4. Verification of cube transportation into the glass using a strain gauge under the glass.  
Verification is also provided by measuring the weight loss of the hopper (see chapter weight measurement).



**Figure 2:** Flow chart of the feeding process

### 3.3. Weight measurement

Four strain gauges provide measuring of weight for identification of the number of cubes in the hopper. Aluminum beams are attached to the cell frame on the four measuring points. Wheatson's bridge is glued on each of them. They are connected in parallel, which means that the subsequent change in resistance on the four Wheatson's bridges is averaged.

The signal from strain gauges needs to be amplified because at the output of the Wheatson's bridge we would notice a very small change in voltage, which cannot be evaluated by a PLC. It is also necessary to adapt the signal to the PLC input so that the evaluation is as accurate as possible. Therefore, we decided to use an electrical device called SKUseCon, which can record the output signal from the bridge to amplify it and then use the software solution to convert it to a rectangular signal in which the frequency changes depending on the load of the bridge.

After determining the weight, the number of cubes is found to divide the total weight by the weight of one cube.

An important aspect of accurate weighing is calibration. Setting the zero voltage value when the hopper is not loaded using a trimmer, which sets the gain and reading of the analog value. This value is then written to the software. The same process is applied in the case of maximum load.

### 3.4. Safety features

- emergency button for immediate stop of the production process
- perimeter plexiglass to prevent human contact with the active cell
- signaling beacon for light signaling the current state of the cell

### 3.5. Heat insulation

To eliminate ice cube melting as long as possible, thermal insulation is required. Therefore, the entire hopper is covered with K-flex insulation material. K-flex is thermal insulation and vapor-tight material

based on rubber with a thermal coefficient of thermal conductivity of 0,031-0,037 W / (mK) and a coefficient of resistance of water vapor diffusion  $\geq 7000$ . It contains a thin aluminum layer for better heat shielding.

#### 4. CONCLUSION

This article aimed to create a mechanical and electrical design, technical documentation, and physical implementation of an autonomous cell Ice cube feeder, which will work in the context of Industry 4.0 ideas.

First, it worked on an iterative process of mechanical cell design in Siemens NX. After the implementation of the mechanical parts on the 3D printer, the individual parts were assembled. The individual sensors and actuators that were installed in the assembly were determined. After the assembly of the whole cell, the software was written. The device was revived and tested.

All goals were met. In the future, the project will deal with the synchronization of all cells in the Self-acting bartender complex.

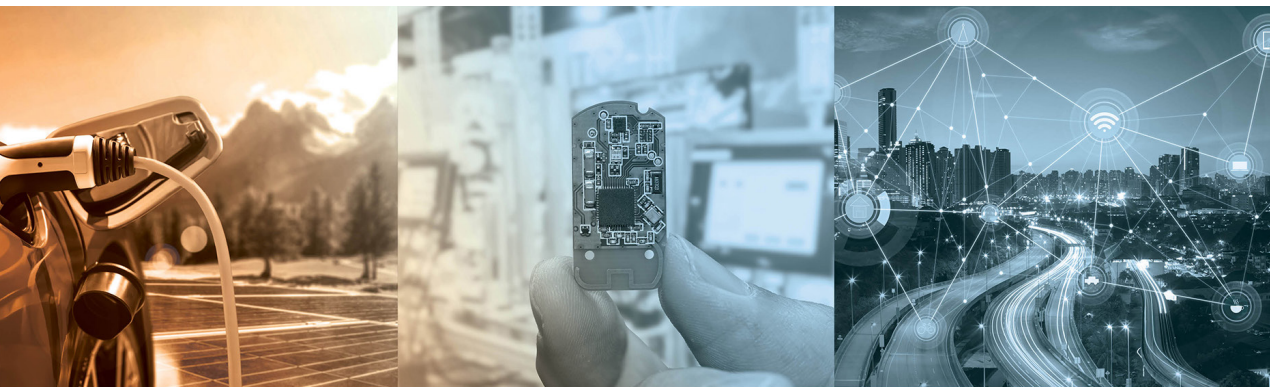
#### ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-20-6205—“Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0” financially supported by the Internal science fund of Brno University of Technology.

#### REFERENCES

- [1] DRHLÍK, Jakub. Bezdrátové technologie: Co je NFC a jak ho využít?. Svět androida [online]. 4.7.2017 [cit. 2022-03-10]. Available from: <https://www.svetandroida.cz/bezdratove-technologie-nfc/>
- [2] KACZMARCZYK, Václav, Ondřej BAŠTÁN, Zdeněk BRADÁČ and Jakub ARM. An Industry 4.0 Testbed (Self-Acting Barman): Principles and Design. IFAC-PapersOnLine [online]. 2018, 51(6), 263-270 [cit. 2022-03-10]. ISSN 24058963. Available from: doi:10.1016/j.ifacol.2018.07.164
- [3] POŽIVIL, Jaroslav. Co lze nalézt v normě ANSI/ISA-88.01. AUTOMA [online]. 10/2001 [cit. 2022-03-9]. Available from: [https://automa.cz/cz/casopis-clanky/co-lze-nalezat-v-norme-ansi/isa-88-01-2001\\_10\\_33691\\_1736//](https://automa.cz/cz/casopis-clanky/co-lze-nalezat-v-norme-ansi/isa-88-01-2001_10_33691_1736//)
- [4] MINISTERSTVO PRŮMYSLU A OBCHODU ČR. Iniciativa průmyslu 4.0 [online]. 10/2016 [cit. 2022-03-11]. Available from: <https://www.mpo.cz/assets/dokumenty/53723/64358/658713/priloha001.pdf>

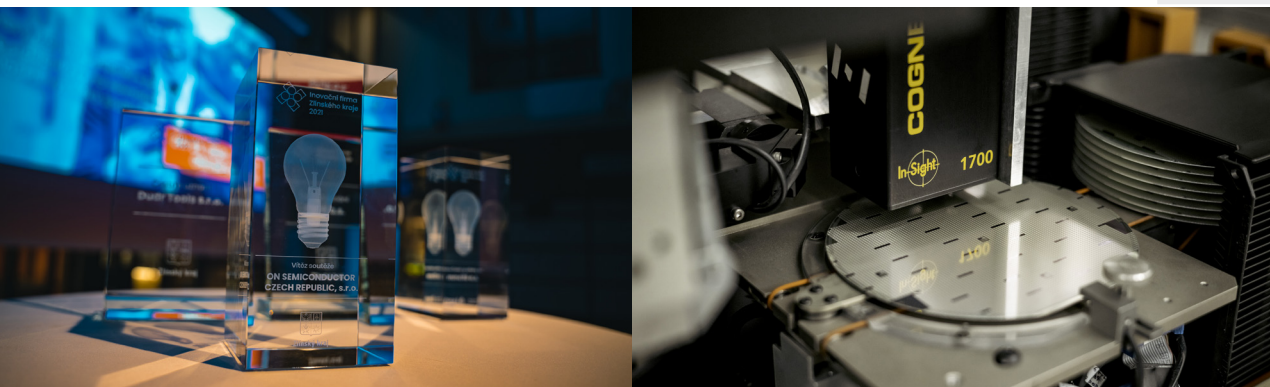
# Budoucnost se děje nyní



Současná expanze výroby nám umožňuje vyrábět špičkové produkty na unikátních technologických zařízeních.

- V posledních čtyřech letech jsme investovali čtyři miliardy do posílení našich výrobních kapacit.
- Kromě výroby čipů na křemíku rozvíjíme i výrobu na karbidu křemíku. Ve výrobě čipů na karbidu křemíku se chceme stát světovou jedničkou.
- Čipy na karbidu křemíku jsou slibnou technologií. Ve srovnání s čipy na křemíku dokážou uspořit až 10% energie, která přes ně prochází.

Získali jsme prestižní ocenění Inovační firma Zlínského kraje 2021



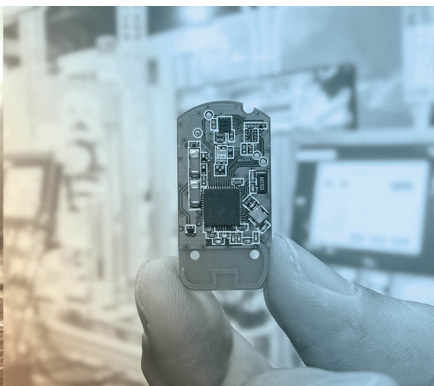
U nás si můžete plnit své sny

Nabízíme:

- Práci na technologiích a produktech, které přispívají k zelenější budoucnosti.
- Šanci být u zrodu převratných novinek v polovodičovém průmyslu.
- Cenné zkušenosti díky nejmodernějším technologiím a systémům.
- Tým odborníků na špičce technologie polovodičů.
- Nejlepší benefity v regionu.
- Zázemí stabilní mezinárodní společnosti.

**onsemi**™





Jsmo úspěšným světovým výrobcem polovodičů s dlouhou tradicí a více než 2000 zaměstnanci v České republice. Naše výrobky mají široké využití v různých oblastech, jako jsou spotřební a průmyslová elektronika, automobilový průmysl, telekomunikační technologie a výpočetní technika. V České republice působíme v Rožnově pod Radhoštěm a v Brně v oblastech:

- návrhu integrovaných obvodů
- výzkumu a vývoje nových technologií
- technologií výroby křemíkových desek a polovodičových čipů

## NABÍZÍME

Pro studenty:

- placené stáže a brigády
- stipendia
- specifikace témat a vedení bakalářských/diplomových prací
- exkurze
- prohlídky pro veřejnost
- pro absolventy nabízíme pracovní uplatnění na zajímavých pracovních pozicích

## HLEDÁME

Preferované obory:

- mikroelektronika
- radioelektronika
- výkonová elektrotechnika
- elektrotechnologie
- chemické, fyzikální a materiálové inženýrství

- anorganická a analytická chemie
- informační technologie
- průmyslové inženýrství a automatizace
- fyzika pevných látek

## PRŮBĚH VÝBĚROVÉHO ŘÍZENÍ

Klademe důraz především na odborné a jazykové znalosti, které uchazeči získali během studií a předchozí praxe. Podstatným kritériem výběru je pro nás také osobnostní hledisko. Během pohovoru se setkáte se zástupcem personálního oddělení a vedoucím, který danou pracovní pozici ve svém týmu obsazuje. Společně se snažíme vytvořit přátelskou a neformální atmosféru.

*„Pro mnoho lidí je práce povinnost, pro mě však práce pro onsemi znamená radost. Považuji za štěstí, že jsem se mohla stát členem skupiny neskutečně vzdělaných lidí a špičekách ve svém oboru. Ve vývoji polovodičových součástek je velká budoucnost. Jsem ráda, že jako analogový designér mám tu možnost kombinovat seberealizaci v tomto technickém odvětví a zároveň spolupráci se skvělými lidmi.“*

Zuzana Kozlová  
Recruiter  
zuzana.kozlova@onsemi.com  
+420 605 842 281

SCG Czech Design Center, s.r.o.  
1. máje 2594  
756 61 Rožnov pod Radhoštěm  
Česká republika



Design Layout Engineer

[www.ikariera.cz/cz/companies/detail/370](http://www.ikariera.cz/cz/companies/detail/370)  
<https://kariera-onsemi.cz/>

STROJNÍ  
ELEKTRO  
APLIKOVANÉ VĚDY  
INFORMATIKA  
EKONOMIE  
CHEMIE  
DESIGN  
JAZYKY  
HUMANITNÍ  
OSTATNÍ

PLNÝ ÚVAZEK

ČÁSTEČNÝ ÚVAZEK

BRIGÁDA

BAKALÁŘKA

DIPLOMKA

STIPENDIUM

PRAXE/STÁŽ

TRAINEE PROGRAM



PROFIL

GALERIE

VIDEO

AKTUÁLNÍ  
PŘÍLEŽITOSTI  
SPOLUPRÁCE

ONLINE  
PRACOVNÍ  
NABÍDKY

ZAMĚSTNANCI

UDÁLOSTI

ČLÁNKY

FAKTA

KONTAKTY



# Automatic check for the certification label

R. Čambál, T. Sýkora

Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech Republic

E-mail: [220968@vut.cz](mailto:220968@vut.cz), [173753@vut.cz](mailto:173753@vut.cz)

**Abstract**—The subject of this article is to propose quality improvement in the field of automated control of presence/absence of the certification label for car seats intended for Chinese market, with the introduction of quality inspection of the label and inspection of the correct affixation position at a defined seat position. The main benefit of this work is to apply a measurement process using Poka-Yoke control by automating the process at Lear Corporation Seating Slovakia s.r.o..

**Keywords**—quality, automation, control, process, sensor, Poka-Yoke competition

## 1. INTRODUCTION

The development of technology in the field of industrial automation is advancing rapidly and this brings a large number of new opportunities for manufacturing companies. This progression affects our production processes and also increases the demands placed on product quality. Improvements in quality control are applied by the replacement of human factor in a faster and more reliable solution. This is currently being implemented by introducing various sensors into the production process. We know this application significantly improves efficiency and is reducing production errors that could lead to potential losses.

The main subject of this article is the application of automated control of presence/absence of a certification label for car seats intended for the Chinese market in Lear Corporation Seating Slovakia s.r.o.. Presence and quality detection label in the process is insufficient and therefore corrective action is required by implementing automated control.

### 1.1. MOTIVATION

Our project was developed at the Lear Corporation Voderady branch located in Slovakia. The main activity of this company branch is the production of passenger seats for cars. The company is tasked with production of the required number of seats with desired quality in a defined time.

The company's production line is established in the form of an automated conveyor belt with individual manual assembly stations. The automation of the seat production process is rarely used due to the complicated assembly process. The process itself contains a large number of very difficult to replace procedures, for which it is more appropriate to use manual work. The production of the company is therefore very dependent on the operators, which brings us various risks. Therefore with bringing simple automated solutions where it's possible and where it can reduce stress placed on operators, we can benefit in reducing production quality errors.

### 1.2. PROBLEMATIC

The CCC label (China Compulsory Certificate) is a product certification label designed for Chinese market. An example of such a label is shown in Figure 1.

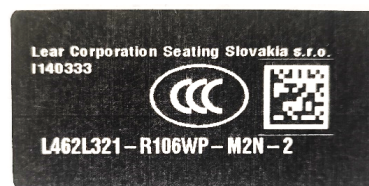


Figure 1: CCC certification label

In its current state, the label is printed on the printer of the assembly station for certified variant of seat according to the customer's order. Operator at the assembly station takes the label and sticks it on a defined place for a specific seat type. The next step is to scan the presence of the label with hand-held scanner which sends information to the Lear production system (LPS), which is used for production control and management [2].

No other control in the following process is implemented and this has caused several complaints:

- The label is missing or remaining - Operator scanned the label but did not stick it (did not follow the standard process) or subsequently pasted it on another order.
- The label has been damaged or illegible - Improper handling or incorrectly set printer (printer did not print the entire contents of the label).
- The label was not affixed to the correct position on the seat - Operator affixed label on the wrong position.

Detection of the presence and quality of the label in the production process is insufficient. Parameters such as material, size, type or location of the label is determined by the customer (carmaker) and cannot be changed. These problems require corrective action. One option is to use Poka-Yoke process automation [3, 4].

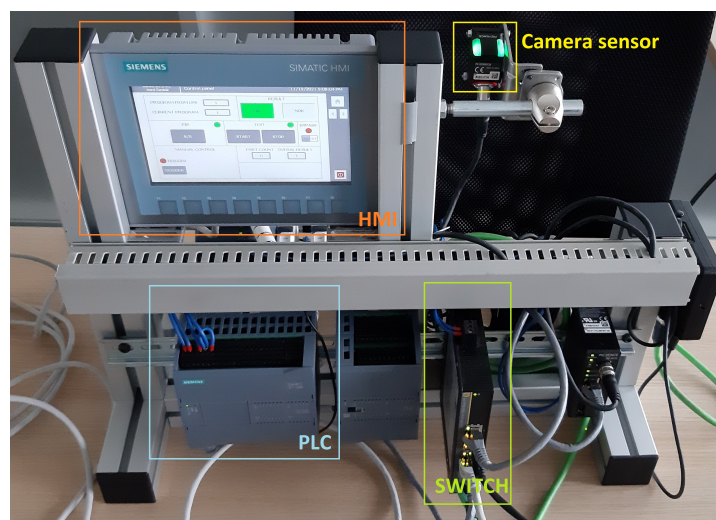
## 2. WORKPLACE FOR TESTING

It is necessary to select appropriate components and test our solution before the implementation on the production line. Designed workplace will provide us with effective testing of the program so that we can avoid potential ones defects and shortcomings which are unacceptable in production. Resources for the implementation of the testing workplace have been provided to us in advance by the company. The components used in the test facility are listed below:

- SIMATIC S7-1200 control system
- SIMATIC HMI KTP700 Basic
- Keyence IV camera sensor

### 2.1. SOFTWARE AND HARDWARE CONFIGURATION

The hardware and software configuration for programming logic controller (PLC) is realized via the Siemens TIA portal environment. It is a tool for efficient control of Siemens components.



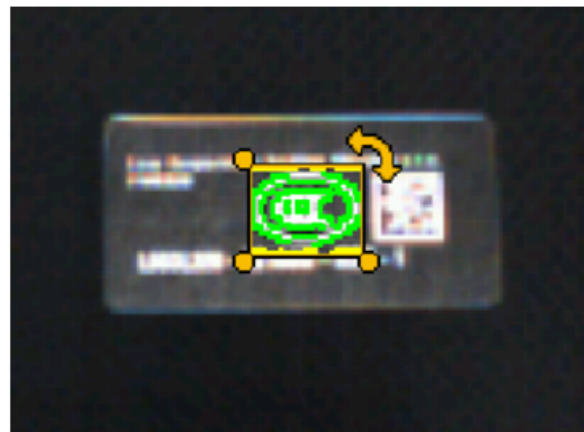
**Figure 2:** Testing workplace

The hardware configuration consists of assigning addresses to devices and the creation of PROFINET based network. Individual devices such as PLC and HMI are connected to a star topology using a industrial switch. The actual appearance of the testing workplace is shown in Figure 2. All components are mounted on a simple aluminum construction for easier work and easier error correction.

The software solution implements the process of checking the presence of the label. It provides communication with the superior LPS system and at the same time solves the communication with the camera system sensor. The basis of the program is a universal function block, which vain part of the automatic control. This function block is also implemented directly on production line and can be used to control identical sensors in production. The rest of the program procures the primary components directly for the test panel, which simulates communication with LPS on human machine interface (HMI).

## 2.2. CAMERA SENSOR SETUP

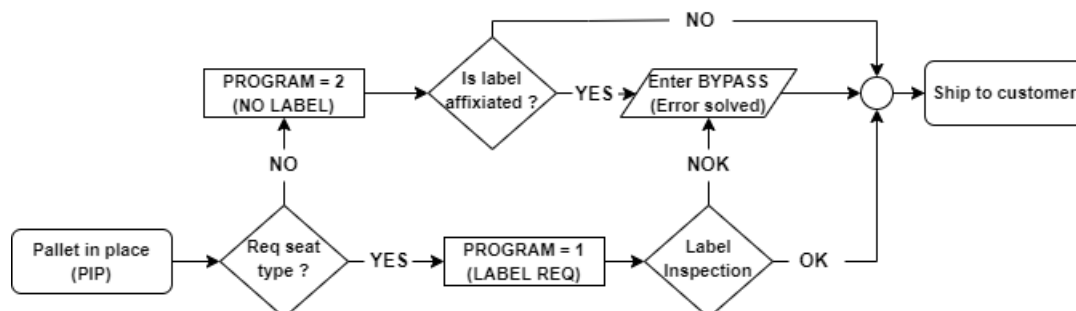
In the IV Navigator environment, which is used to configure the sensor, we implement two programs, which evaluates and sends the label information to the PLC. The first program performs CCC character recognition and the other recognizes black background if the label is not affixed. Figure 3 shows a sample image from the Keyence IV camera sensor. Obviously, the image quality is not high, but it is sufficient for the given issue.



**Figure 3:** Camera program for outline detection

## 3. FINAL IMPLEMENTATION ON THE PRODUCTION LINE

The solution on the production line is implemented by two camera sensors which simultaneously scan the position of the label on two seats. The sensors are checking the presence/absence of the label in the correct position and quality. Evaluation of the presence/absence, position and quality of the label is ensured by Poka-Yoke control. In case of an undesirable NOK (Not OK) result from any specified condition. The LPS system ensures that the seat does not leave the station and therefore cannot be sent to the customer. A positive OK result releases the seat from the station.



**Figure 4:** Label inspection flow chart



**Figure 5:** Final implementation

#### 4. CONCLUSION

The work deals with the introduction of automated presence/absence control certification label for car seats placed on the Chinese market, with the introduction of inspection label quality using a camera sensor. This solution is a demonstration of the implementation of simple automation. It has brought a significant improvement in the quality of products in the company and is being implemented on several other production lines.

A similar approach can be applied in different companies and can thus help reduce production costs and increase product quality.

#### ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-20-6205—“Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0” financially supported by the Internal science fund of Brno University of Technology.

#### REFERENCES

- [1] Lear [online]. Lear Corporation, 2021 [cit. 2021-02-15]. URL <<https://www.lear.com>>
- [2] Fundamentals of Industrial Quality Control, Third Edition [online]. CRC Press, 2018. ISBN 9780203755266. Dostupné z: doi:10.4324/9780203755266
- [3] PÖTTERS, Patrick, Robert SCHMITT a Bert LEYENDECKER. Effectivity of quality methods used on the shop floor of a serial production - how important is Poka Yoke?. Total quality management & business excellence [online]. ABINGDON: Routledge, 2018, 29(9-10), 1200-1212 [cit. 2022-03-05]. ISSN 1478-3363. doi:10.1080/14783363.2018.1488559
- [4] LV, Zhihan, Jinkang GUO a Haibin LV. Safety Poka Yoke in Zero-Defect Manufacturing Based on Digital Twins. IEEE transactions on industrial informatics [online]. IEEE, 2022, 1-1 [cit. 2022-02-27]. ISSN 1551-3203. doi:10.1109/TII.2021.3139897

# Acoustic detection of starlings through artificial intelligence

B. Zezula<sup>1</sup>, P. Marcoň<sup>2</sup>

<sup>1</sup>Bachelor Degree Programme, Brno University of Technology

<sup>2</sup>Associate professor, Brno University of Technology

E-mail: [xzezul08@vut.cz](mailto:xzezul08@vut.cz), [marcon@vut.cz](mailto:marcon@vut.cz)

**Abstract**—To protect the vineyards from starlings is very costly and ineffective with the available resources. In my work, I have designed a detection module which consist of four microphones, each one is directed to the cardinal point and based on the intensity of sound, the module will provide the information about the direction where the flock is located, in the vineyard that passes to the control module and scares the flocks. The detection module processes the signal in a Raspberry Pi 4 single board computer using an artificial neural network algorithm.

**Keywords**—CNN, spectrogram, starlings, Raspberry Pi, Artificial neural network, microphone array, detection of bird singing

## 1. INTRODUCTION

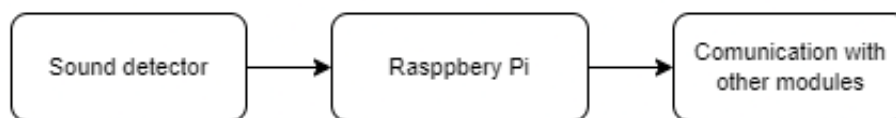
Flocks of starlings pose a great problem on viniculture because they can damage a great part of wine crops in a short time. Vine growers defend their plants mostly by using propane sound cannons, which scare the starlings in regular intervals. Sound cannons are a cheap solution, but the starlings get used to the regular sound blasts, and so the cannons come out ineffective. Growers can scare the flocks also by using armed guards but it's harder to protect the vineyard all the time. There are many methods to protects the crop however, they may be either over time become ineffective or are too expensive.

Previously, Marcoň and his colleagues used cameras and artificial neuron sites to detect and scare birds from ripening fruits [1]. Recently, microphones arrays combined with delay-and-sum beamforming methods were used to detect mallards' hat are occluded by stalks or grass [2]. The author revealed that using a properly assembled microphone system, the source of a certain sound can be recognized together with its position in the field. Several methods showed an efficient in the detection however, there is some difficulties to record or visualize the starling. To overcome these drawbacks, my research work focused on the possible extension of a system incorporating a camera by a module, which will process the acoustic signal. The designed module uses a system of directional microphones to record the sounds in a vineyard. The recorded sound is pre-processed and sent to a Mel spectrogram and then classified CNN (Convolutional neural network). The layout and directions of each microphone in the system is crucial to evaluate the direction in which the starlings are. The goal is to detect direction with accuracy circa  $\pm 45$  degrees.

## 2. MATERIALS AND METHODS

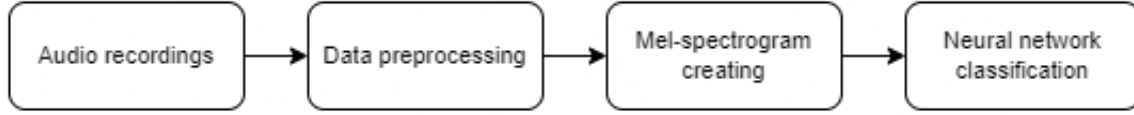
The proposed module in Fig. 1 consists of three parts, a sound detector, a single-board RPi computer (Raspberry Pi 4) and communication with the scaring system.

The detection module software consists of several parts, which are shown in Fig. 2. In the first phase, the sound needs to be recorded and initially processed. The preprocessed signal is converted to a mel spectrogram. The created images are inserted into the CNN input.



**Figure 1:** A block diagram of the detection module.



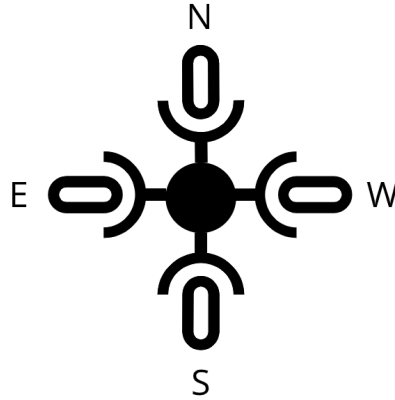


**Figure 2:** A block diagram of the algorithm.

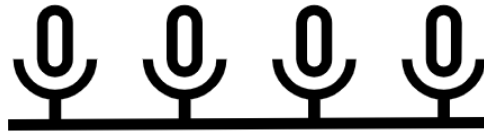
### 2.1. Sound recording

To record birdsong in the vineyard, it is necessary to use microphones with a large detection distance and to arrange them in such a way that the entire area of the vineyard is covered. Outdoor sound recording is affected by noise caused by wind, leaf movement and traffic with frequency below 1 kHz. To process further recording it is necessary to filter the amount as much as possible. Therefore, suitable microphone types include a shotgun type condenser microphone or an electret microphone with a parabola. In our work, we used condenser microphones to have efficient sensitivity.

Four directional microphones connected to a sound card to record the sound. These microphones are each pointed in one direction as shown in Fig. 3. The resulting direction of the signal source is determined from the magnitude of the amplitudes of each microphone. Another option was considered as to use an array of four or more microphones that would be pointed in one direction, as in Fig. 4.



**Figure 3:** Omnidirectional arrangement of microphones



**Figure 4:** Arrangement of microphones in one row

### 2.2. Data preprocessing

The conversion of the audio signal to a mel-spectrogram [4] were used to determine the bird species based on the recording as accurately as possible. The mel-spectrogram has frequencies scale which is converted to the mel-scale as compared to the normal spectrogram. The Mel scale is based on knowledge that one cannot perceive sound on a linear scale. The variation among the high frequencies is very hard to distinguish however at low frequencies it is easier to distinguish. The pitch unit in the mel-scale is such that the difference in tones throughout the range sounds the same [5]. To convert the frequency to the frequency expressed in mel scale, formula 1 is used, where  $f$  is the frequency of the sound.

$$Mel(f) = 1127 \ln 1 + \frac{f}{700} \quad (1)$$

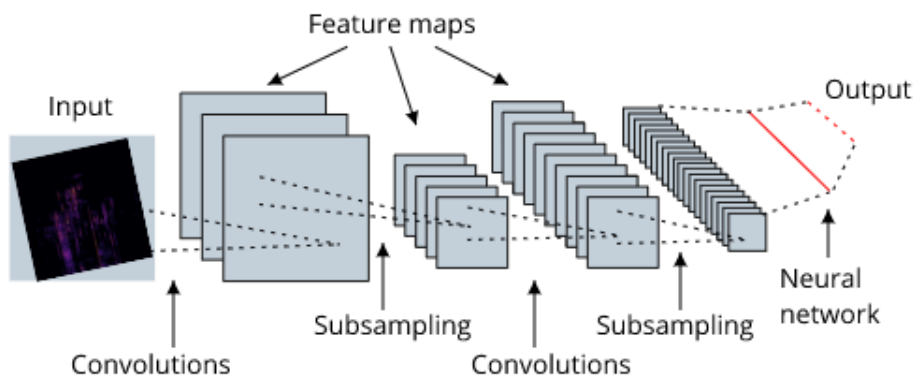


### 2.3. Software for classification

Different neural network topologies and algorithms can be used to classify the audio recordings. The algorithms considered as a different property. An important aspect of choosing the right algorithm is based on the resulting accuracy. It is necessary to use a suitable neural network for different input data.

Convolutional neural network is a suitable method for processing images or audio signals to converted into image forms. The architecture can be divided into a convolution part, a subsampling part and a fully connected neural network. The alternation of the convolution part with subsampling is shown in Fig. 5. These layers can intersect more times than shown in the figure. Finally, the data is processed in a fully connected neural network.

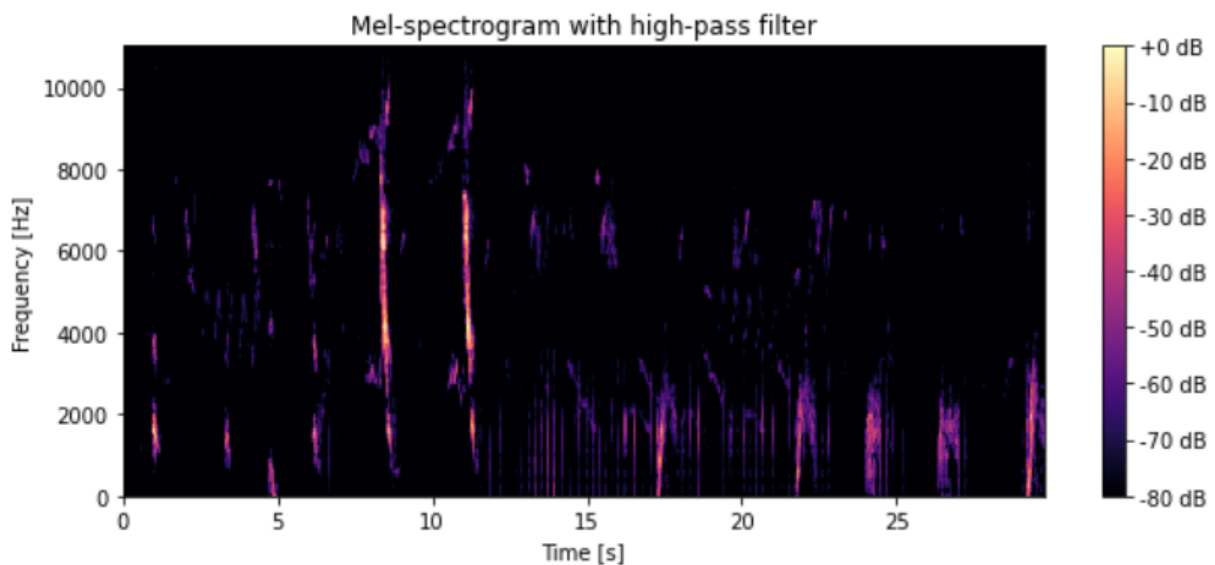
In my work, I deal with the implementation of CNN using the library for Python TensorFlow. The advantage of using a framework such as TensorFlow is that it simplifies the debugging of the CNN algorithm. I use TensorFlow because it is one of the most widely used frameworks.



**Figure 5:** Convolutional neural network architecture

### 3. RESULTS

The work is based on the preprocessing of sound and creation of mel-spectrogram data set for CNN software. The microphone field will be realized according to Fig. 3 and the output of the preprocessing is on Fig. 6 in the figure are seen patterns which is typically for starlings (7s - 13s).



**Figure 6:** Mel-spectrogram with filtered low frequency

The dataset consists of recordings of sounds without starlings singing and recordings with too much noise and the low frequencies were removed using high-pass filters as mentioned earlier (2.2. Data preprocessing). To convert raw signal to spectrogram we used higher frequency spectrum range of 1.5 kHz and the setting of the processed frequency band is based on the knowledge of the starling singing frequency [3]. The mel-spectrograms are made using Librosa library for Python. The spectrogram is created from the time windows of the signal frequency spectrum. The window length is 1024 and the hop length is also 1024.

#### 4. CONCLUSION

The Xeno-Canto database was used to create our own dataset of birds' songs. The Fig. 6 is part of the training data that we used to learn CNN. The dataset consists of mel-spectrogram images containing the sound of starlings, other birds, and the sound of leaves. The mel-spectrogram method is one of the most accurate and widely used methods in the recognition of birds singing [6] by CNN.

Future work will be to program CNN software and create the detection module hardware and this is based on the python module by using the TensorFlow library. Using a neural network framework will shorten the software development in time because the available frameworks were already tested.

#### ACKNOWLEDGMENT

The authors acknowledge financial support from the Technology Agency of the Czech Republic (TJ04000441—A Starling Scaring System Using a Passive Optical Locator) and the general student development project supervised by Brno University of Technology.

#### REFERENCES

- [1] MARCOŇ, Petr, Jiří JANOUŠEK, Josef POKORNÝ, et al. A System Using Artificial Intelligence to Detect and Scare Bird Flocks in the Protection of Ripening Fruit. *Sensors* [online]. 2021, 21(12) [cit. 2021-12-31]. ISSN 1424-8220. Dostupné z: doi:10.3390/s21124244
- [2] MADOKORO, Hirokazu, Satoshi YAMAMOTO, Kanji WATANABE, Masayuki NISHIGUCHI, Stephanie NIX, Hanwool WOO a Kazuhito SATO. Mallard Detection Using Microphone Arrays Combined with Delay-and-Sum Beamforming for Smart and Remote Rice—Duck Farming. *Applied Sciences* [online]. 2022, 12(1) [cit. 2021-12-31]. ISSN 2076-3417. Dostupné z: doi:10.3390/app12010108
- [3] BEASON, Robert C. What can birds hear?. *Proceedings of the Vertebrate Pest Conference*. 2004. Dostupné také z: <https://escholarship.org/uc/item/1kp2r437>
- [4] KAHL, Stefan, Connor M. WOOD, Maximilian EIBL a Holger KLINCK. BirdNET: A deep learning solution for avian diversity monitoring. *Ecological Informatics* [online]. 2021, 61 [cit. 2021-12-31]. ISSN 15749541. Dostupné z: doi:10.1016/j.ecoinf.2021.101236
- [5] LELAND, Roberts. Understanding the Mel Spectrogram. *Medium*[online]. 2020 [cit.2021-12-30]. Dostupné z: <https://medium.com/analytics-vidhya/understanding-the-mel-spectrogram-fca2afa2ce53>
- [6] BRIGGS, Forrest, Balaji LAKSHMINARAYANAN, Lawrence NEAL, Xiaoli Z. FERN, Raviv RAICH, Sarah J. K. HADLEY, Adam S. HADLEY a Matthew G. BETTS. Acoustic classification of multiple simultaneous bird species: A multi-instance multi-label approach. *The Journal of the Acoustical Society of America* [online]. 2012, 131(6), 4640-4650 [cit. 2021-12-31]. ISSN 00014966. Dostupné z: doi:10.1121/1.4707424

# Properties of Current Signals in Nanopore Sequencing

V. Plocková<sup>1</sup> and K. Sedlář<sup>1</sup>

<sup>1</sup>Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology, Czechia

E-mail: [xplock00@vut.cz](mailto:xplock00@vut.cz), [sedlar@vut.cz](mailto:sedlar@vut.cz)

**Abstract**—Oxford Nanopore technologies brought new and revolutionary technology in the field of DNA sequencing. Their sequencing device measures changes in the electric current flowing through pores together with DNA. This work aims to describe differences between raw signals produced by various sequencing kits and sequencing flowcells while sequencing several different bacteria. Two datasets combining five different organisms, two sequencing kits, and two types of flowcells were used to analyze various statistical parameters that would be suitable for the description of current signals gathered from nanopores.

**Keywords**— current signal, Oxford Nanopore Technologies, sequencing, statistical parameters, clustering.

## 1. INTRODUCTION

Nanopore sequencing is a hot topic of bioinformatics, currently being mentioned in a large number of scientific articles. In 2014, the company Oxford Nanopore Technologies (ONT) released its first portable nanopore sequencing device, thus enabling a revolution in sequencing, mainly due to the ability of DNA sequencing anytime and anywhere without the need for a laboratory. Nanopore sequencing has the potential to offer cost-effective genotyping, high mobility for testing, and fast real-time sample processing [1]. The principle of this technology is quite simple. First, you need to prepare the library and then place the sample in the sequencing device. After the voltage is applied, the DNA molecule begins to pass through the nanopore and begins to generate ion current. Current changes correspond to individual nucleotides of DNA. It allows fast sequencing of long individual DNA molecules [2]. This technology also offers an important tool in the fight against antimicrobial resistance. Usually, raw currents are immediately decoded into DNA sequences in a process called basecalling. A majority of studies sees basecalling as a black box using artificial intelligence and neural networks without working with signals themselves [3]. Here, we are dealing with the properties of raw signals, which is a neglected topic of nanopore sequencing.

## 2. MATERIALS AND METHODS

Data for this article were provided from a database of sequences at Department of Biomedical Engineering, FEEC, BUT. The raw data were stored in FAST5 format. Basecalling was done using Guppy [4] to get FASTQ files. For subsequent analysis, two datasets were created.

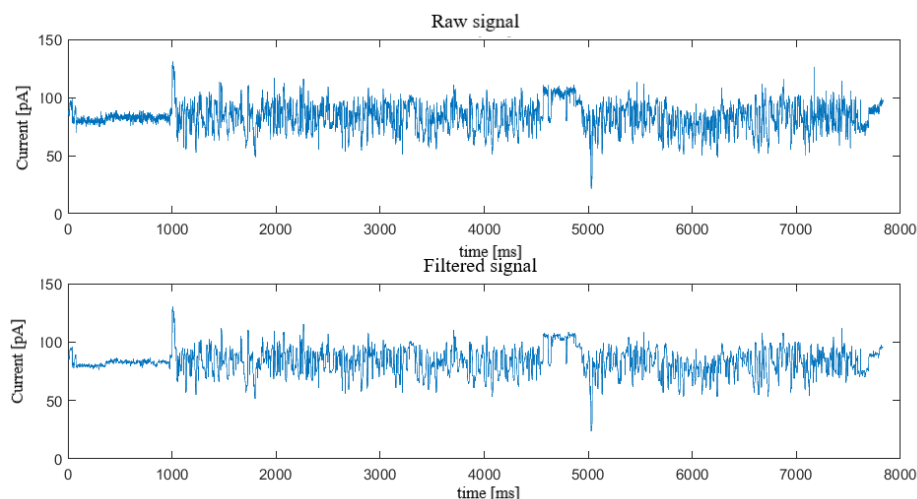
The first dataset was created using a single organism *Clostridium beijerinckii* to find out if the calculated statistical parameters of different signals, sequenced from one organism, differ.

The second dataset consisted of eight samples containing data from five organisms. The samples were selected to represent as many types of sequencing kits (LSK-Ligation Sequencing Kit vs. RBK-Rapid Barcoding Kit) and flowcells (MinION vs. Flongue) as possible. The dataset was created from the following organisms: *C. beijerinckii* (CB; MinION, LSK), *Klebsiella pneumoniae* (KP\_1, MinION, RBK), *K. pneumoniae* (KP\_2; Flongue, RBK), *K. pneumoniae* (KP\_3; Flongue, LSK), *K. pneumoniae* (KP\_4; MinION, LSK), *Pantoea agglomerans* (PA; MinION, LSK), *Schlegelella thermodepolymerans* (ST; MinION, LSK), and *Janthinobacterium lividum* (JL; MinION, RBK).

The first thing that needed to be done was signal preprocessing. The signals in FAST5 files are compressed, i.e., originally measured signals in pA are further converted and stored as 16-bit integer

values, so it was necessary to convert the signals back to picoampere values. The next step was median filtering, despite that ONT states on its website that signals no longer need to be further filtered [2].

The signals turned out to be disturbed by impulse noise, see Fig 1. If the data were not filtered, the following analysis would be affected by outliers and calculations could indicate erroneous results. This could lead to bias in the cluster analysis. To filter signals a median filter was used, the window length was set to five. This window size was selected because a larger window could filter out individual k-mers of length five that correspond to individual current levels [5].

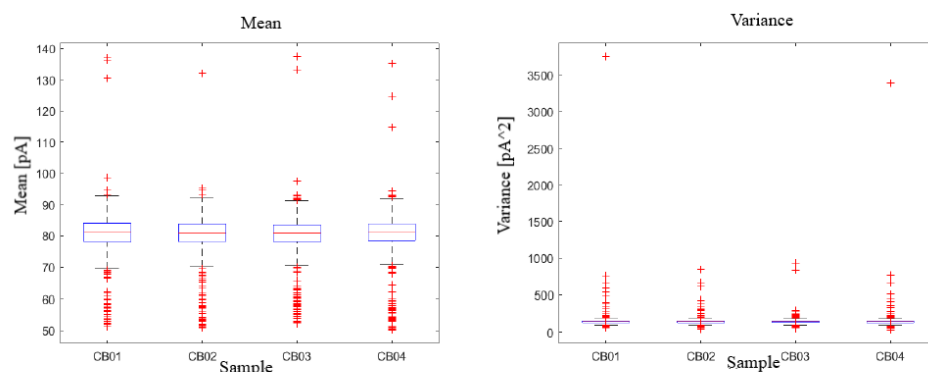


**Figure 1:** Signals before and after preprocessing

Then signals were described by chosen statistical parameters and statistical tests listed below were performed. The last step was hierarchical clustering. First, the distance matrix (or p-distance) was calculated when the Euclidean distance was chosen for calculation of the distance matrix. Furthermore, for mean and basecalled sequences hierarchical clustering using the c-link method was performed.

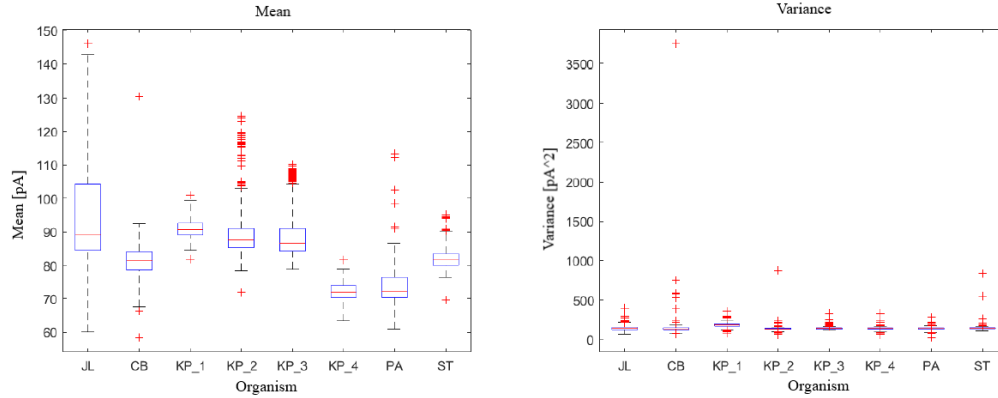
### 3. RESULTS AND DISCUSSION

Each signal in the dataset was described by statistical parameters. These are the mean, variance, standard deviation, coefficient of variation, skew coefficient, sharpness coefficient, and Hjorth descriptors, which include activity, complexity, and mobility. The data were plotted using box plots, see Fig 2 and Fig 3. (other parameters has not been shown)



**Figure 2:** Mean value and variance of signals of the first dataset

Significant differences between the Fig 2 and Fig 3 can be seen. While the values of mean and variance for the first dataset do not differ significantly, the boxplots overlap, for the second dataset, we can see that each organism produces different signals, and even sequencing kits and flow cells produce different signals and are distinguishable from each other. This demonstrates that filtered nanopore current signals could be used to distinguish between various organisms and various sequencing kits and flowcells.



**Figure 3:** Mean value and variance of signals of the second dataset

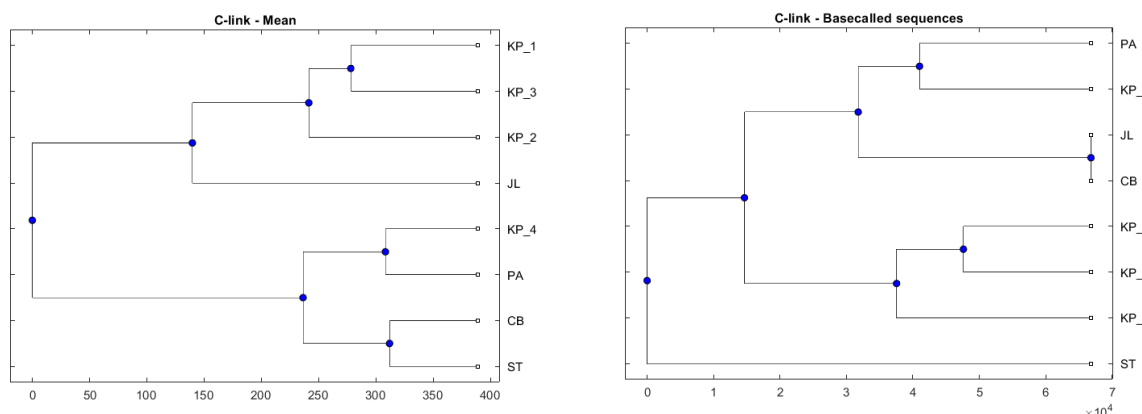
In order to make the evaluation objective, statistical tests were calculated. The first – Shapiro-Wilk to find out if the data has a normal distribution and what statistical test to use next.

Since the data were not normally distributed, the Kruskal-Wallis test was used. This test is just an intermediate step before Tukey’s test, in order to exclude parameters unsuitable for further analysis. In Table 1 you can see the results for selected statistical parameters for the first and the second datasets. The results show that for the first dataset, no statistical parameter can distinguish the samples from each other, but the results for the second dataset were already more satisfactory – most parameters could distinguish one sample from at least one other

**Table I:** The result of Kruskal-Wallis test for the first and second dataset

Statistical parameters	p-values of the first dataset	p-values of the second dataset
mean value	0,2799	0
standard deviation	0,2687	0
coefficient of variation	0,2349	0
skew coefficient	0,3823	2.4889e-198
sharpness coefficient	0,0283	4.1757e-168
variation	0,2687	0
mobility	0,2071	0
activity	0,2687	0
complexity	0,7325	0

Tukey’s test was then performed (results has not been shown). The results for the first dataset are that the individual sample pairs do not differ from each other. In the second dataset, the samples are best distinguished by the coefficient of variation, mobility and mean. Other parameters with satisfactory results are skewness and complexity because the sample KP\_1 to KP\_4 is one organism, so they should be difficult to distinguish from each other and this is not entirely the rule. The only difference is the different library preparation procedure and that the samples were sequenced on different flowcells. The comparison of this parameter with standard clustering of sequences is shown in Fig 4, you can see differences between results of cluster analysis. In the basecalled sequences we can see the correct evaluation of the *Klebsiella* samples, only the KP\_3 sample clustered far from the other samples from the same organism. This may be due to samples sequenced on the Flongse flowcell that can give poor results. But if we compare it to clustering of mean parameter where *Klebsiella* samples (and not only *Klebsiella* ones) are clustered completely differently, this leads us to the idea that sequencing kits and flowcells also need to be considered when a following analysis and basecalling are performed.



**Figure 4:** Example of clustering basecalled reference sequence and one of the statistical parameters (in this in it is mean)

#### 4. CONCLUSIONS

The analysis of the first dataset showed that the signals are species-specific. It does not matter which signals we choose. Although they represent random parts of a genome, their signals properties are the same. This is important for the classification of organisms directly from the raw signals without the need of basecalling. The same analysis was repeated on the second dataset containing five organisms, two different sequencing kits, and two flowcells. Some parameters were able to distinguish among organisms and devices better than others, for example, mean value. Some results brought a surprising finding that even in the case of sequencing one organism we can get completely different results. This is most obvious for the parameter of mobility, where various samples of the organism *K. pneumoniae* show significant differences. The possible explanation is, as mentioned above, the use of various ONT tools. This may play a role in the following decoding of signals by different neural networks and, if this fact is considered, the accuracy of ONT technology might be theoretically improved. In conclusion, we can say that raw signals can distinguish organisms, but they don't have to always be completely correct, and it will probably be necessary to filter the signals and eliminate outliers.

#### ACKNOWLEDGMENT

Computational resources were supplied by the project "e-Infrastruktura CZ" (e-INFRA CZ LM2018140) supported by the Ministry of Education, Youth and Sports of the Czech Republic.

#### REFERENCES

- [1] J. Miten. "Bioinformatic Analysis of Nanopore Data," in Nanopore sequencing: An introduction, vol 1. Word Scientific., 2019, pp. 147-158. [Online]. Available: [https://www.worldscientific.com/doi/abs/10.1142/9789813270619\\_0009](https://www.worldscientific.com/doi/abs/10.1142/9789813270619_0009)
- [2] "How nanopore sequencing works." Nanoporetech.com Available: <https://nanoporetech.com/how-it-works>
- [3] R. R. Wick, L. M. Judd, and K. E. Holt, "Performance of neural network basecalling tools for Oxford Nanopore sequencing", *Genome Biology*, vol. 20, no. 1, 2019. doi 10.1186/s13059-019-1727-y
- [4] Oxford Nanopore Technologies, Oxford, United Kingdom. *Guppy Protocol*. (2020). [Online]. Available: [https://community.nanoporetech.com/protocols/Guppy-protocol/v/gpb\\_2003\\_v1\\_rev14dec2018](https://community.nanoporetech.com/protocols/Guppy-protocol/v/gpb_2003_v1_rev14dec2018)
- [5] F. J. Rang, W. P. Kloosterman and J. De Ridder, „From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy“, *Genome Biology*, vol. 19, no. 1, 2018



# Analyzing parameters of drawing

T. Šilerová

Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech Republic

E-mail: xsiler06@vutbr.cz

**Abstract:** This article deals with analyzing parameters of drawing tests that are used in psychodiagnostics. A tablet and stylus were used to acquire the data, replacing a paper and a pencil used nowadays in conventional drawing tests administration. New parameters have been developed to assess proband performance and the recorded data were evaluated.

**Keywords:** Drawing tests, Rey-Osterrieth figure, drawing parameters

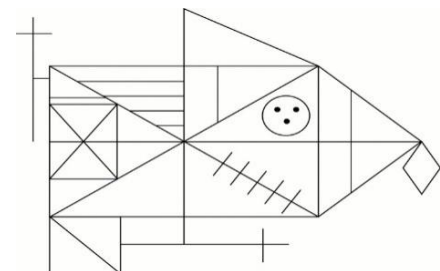
## 1. INTRODUCTION

Drawing tests are widely used in psychodiagnostics. To assess the level of development, to diagnose perceptual-motor functions and neuropsychological disorders, to measure an individual's attention, to diagnose cognitive deficits or they can be also used as intelligence tests. [1]

Currently, drawing tests are subjectively evaluated by an evaluator who observes the behavior and drawing of the proband during the test. Thus, the final evaluation of a drawing may be influenced by the evaluator's mental state, his or her experience, and how he or she focuses on changes in an individual's behavior during testing. The current approach also cannot evaluate the dynamics of the drawing. The ideal way to eliminate these problems, simplify the evaluation of the drawing and enable the evaluation of the drawing dynamics is to replace pen and paper approach with a digital record of the drawing and then perform the evaluation using the parameters that were designed. Tablet and stylus recording is used, for example, to evaluate handwriting in patients with Parkinson's disease or to identify and evaluate developmental dysgraphia by handwriting analysis. [2] [3]

## 2. DATA RECORDING

Data were measured on 25 adult probands aged 20 to 24 years. The probands were asked to redraw Rey-Osterrieth's figure according to the presented template. After three minutes and then thirty minutes without warning, they drew a figure from memory. This is common practice. Rey-Osterrieth's figure consists of rectangles, lines, circles, triangles, and other geometric components designed to require a minimum of drawing requirements. (Figure 1)

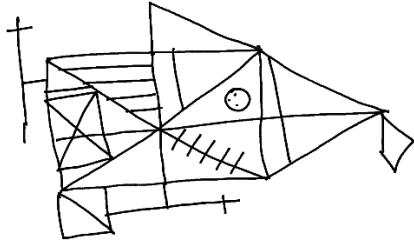


**Figure 1:** Rey-Osterrieth figure [4]

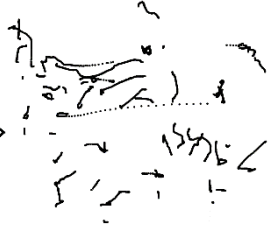
Before drawing, each proband filled out a questionnaire that included information about age, gender, education, whether the proband is left-handed or right-handed, and questions about how the proband's performance may be affected (e.g., drug effect, migraine, whether the proband is being treated for depression or anxiety, whether the proband has been diagnosed with some form of epilepsy, learning disability or attention deficit disorder, etc.). The same conditions were ensured for all probands, the light was switched on so that there was no shading on the tablet, everyone was sitting on a chair in a position that was comfortable for them, and the chair was always placed in the same place at the table. They all drew with their dominant hand and were also instructed not to touch the template. Another important piece of information was whether the proband wore glasses and if so, he had to wear them. None of the participants had a hand injury that could affect the drawing.

## 3. LINE SEPARATION

During the drawing, the movements of the stylus when touching the tablet – touch movements - are recorded, as well as the movements above the tablet in proximity the drawing surface - in-air movements. To quantify parameters and evaluate the dynamics of the drawing, it is first necessary to separate touch and in-air movements from each other, since only touch movement are further used. Information on the intensity of the stylus pressure on the tablet was used for this purpose. The touch movement curves were displayed in color as individual lines, as can be seen at figure 2 - 4.

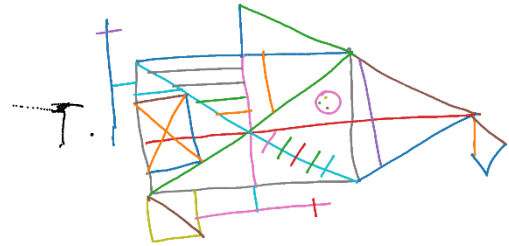


**Figure 2:** Touch movements



**Figure 3:** In-air

**Figure 4:** Individual lines



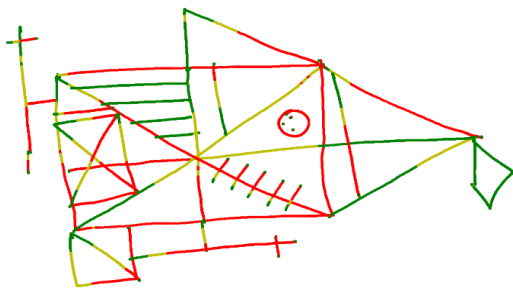
#### 4. DESIGN OF PARAMETERS

New parameters have been designed to objectively evaluate the dynamics of the drawing: tablet pressure rate, line speed and acceleration, line drawing sequence and whole drawing procedure, drawing time, in-air time, number of pauses and total testing time.

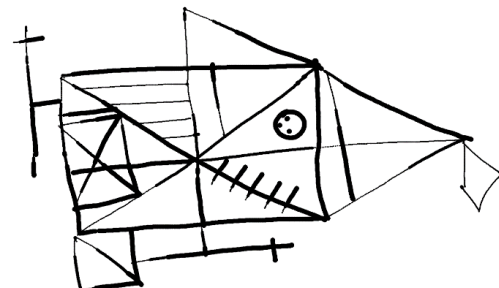
##### 4.1 TABLET PRESSURE RATE

The proband's mental effort is reflected in the pressure rate. However, the change in pressure can be caused both by the stylus's holding style and the proband's mental state. In example, if the proband is angry, lines that have been drawn with great force may appear. The same line appearance is observed when holding the stylus with fist, such as in children's drawings.

To differentiate mental effort from stylus holding style, the relative value of pressure that the proband developed during the drawing on the tablet, was calculated. The relative pressure rate is expressed as a pressure relative to the highest value that the proband has developed over the entire drawing period. In this view, evaluator can see how the test person changes the pressure on the pencil during the drawing. The result is shown in two ways: by the color code (green = light, yellow = moderate, red = strong) and the line thickness (thin line = light, thick line = strong), as can be seen at figure 5 and figure 6.



**Figure 5:** Pressure rate in color



**Figure 6:** Pressure rate by line thickness

##### 4.2 LINE SPEED AND ACCELERATION

The time stamps and line coordinates were used to calculate this parameter. The total drawing time of a given line was obtained by subtracting the line end time value from the line start time value. The length was determined by summing the distances between all recorded line points. The speed was calculated:

$$v = \frac{\Delta s}{\Delta t} \quad (1)$$

$s$  is the calculated line length and  $t$  is the time the proband drew the line

First, the speed of all lines was calculated according to formula 4-1. These times were stored and sorted

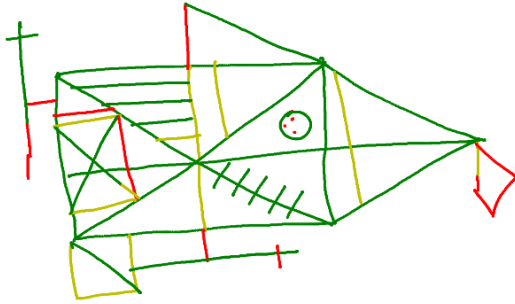
in ascending order, and they were divided using to percentile range, into three groups and color-coded the slowest (red) and fastest drawn lines (green). Result is shown at figure 7.

Acceleration expresses the change in speed per unit time. The display therefore shows in which parts of the figure the proband slowed down or accelerated his or her drawing. The acceleration of each point of the line was calculated according to formula 4-2.

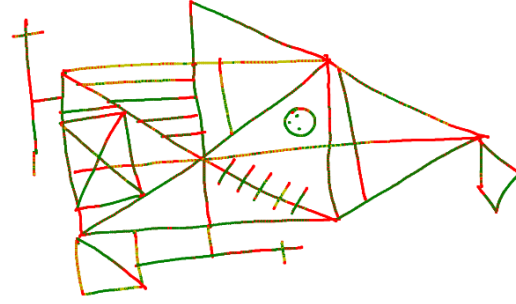
$$a = \frac{\Delta v}{\Delta t} \quad (2)$$

$v$  is the speed and  $t$  is the time

The speed of each line was calculated as in the previous chapter and time was calculated as the difference between two consecutive coordinates. The parts of the figure that were drawn at low acceleration are drawn in red, and the parts that were drawn at high acceleration are drawn in green. (Figure 8)



**Figure 7: Line speed**

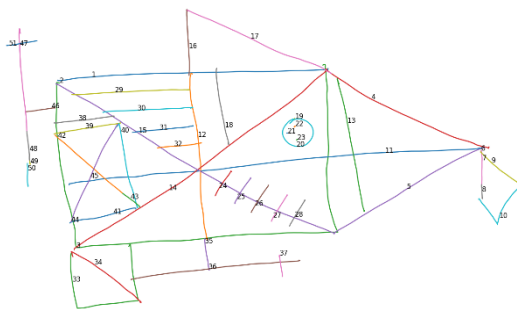


**Figure 8: Acceleration**

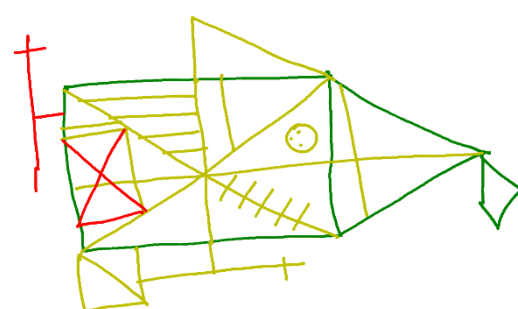
#### 4.3 LINE DRAWING SEQUENCE AND WHOLE DRAWING PROCEDURE

Due to the information obtained during proband testing, it is possible to separate the individual lines and assign them a number based on the order in which they are drawn. From this view, it is then possible to read whether the proband starts drawing outlines and then details or vice versa. (Figure 9)

By viewing the line drawing procedure, you can see the order in which the proband drew the figures. First, the ones created by the movement of the stylus over the tablet - in-air movements - were removed from all recorded lines. The first 20% of the drawn lines were drawn in green and the last 20% of the lines were drawn in red. The range 20% - 80% of the lines were drawn in yellow (Figure 10). This range was chosen because the figure consists of 18 elements and the first 20% of the lines should represent the 4 large elements of the figure. It is further assumed that the proband will continue with the details of the figure, which make up 80% of the figure, and the last 20% should be the side details of the figure.



**Figure 9: Line drawing sequence**



**Figure 10: Whole drawing procedure**

#### 4.4 DRAWING TIME, IN-AIR TIME

The time of individual drawings consists of summary of touch and in-air movements. This time was computed for 1<sup>st</sup> drawing, 2<sup>nd</sup> drawing and 3<sup>rd</sup> drawing. Drawing time is the time a proband takes to draw the figure. It is therefore the sum of the drawing times of all the lines categorized as touch data. The time of in-air movements expresses the time when the proband did not draw but moved the stylus over the tablet. The times of all in-air movements were summed to obtain the in-air time. Parameters were calculated and displayed in a table 1.

#### 4.5 NUMBER OF PAUSES

The total in-air movement time is related to a parameter that indicates the number of pauses the proband made during drawing. In the evaluation, it is important whether the proband's time over the tablet is long and has a small number of pauses lasting a long time, or whether the proband's drawing consists of a large number of pauses that last a short time. (Table 1)

Time of individual drawing [s]	Drawing time [s]	In-air time [s]	Number of pauses
132.515	51.519	80.989	46

**Table 1:** Time information and number of pauses

#### 4.6 TOTAL TESTING TIME

This parameter provides information about the testing duration of all three drawings. At this time, the drawing time and in-air movements of each drawing that was performed are included (1st drawing, 2nd drawing, and 3rd drawing). This time is again recorded in the table together with the times of the individual drawings. (Table 2)

First drawing [s]	Second drawing [s]	Third drawing [s]	Total testing time [s]
132.515	86.042	57.255	275.812

**Table 2:** Information about first drawing, second drawing, third drawing and total testing time

### 5. CONCLUSION

The article presents the advantages of digital recording and processing of drawing tests. Application of objective parameters brings a lot of advantages in testing evaluation. This evaluation is fast, objective and the dynamics of the drawing can be evaluated. It is also possible to collect data and compare individuals with each other.

The parameters that were designed are the total time of the drawing, the time of in-air movements, the time of the drawing itself (time when the proband touched the tablet stylus), number of pauses, speed and acceleration of individual lines, order of drawn lines and display with which units the proband proceeded drawing. The figure with each parameter is always drawn using color code (green, yellow and red) and some parameters values are expressed using different line weights, which allows the evaluator to quickly obtain diagnostic information from the overall figure preview.

### REFERENCES

- [1] Test obkreslování. Psychodiagnostika [online]. Brno: Psychodiagnostika [cit. 2021-11-30]. Dostupné z: [http://www.psychodiagnostika-sro.cz/cz/Katalog\\_popis.asp?kod=552&ZozArg=1&Kateg=1](http://www.psychodiagnostika-sro.cz/cz/Katalog_popis.asp?kod=552&ZozArg=1&Kateg=1)
- [2] Masarova, Lucia & Drotar, Peter & Mekyska, Jiri & Smekal, Zdenek & Rektorova, Irena. (2014). Assessing Handwriting in Patients with Parkinson's Disease. *Československá neurologie a neurochirurgie*. 77/110. 456-462
- [3] Mekyska, Jiri & Faundez-Zanuy, Marcos & Mzourek, Zdenek & Galáž, Zoltán & Smekal, Zdenek & Rosenblum, Sara. (2016). Identification and Rating of Developmental Dysgraphia by Handwriting Analysis. *IEEE Transactions on Human-Machine Systems*. 47. 10.1109/THMS.2016.2586605
- [4] Bosson S, Holland PC, Barrow SA visual motor psychological test as a predictor to treatment in nocturnal enuresis *Archives of Disease in Childhood* 2002;87:188-19

# Concept Drift Detection in Prediction Classifiers for Determining Gender in Metabolomics Analysis

A. Kostoval<sup>1</sup>, J. Schwarzerova<sup>1,2</sup>

<sup>1</sup>Department of Biomedical Engineering Faculty of Electrical Engineering and Communication, Brno University of Technology Brno, Czech Republic

<sup>2</sup>Molecular Systems Biology (MOSYS), University of Vienna, Vienna, Austria

E-mail: [221515@vut.cz](mailto:221515@vut.cz), [Jana.Schwarzerova@vut.cz](mailto:Jana.Schwarzerova@vut.cz)

**Abstract**—Currently, one of the most challenges in data analysis is connected to prediction modeling including dynamic information. Metabolomics analysis focuses on data presented dynamic information in real-time such as time-series data. Unfortunately, prediction models based on time series data are often affected by a phenomenon called concept drift. This phenomenon can reduce the accuracy of prediction models which is an unwanted effect. On the other hand, concept drift analysis can be useful in finding confounding factors. This study is divided into two parts. The first part presents the modeling of prediction classifiers based on metabolite data. The second part of this study brings concept drift detection in the created classified models. This study presented approaches to identify one of the confounding factors in human biology.

**Keywords**—Concept drift, Concept drift detection, Metabolomics, Machine learning, Prediction modeling

## 1. INTRODUCTION

The concept drift is defined as unexpected changes between input and output data. Sometimes concept drift occurs when test sets are changing unpredictably. Subsequently, the model is unable to respond correctly because training sets were different. It follows, that concept drift negatively affects prediction models which were being trained on dynamically changing data.

Concept drift analysis studies the detection and correction of an unwanted phenomenon that is caused by obvious changes in real-time [1]. These changes are characterized in data distribution and can be detected in a prediction model [2]. Thus, the main goal of concept drift detection focuses to bring algorithms that can detect changes in data distribution. As a result of concept drift detection, we would ensure the long-term accuracy, reliability, and stability of prediction models.

Metabolomics is one of omics science focused on detection, quantification, and analysis of metabolites in an organism [3]. Metabolites are small molecules, smaller than 1500 *Da*. Metabolomics and the study of prediction models based on metabolites data leads to earlier detection of human disease, which is reflected in metabolomes such as diabetes mellitus, cancer, etc. [3]. Nevertheless, prediction models based on metabolite data are affected by the concept drift phenomenon [4]. Therefore, this concept drift phenomenon is necessary to detect and reveal confounding factors that are hidden.

## 2. MATERIALS AND DATA

In this study, data were taken from the cohort study by Chu et al. [5]. Metabolomics datasets include 534 healthy subjects (237 males and 296 females) in ages 18 – 75. The study brought two different datasets which were measured using different techniques. The first dataset represented platform Brainshake Metabolomics (BM) which was measured by principle nuclear magnetic resonance and included 231 features with 200 absolute concentrations. The second dataset represented platform General Metabolomics (GM) based on flow injection TOF-M. GM included 1586 features with 257 absolute concentrations. These data are freely available at: <https://hfgp.bbmri.nl/>.

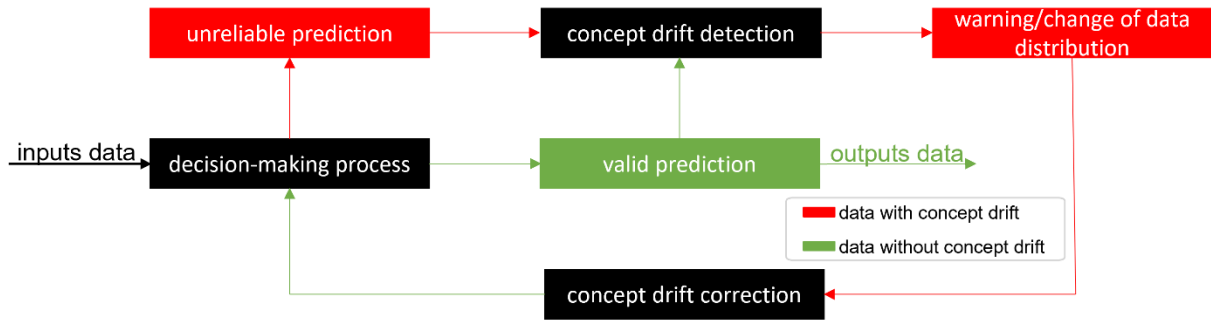
### 3. METHODS

The whole methodology of this study is divided into three parts. The first step includes pre-processing phasis focused on data preparation. The second part is focused on model prediction classifiers using different approaches. The last part is connected to concept drift analysis. In the last part, the concept drift detection was performed, and the confounding factor was identified.

Firstly, datasets were divided into inputs and targets data. Inputs data represent concentrations of metabolites. Targets data classify phenotype of gender into two binary classes (male – 0, female – 1). Furthermore, both datasets were split by 10-fold cross-validation to estimate classify skill of prediction models. Training and testing datasets were divided into a ratio of 9:1 and cross-validation ran 10 times.

Secondly, the prediction classifiers were modeled. The classifiers were trained to determine the gender of a patient based on measured metabolite concentration. Prediction models were implemented in Python using the library Scikit-learn [6]. Namely, we used: Logistic Regression (LR), Gradient Boosting (GB), Random Forest (RF), and Naïve Bayes (NB). All these methods were applied to model classifiers.

The last step of our methodology included concept drift detection. The concept drift detectors were used on metabolomics models. The concept drift principle is shown in Figure 1. Firstly, input data are predicted by a decision-making process. In the next step, the concept drift detector detects deflection in data distribution, detector warns possibility of concept drift presence. Nevertheless, concept drift detectors have two thresholds. The first limit is warning but the second limit announces exact detection of change in data distribution. Finally, these detected changes are appropriately revised and again predicted by a decision-making process.



**Figure 1:** Pipeline of methodology for concept drift detection and correction

The first method which was used in our study is Drift Detection Method (DDM) [3]. DDM is based on the monitor error rate of the classification, which is understood as the probability of incorrect prediction. The second method is the Early Drift Detection Method (EDDM) [3]. EDDM is very similar to DDM, but EDDM analyzes also changes in distance between two consecutive misclassified objects. EDDM is useful to detect gradual changes in data distribution. On the other hand, DDM is better for detecting sudden concept drift. DDM and EDDM were implemented from Scikit-multiflow [7] for each created classifier model.

### 4. RESULTS AND DISCUSSION

The accuracy of our prediction models was estimated using 10-fold cross-validation. Table I shows evaluation parameters presented final accuracy for each method. The highest value of accuracy connects to the GB method, see Table I. Table II presents another metric for evaluating the accuracy of the classification was the F1-score. This metric compares the prediction of classification models with the target. The best accuracy and F1-score achieved the LR model in the case of the BM dataset and the GB model in the case of the GM dataset. On the other hand, the lowest value of accuracy is identified in prediction models based on the NB approach and F1-score values confirm this statement.

**Table I:** The accuracy (dimensionless quantity) values of classification models for datasets BM/GM

	LR	GB	RF	NB
BM	0,85	0,83	0,79	0,69
GM	0,85	0,88	0,84	0,78



**Table II:** The F1-score (dimensionless quantity) values of classification models for datasets BM/GM

	LR	GB	RF	NB
BM	0,81	0,81	0,79	0,73
GM	0,83	0,89	0,87	0,76

The concept drift detectors bring revealing of concept drift in our created models. Table III shows the number of concept drift detection for models created by BM data. Table IV includes the number of concept drift detection in models based on GM data.

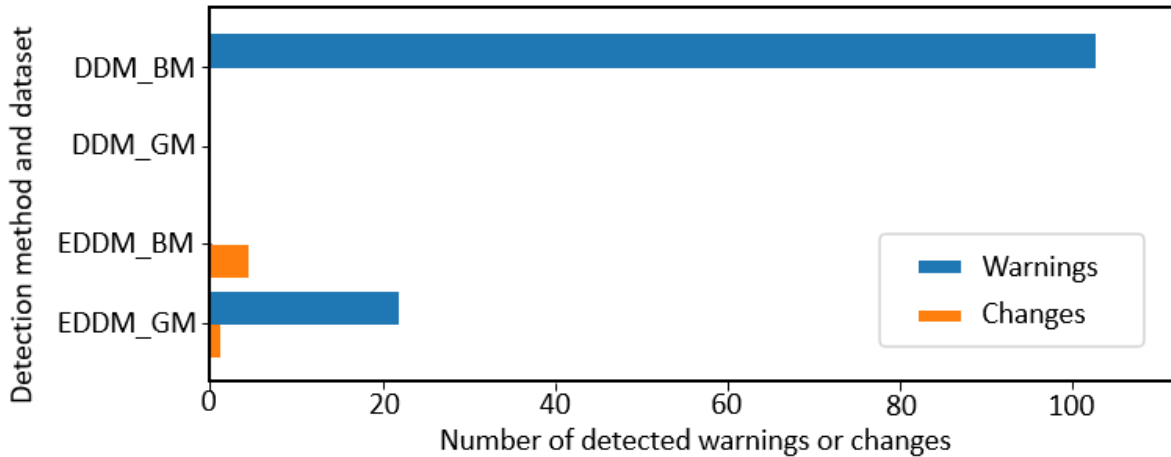
**Table III:** Number of warnings (W) and changes (CH) detection in data distribution in BM dataset

	LR	GB	RF	NB
DDM	35 W / 0 CH	22 W / 0 CH	39 W / 0 CH	17 W / 0 CH
EDDM	0 W / 1 CH	0 W / 1 CH	0 W / 1 CH	0 W / 1 CH

**Table IV:** Number of warnings (W) and changes (CH) detection in data distribution in GM dataset

	LR	GB	RF	NB
DDM	0 W / 0 CH	0 W / 0 CH	0 W / 0 CH	0 W / 0 CH
EDDM	5 W / 0 CH	0 W / 1 CH	18 W / 0 CH	5 W / 0 CH

Precisely, Figure 2 shows the plot for comparison number of concept drift detection using methods DDM and EDDM.



**Figure 2:** Summary concept drifts by detection methods DDM and EDDM. DDM\_BM (EDDM\_BM) is the sum of concept drifts detection by DDM (EDDM) for each classifier model based on BM data. Similarly, DDM\_GM (EDDM\_GM) is the sum of concept drift detection for classifier models based on GM data.

The highest number of warning levels was detected by DDM in RF models. In each of the RF models, data distribution includes significant changes detected as concept drift phenomenon. EDDM detector revealed more concept drift detection than DDM. Therefore, EDDM is more suitable for detecting the concept drift in metabolomic prediction, which corresponds with the findings in the study [4]. In the end, we identified concept drift according to patient age which is a promising factor as a confounding factor in the human metabolomics analysis. Figure 3 shows the detection of concept drift occurring in the adolescent period. Regarding it, the confounding factor presented the age of a patient is identified. Thus, our study brings confirmation revealed confounding factor from the study [4].

**Figure 3:** Visualization of concept drifts detection according to the age of patients

## 5. CONCLUSION

Metabolomics brings a new era allowing to deal with prediction phenotype based on metabolites concentrations. However, metabolomics focuses on data presented dynamic information in real-time such as time-series data. Thus, prediction models based on metabolite data can be included to concept drift which reduces prediction accuracy. This study brings 8 prediction classifiers that predict gender based on metabolites concentrations. The main message of the study is to reveal an innovative view on metabolomics analysis with detecting confounding factors like is concept drift.

In created classification models, the concept drift was detected using DDM and EDDM. The EDDM is more appropriate for concept drift detection in metabolomic prediction than DDM. Thanks to concept drift detection, confounding factor related metabolomics analysis was identified as the age of a patient. This finding will help to create more accurate models for the early diagnosis, which is essential to a full recovery, or economically less demanding treatment.

## ACKNOWLEDGMENT

This work has been supported by grant FEKT-K-21-6878 realised within the project Quality Internal Grants of BUT (KInG BUT), Reg. No. CZ.02.2.69 /0.0/0.0/19\_073/ 0016948, which is financed from the OP RDE.

## REFERENCES

- [1] WEBB, Geoffrey I., et al. Characterizing concept drift. *Data Mining and Knowledge Discovery*, 2016, 30.4: 964-994.
- [2] YU, Shujian, et al. Concept drift detection and adaptation with hierarchical hypothesis testing. *Journal of the Franklin Institute*, 2019, 356.5: 3187-3215.
- [3] IDLE, Jeffrey R.; GONZALEZ, Frank J. Metabolomics. *Cell metabolism*, 2007, 6.5: 348-351.
- [4] SCHWARZEROVA, Jana, et al. An Innovative Perspective on Metabolomics Data Analysis in Biomedical Research Using Concept Drift Detection. In: 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2021. p. 3075-3082.
- [5] CHU, Xiaojing, et al. Integration of metabolomics, genomics, and immune phenotypes reveals the causal roles of metabolites in disease. *Genome biology*, 2021, 22.1: 1-22.
- [6] Scikit-learn: Machine Learning in Python, Pedregosa et al., *JMLR* 12, pp. 2825-2830, 2011.
- [7] MONTIEL Jacob et al., "Scikit-multiflow: A multi-output streaming framework", *The Journal of Machine Learning Research*, vol. 19, no. 1, pp. 2915-2914, 2018.



**Vyvíjíme a testujeme chytrá auta**

Jdeš do toho?



**valeo.jobs.cz**

SMART TECHNOLOGY FOR SMARTER MOBILITY

# Kernel latency analysis on Raspberry Pi OS

M. Mitěv<sup>1</sup> and L. Pohl<sup>2</sup>

<sup>1</sup>Brno University of Technology, Czech Republic

<sup>2</sup>Brno University of Technology, Czech Republic

E-mail: [191958@vut.cz](mailto:191958@vut.cz), [pohl@vut.cz](mailto:pohl@vut.cz)

**Abstract**—This article deals with latency analyze on Raspberry Pi OS and finds out how PREEMPT\_RT patch will affect latency. Article provides insight into creation Real-time application development and how to use this application in latency measurement. Further thesis describing methods for latency measurement.

**Keywords**— Raspberry Pi, GPIO, PREEMPT\_RT patch, cyclictst

## 1. INTRODUCTION

There are currently several methods to create Linux operating system as real-time system. In addition, these efforts have recently been motivated by a growing number of microcomputers based mainly on the ARM platform. Raspberry Pi microcomputers from the Raspberry Pi Foundation are one of the most important representatives of this area. The aim of the article is to apply the PREEMPT\_RT patch to the Raspberry Pi OS operating system and to measure the OS latency after applying this patch. The first section deals with the introduction of real-time system. It also describes the possibilities of implementing a real-time system on Linux and deals with the PREEMPT\_RT patch. The second section describes the RT application, which is used to measure latency. The third section describes the procedures for measuring OS latency and evaluating the results.

## 2. REAL-TIME SYSTEMS

The basic requirement for real-time systems is determinism. Operations are performed at fixed, predetermined times or time intervals. The interrupt response must be such that the system is able to handle all requests at the required time. [2] In other words, no matter what role a real-time application performs, it must perform it not only correctly but also on time. Real-time does not mean that the application should run as fast as possible, but it must perform critical tasks at a specific time, or real-time deals with a guarantee not with pure time.[1]

## 3. REAL-TIME SYSTEM IMPEMANTION ON LINUX

Linux as such is not designed as a real-time operating system but there are methods to adapt it for real-time use. One of the possible solutions used in practice is dual-kernel archiceture. The dual-kernel architecture includes real-time kernel, and Linux runs as low-priority task. The disadvantage of this solution is that it is necessary to maintain micro kernel for new hardware. Furthermore, this solution requires an abstract hardware layer (HAL) on which Linux will run. These systems also often do not allow the use of standard libraries when developing applications, but special tools are needed. Dual-kernel systems include:

- RTAI - Real time application Interface
- Xenomai
- RTILinux

Another solution is to make the Linux kernel itself real-time. This solution is called a single-kernel. The PREEMPT\_RT patch is currently working on this solution. The advantage of the PREEMT\_RT patch is that most features are part of the "mainline" and standard C libraries and POSIX standards can be used when writing real-time applications.[2]

## **PREEMPT\_RT PATCH**

The goal is to make the Linux kernel predictable and provide limited latency. The PREEMPT\_RT patch allows the kernel to be interrupted during system calls to service a higher priority task. The main task of this patch is to increase the degree of code preemption towards the fully pre-emptable kernel (PREEMPT\_RT\_FUUL). This level of preemption allows real-time tasks to preempt the kernel even in critical sections. However, some areas are still not pre-emptive. For example, the upper half of the interrupt handler and critical sections protected by raw spinlock (raw\_spinlock\_t). All "sleeping mutex" have been replaced by rt\_mutex, which implement priority inheritance. The advantage of the PREEMPT\_RT patch is that most features are part of the "mainline" and standard C libraries and POSIX standards can be used when writing real-time applications. [2]

## **4. PROPERTIES OF PREEMPT\_RT PATCH**

The next section will focus on some of the features that the RT patch brings to Linux to meet the requirements of RT systems. This section will address:

- Interrupts as threads
- Hard IRQ like threads
- Softirq timers threads

### **Interrupts as threads**

These are methods that Linux provides to finish the work needed by asynchronous events. These methods are kernel thread, a tasklet, or a softirq. The aim is for the device driver to prioritize these methods instead of putting work into the ISR.[1]

### **Hard IRQ like threads**

Hard IRQ (interrupt request) is kind of envelope around an ISR. Usage of this envelope can lead to interrupt inversion and high priority process does not get CPU time because interrupt service routine performs process with lower priority. RT patch tries to shorten the time of interrupt inversion to a bare minimum by converting the interrupt routine to thread.[1]

### **Softirq timers threads**

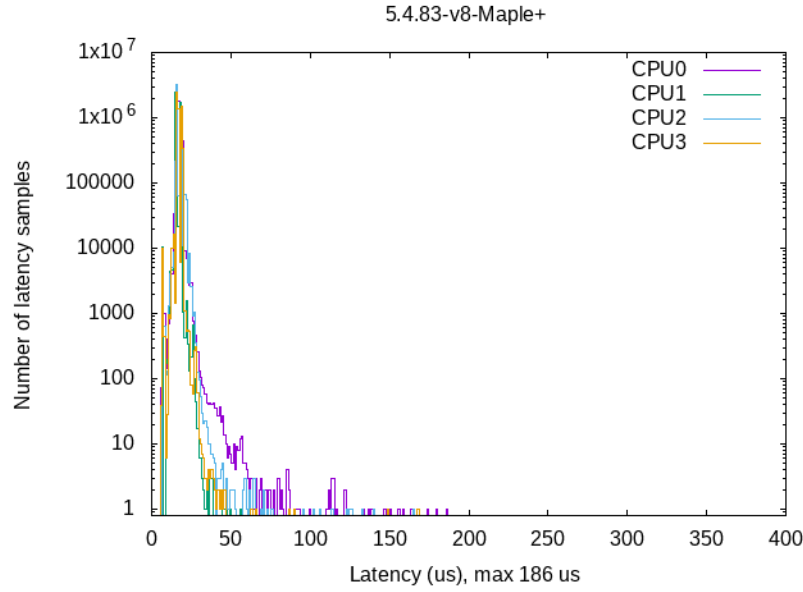
RT patch forces each timer intervals to be at least one jiffy. This is to prevent softirq-hrtimer thread from starving if the process ignores the signal that comes from the timer. Also lower-priority process can not do more than one event per jiffy that would preempt higher-priority process.[1]

## **5. CYCLIC RT APPLICATION**

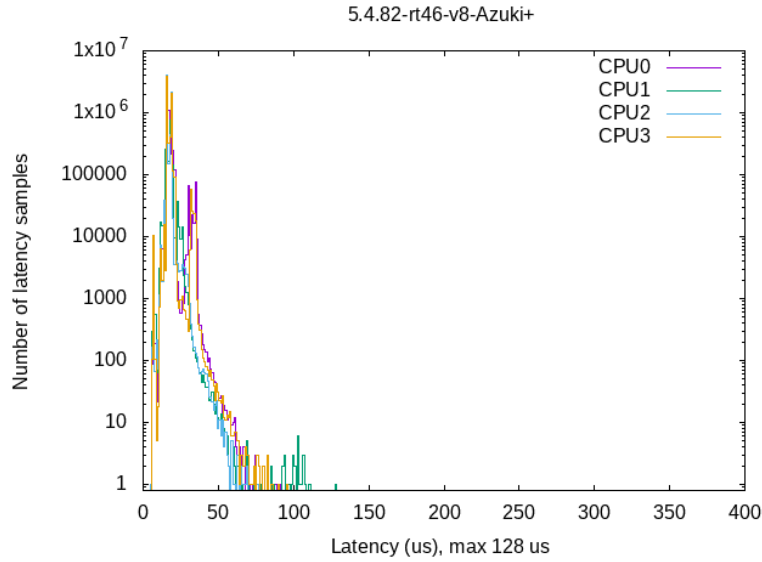
A cyclic RT application is one that repeats after a fixed amount of time, such as reading data from a sensor. The time to execute this application should be less than the period of this task. [3] Two versions of this application were created to measure latency. Where the first version of the application is written for FIFO and RR (round-robin) planning policy and the second version is created for deadline planning policy. Both applications control output pin 18 on the GPIO, which is located on the Raspberry Pi with a period of 1 ms. The gpiod.h function library is used to control GPIO pins. When using a FIFO or RR scheduling policy, the Clock\_nanosleep function is used, which puts the process to sleep for a predetermined time. The CLOCK\_MONOTIC timer is used to determine if this time has already elapsed. This timer is an absolute time that is calculated from a fixed point in the past and is not affected by changes to the system clock. When using deadline scheduling policies, it is necessary to define a system call and a sched\_attr structure that contains the parameters with which the deadline works.

## **6. LATENCY MEASUREMENT USING CYCLICTEST**

Cyclictest is a tool for repeatedly measuring the time difference between the expected time of waking the fiber and the actual time when the thread wakes up. This tool can measure latency caused by hardware, firmware, and OS. This tool is part of the rt-test package. [4] Mklatencyplot is a bash script that runs a cyclictest and creates a histogram from the measured values. the script is dusty from [5].



**Figure 1:** Measurement result using mkl latency plot on Linux kernel without PREEMPT\_RT patch



**Figure 2:** Measurement result using mkl latency plot on Linux kernel with PREEMPT\_RT patches

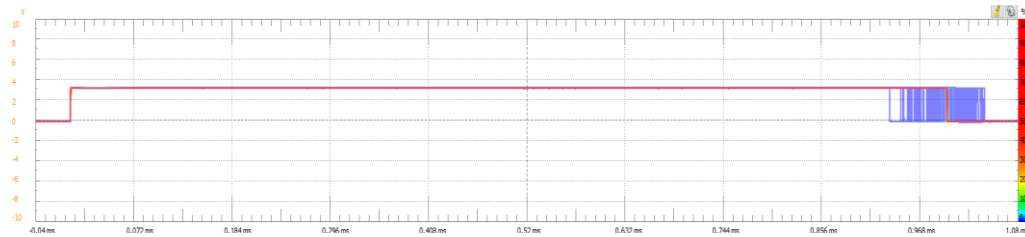
The first measurement took place on a Linux kernel without the PREEMPT\_RT patch and the results are summarized in Figure 2. The graph shows that the highest measured latency is 186  $\mu$ s and that most samples of the recorded latency are below 50  $\mu$ s.

The second measurement was performed on a kernel with the PREEMPT\_RT patch and the results are summarized in Figure 2. The graph shows that most samples of recorded latency are below 50  $\mu$ s and the largest recorded latency is 128  $\mu$ s.

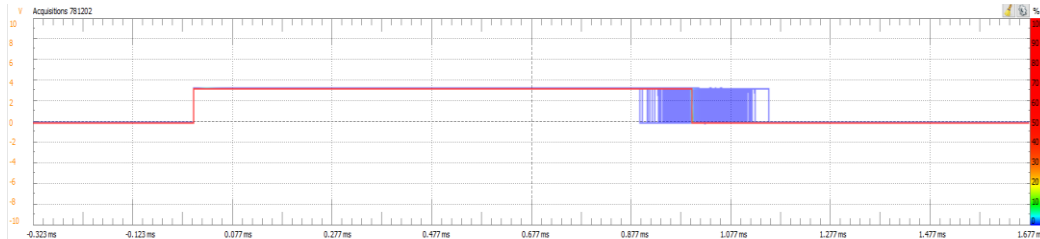
## 7. LATENCY MEASUREMENT USING ANALOG DISCOVERY 2

In the next measurement, the RT application described in Chapter 5 and the Analog Discovery 2 device were used. The measurement was performed by monitoring the time required for the RT application to complete at the output of the digital oscilloscope. Along with the RT application in the background, cyclicttest was also run to make the system as busy as possible during the measurement. The results are summarized in Figures 3, 4 and 5. The graphs show that FIFO and RR give similar results and deadline planning policy gives the best results.

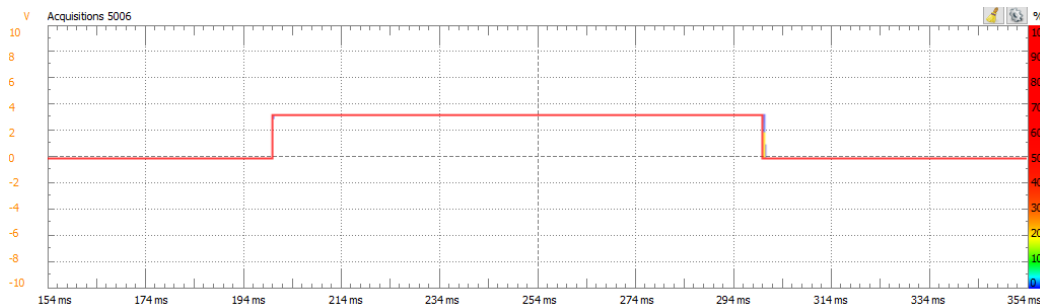




**Figure 3:** Measurement result using Analog Discovery 2 for RR planning policy



**Figure 4:** Measurement result using Analog Discovery 2 for FIFO planning policy



**Figure 5:** Measurement result using Analog Discovery 2 for deadline planning policy

## 8. CONCLUSION

The first method for measuring latency was the program `cyclicttest` and the script `mklatencyplot` to create a histogram from the measured values. This measurement compared the latency on the Linux kernel without the `PREEMPT_RT` patch. The results are summarized in Figures 1 and 2. It can be seen from the graphs that the `PREEMPT_RT` patch application reduces the maximum latency by 58  $\mu$ s. However, this can be skewed because both measurements took a relatively short time (approximately 2 hours each measurement). Another measurement looked at the impact of the choice of planning policy on latency. The RT application that controlled the output pin GPIO was used to verify the impact of the planning policy. The measurement was performed using an Analog Discovery 2 device and the WaveForms program. The measured results show that best results gives deadline planning policy.

## REFERENCES

- [1] Karim Yaghmour, Jon Masters, Gilad Ben-Yossef, Philipe Gerum, *Building Embedded Linux Systems*. 2008, Second edition, Sebastopol: O'Reilly ISBN 978-0-596-52968-0
- [2] Jan Altenberg, linutronix GmbH, Introduction to Realtime Linux [online], 8.3.2022, available from: <https://youtu.be/BKkX9WASfpl>
- [3] HOWTO build a simple a basic cyclic application [online], 8.3.2022, available from: <https://wiki.linuxfoundation.org/realtime/documentation/howto/applications/cyclic>
- [4] Cyclicttest [online], 8.3.2022, available from: <https://wiki.linuxfoundation.org/realtime/documentation/howto/tools/cyclicttest/start>
- [5] mklatencyplot [online], 8.3.2022, available from: <https://www.osadl.org/uploads/media/mklatencyplot.bash>

# A proposal of a method to detect spam from information messages

Michal Rickwood<sup>1</sup>, and V. Oujezský<sup>2</sup>

<sup>1</sup>Brno University of Technology, The Czech Republic

<sup>2</sup>Brno University of Technology, The Czech Republic

E-mail: [221568@vut.cz](mailto:221568@vut.cz), [oujezsky@vut.cz](mailto:oujezsky@vut.cz)

**Abstract**—This paper presents a spam detection algorithm that uses solely traffic flow logs in the form of Netflow messages. Internet service providers must detect spam in order for their entire subnets not to be marked as spamming stations. The algorithm was drafted based on an analysis of various datasets containing Netflow records. These datasets consist of valid e-mails, spam and common non e-mail related traffic. The algorithm uses domain name system blacklist verification as the first step of identifying a spamming station. Furthermore, theoretical models of valid clients and spammers have been laid out. In continuation of this work, the dataset will be studied to find correlation with the models. Included in the tracked parameters one can find the number of incoming and outgoing messages, timestamps amongst others.

**Keywords**—Detection, e-mail, flow, security, spam

## 1. INTRODUCTION

This paper presents a topic that is encountered by every e-mail user, which in today's society corresponds to approximately 50 % of the world's population. This topic is spam. Spam is a method of sending unwanted e-mail messages to a large number of recipients. Although the ratio of spam to legitimate mail has shown a downward trend in recent years, approximately 45 % of traffic falls into the first category [1].

Many systems have been designed and implemented for detection. Most of them require scanning the content of the messages sent. This approach appears to be relatively simple and highly effective, since artificial intelligences can identify spam mail essentially flawlessly from learned datasets. However, the law side of the equation enters the picture and the whole situation becomes significantly more complicated. Filtering by email content is an obvious and serious invasion of privacy. The problem arises for the ISPs that provide the email server. If a client uses that provider's services for spamming purposes and is flagged as a spammer in the UCEPROTECT project, other users in the subnet are affected. The provider is of course then subject to legal and mostly financial consequences. For this reason, the provider must monitor outgoing traffic from its server in a way other than by scanning the content of the messages in order to be able to detect a possible spamming station earlier.

One widely used detection method is using NetFlow messages. These records contain only metadata about the communication and the users, not the content of the communication itself. This information, which is included in the packet headers, includes such things as Internet Protocol sender and receiver addresses, port numbers, and more. An approach based on monitoring network traffic from NetFlow messages is also addressed in this paper. In order to identify and model the behavior of spamming devices, we use the already created NetFlow message dataset [2]. As a first step, the algorithm looks up the IP address in the domain name system blacklist. It then compares the metadata about the station with the previously created models of the spammer and legitimate user based on the datasets.

## 2. THE CURRENT METHODS USED

Many of today's spam detection algorithms still work by reading the content of messages, even if only the subject field, for example, and not the text itself. Providers that interfere with messages include leaders such as Google, Yahoo and Outlook [3]. They use various machine learning methods such as neural networks, k-nearest neighbor algorithm and others for detection. The methods are then learned on large sets of spam and valid emails. In addition to using already established rules, these learning-based algorithms can create their own rules. Google's spam detection algorithm has advanced to the stage where it can detect up to 99.9 % of spam. Conventional machine learning algorithms compare each message

against valid and spam data before setting spam detection rules. However, there are also design systems that have been inspired by artificial immune systems and use special features to generate detectors to cover the spam space [4].

Providers use methods that have a very high success rate, but the problem still lies in the exploitation of the content. However, there are many algorithms or method designs that detect spam without reading the messages. These methods observe and collect data about traffic on the network, which they then aggregate by source address [5]. For example, the following table I might result, where an observer has observed the number of incoming connections, outgoing connections on port 25, and the number of distinct addresses in both directions. The table I is sorted by the number of outgoing connections.

IP	dist out	out	dist in	in
1	1980	334356	36354	675381
2	3227	247588	36354	17645
3	11459	11459	36354	745408
<b>4</b>	<b>39460</b>	<b>244117</b>	<b>0</b>	<b>0</b>
5	11280	240733	153275	675632
6	3512	238665	788	27738
7	7943	195573	132616	539297
<b>8</b>	<b>2</b>	<b>184698</b>	<b>0</b>	<b>0</b>
9	2252	136847	10	187
<b>10</b>	<b>24213</b>	<b>116898</b>	<b>1</b>	<b>2</b>
11	7774	115746	8	24972
12	8376	68413	24	172464
<b>13</b>	<b>17532</b>	<b>64685</b>	<b>0</b>	<b>0</b>
14	341	57251	66237	901280
15	443	54212	10	578

**Table I:** Example of the result of traffic data aggregation by IP addresses [5].

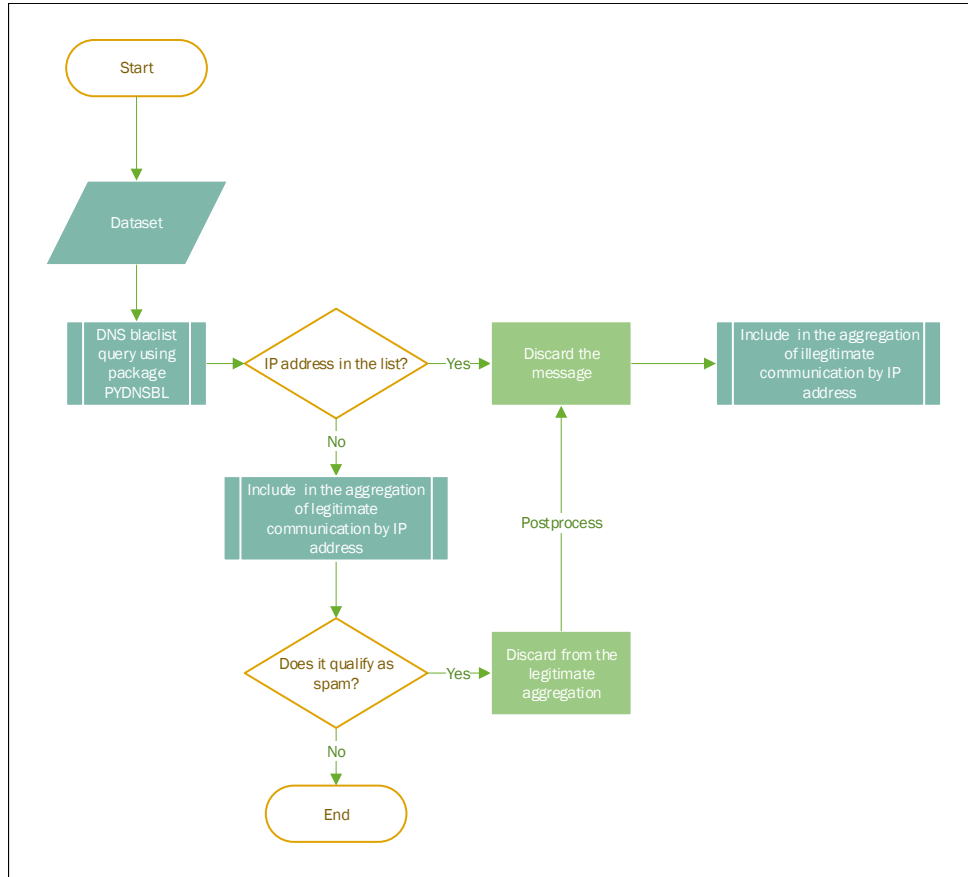
In the table I we can see that most IP (Internet Protocol) addresses that have a high number of outgoing connections also have a similar number of incoming connections. In this case, author Vliek presents the assumption that these are active but valid stations. Highlighted, on the other hand, are stations that have a high number of outgoing connections but few or no incoming connections. According to Vliek's assumption, these stations are candidates for spam clients. However, he also states that there are legitimate stations that are only used for messaging. This category includes, for example, accounts set up by banks to send account information, online newsletters that send e-mail in some cases even several times a day. However, most of these cases will have a relatively small number of different IP addresses to which they send mail.

### 3. THE PROPOSED METHOD

The development of the detection algorithm can be carried out in several different environments with different software and hardware resources. The main hardware requirement for the environment is a relatively high computational capacity. The most practical solution turned out to be the use of the cloud. One variant of this solution is from Kaggle [6].

The detection algorithm is applied to the dataset. It consists of data items of the same type, in our case a packet that was sent over the monitored network. The parameters tracked include the following – send time, protocol type, source IP address and port, destination IP address and port and other.

We use already created datasets. Extensive datasets are available captured from CTU in Prague [2]. Figure 1 shows a schematic of the proposed detection algorithm. The first step is to convert from .pcap



**Figure 1:** The basic schema of the algorithm.

format to .csv format and then retrieve it using the pandas package [7]. The next step is to query the DNS (Domain Name System) blacklists using the pydnsbl package [8]. If the IP addresses are listed, the message would be dropped by the provider. In our case, it just adds the data from the netflow message to the aggregation of illegitimate traffic. If the response from the DNS blacklist query is negative, the information from the netflow record is added to the aggregation of legitimate communication. It is then checked that the IP address from which the communication is sent still falls into the category of a legitimate station and has not crossed any of the specified thresholds to further verify that it is not a spam station. If no threshold is crossed, the process is terminated.

The main indicator of suspicious behavior is the threshold of a large amount of activity in terms of outgoing connections and on the contrary a small amount of incoming traffic. The second threshold is the traffic behaviour in time computed from the send time NetFlow value. We are looking for a specific repetitive behaviour in terms of time / repetitions. The main value of the threshold is weighted centrality measurement of the average duration. However, this information is definitely not enough. Another important indicator is the ratio of active to idle time. If all communication is sent at one time every day or periodically, for example, every 20 minutes, this station has a higher chance of being a spamming station. The algorithm itself is implemented in Python in the Kaggle web environment.

#### 4. CONCLUSION

The aim of the research was to design an algorithm that can detect end stations in the network generating “spam” traffic from the traffic log in the form of Netflow messages. The primary reason for filtering using logs without interfering with the messages themselves is to protect privacy and personal information.

As part of the design, an analysis of the current spam filtering solution with and without interference to the content of the communication was first performed. The algorithm in its current version retrieves the underlying data, can check the IP address against a list and produce basic aggregated traffic data from the perspective of individual stations. Due to time constraints, not all IP addresses in the dataset used

have been scanned yet, but only the first 10,000 records. Of those addresses, 27,66 % were listed on at least one list. On average, then, an address was listed on 1,9378 lists in the case of a positive response.

Several IP addresses with suspicious behavior were found in the statistics. The main indicator of such suspicious behavior was a large amount of activity in terms of outgoing connections and on the contrary a small amount of incoming traffic. For these stations, there was also a very low number of stations being sent to. At the moment, it is not possible to firmly establish whether these are end stations generating spam traffic. It is possible that these are automated informational e-mail boxes.

Based on the analyses, parameters will be determined which will be monitored in terms of IP addresses and by which it will be decided whether the station is a legitimate user or a spammer. At this stage, any reasonable statistics of the proposed algorithm are not available. Currently, two parameters are used that still need to be tuned. Depending on the accuracy, additional parameters can be added. One of the approaches that will be developed in the continuation of this work is traffic monitoring according to the author Vliek.

## REFERENCES

- [1] J. Johnson, "Global spam volume as percentage of total e-mail traffic from january 2014 to march 2021, by month." [Online]. Available: <https://www.statista.com/statistics/420391/spam-email-traffic-share/>
- [2] S. Garcia, M. Grill, J. Stiborek, and A. Zunino, "An empirical comparison of botnet detection methods." [Online]. Available: <http://dx.doi.org/10.1016/j.cose.2014.05.011>
- [3] E. G. Dada, J. S. Bassi, H. Chiroma, A. O. Adetunmbi, O. E. Ajibuwa *et al.*, "Machine learning for email spam filtering: review, approaches and open research problems," *Heliyon*, vol. 5, no. 6, p. e01802, 2019.
- [4] I. Idris and A. Selamat, "Improved email spam detection model with negative selection algorithm and particle swarm optimization," *Applied Soft Computing*, vol. 22, pp. 11–27, 2014.
- [5] G. Vliek, "Detecting spam machines, a netflow-data based approach," University of Twente, The Netherlands, 2009.
- [6] Kaggle. [Online]. Available: <https://www.kaggle.com/>
- [7] T. pandas development team, "pandas-dev/pandas: Pandas," Feb. 2020. [Online]. Available: <https://doi.org/10.5281/zenodo.3509134>
- [8] dmippolitov, "Pydnsbl," Sep. 2017. [Online]. Available: <https://github.com/dmippolitov/pydnsbl/>

# COMMUNICATION NETWORK FOR A SWARM OF UNMANNED AERIAL VEHICLES

<sup>1</sup>Jan Rajm, <sup>2</sup>Jiří Janoušek

<sup>1</sup>Bachelor Degree Programme (3.), FEEC, Brno University of Technology, Czech Republic

<sup>2</sup>DTEE, FEEC, Brno University of Technology, Czech Republic

E-mail: [xrajmj00@vutbr.cz](mailto:xrajmj00@vutbr.cz), [xjanou09@vutbr.cz](mailto:xjanou09@vutbr.cz)

**Abstract**—This paper deals with the creation of a wireless network for a swarm of unmanned aerial vehicles which enables high enough data throughput for transition of telemetric and application data between unmanned aerial vehicles and ground control station and also enables communication at a greater distance. Even when unmanned aerial vehicles are beyond visual line of sight.

**Keywords**—4G, 5G, cellular network, LTE, swarm of UAVs, UAV

## 1. INTRODUCTION

Unmanned aerial vehicles (UAVs), also called drones, have a wide area of use-cases because of their high mobility. Examples of use-cases may be filming and taking photos, transportation of small packages, military purposes, precision agriculture, surveillance, and many others. Since UAVs serve as instruments for remote tasks or as devices for data collection, a reliable communication network is essential. Only with it is it possible to control UAVs by a ground control station (GCS), where all important decisions are made. UAVs serve as peripheral devices of the GCS.

One UAV can handle a large number of tasks, adding another UAV and creating a swarm can expand the number of tasks or save time. In a single swarm can be many types of UAVs, each type with its special abilities. On the other hand, it is more challenging not only for a GCS to manage the swarm, but also demands higher robustness of the communication network.

Since ad-hoc networks are usually limited to the line of sight (LOS) communication between the GCS and its subordinated swarm of UAVs, cellular network technology is proposed as a solution for the beyond visual line of sight (BVLOS) communication. In this paper, a communication network for a swarm of UAVs, enabling BVLOS communication, is designed.

## 2. SELECTION OF STANDARDS

The main block of communication networks are communication modules. They are made according to a specific standard, which defines its features, such as the used frequency band, modulation, type of forward error correction, cryptography and so on. Other important aspects of communication standards are the network technology, which can be created using modules of the standard, and the topology of the network. What we are searching for are wireless communication standards with data rate enabling transmission of telemetry data and application data (data from sensors, video, etc.) in real time, communication range higher than 100 meters, the ability to communicate BVLOS and the ability to maintain a swarm containing hundreds of UAVs. Maximum allowable latency depends on the specific use-case and may range from the tens to the hundreds of milliseconds [1].

The most suitable possibility is LTE, which is the building block of 4G cellular networks. This means an advantage of already created network infrastructure, so we do not need to build our own infrastructure. In addition, there is an advantage of using commercial frequency bands, in which there is less interference. Just the subject with the permission may use these frequencies. In our case the permission belongs to network providers. The other advantage of cellular network is its coverage of an area by its signal enabling BVLOS communication between a GCS and a swarm of UAVs. Data rate for both downlink and uplink of 4G LTE network may be in orders of hundreds of Mb/s and theoretically can be data rate for downlink up to 1Gb/s and uplink 500 Mb/s. That is the opposite of our requirement.



For UAVs it would be better to have a higher data rate for uplink (because of sending application data back to the GCS), than downlink (the only data the UAV receive are control data). Data rate closely depends on the environment. Higher levels of noise, obstacles and many other phenomena will decrease data rate. Nevertheless, data rate of hundreds (even tens) of Mb/s for uplink is still high enough to stream a video and transmit telemetric data. The only limitations of using 4G networks are the maximal height and the fee for using a provider's network. Since a base transceiver station (BTS) is supposed to deliver a terrestrial signal, the signal level above BTS is minimal. The latency of 4G networks ranges between 30 to 80 milliseconds, which is good enough for a most of use-cases. Telemetry requires data rate ranging 60-100 kb/s for both uplink and downlink [1][2].

Use-cases closely depend on network properties and the technology of the network. The use of 4G networks is more suitable for urban areas than for rural areas. It is also worth mentioning that the 4G network allows the swarm of UAVs to be spread over a large area. In general, the use-cases, for which our network is most suitable, require both large range and high data rate, for example delivery, surveillance, infrastructure control, first aid, and so on.

Other considered cellular standards for this paper included LTE Cat-M, NB-IoT and LoRaWAN. For ad-hoc network technology are suitable standards ZigBee (802.15.4), HaLow (802.11ah), RFD868, FrSky and many others. For satellite network technology is suitable for example SPL Satellite Telemetry.

### 3. SOFTWARE FOR UAV COMMUNICATION

Mission Planner is a software for a GCS, enabling planning a mission for one and more UAVs, enabling configuration of flight controller settings for UAVs, including control of mission commands. Mission Planner also provides an overview of telemetry data of every single UAV of our swarm and can run its own system in the loop (SITL) simulation as shown in a Figure 1. Mission Planner is an open-source software [3].



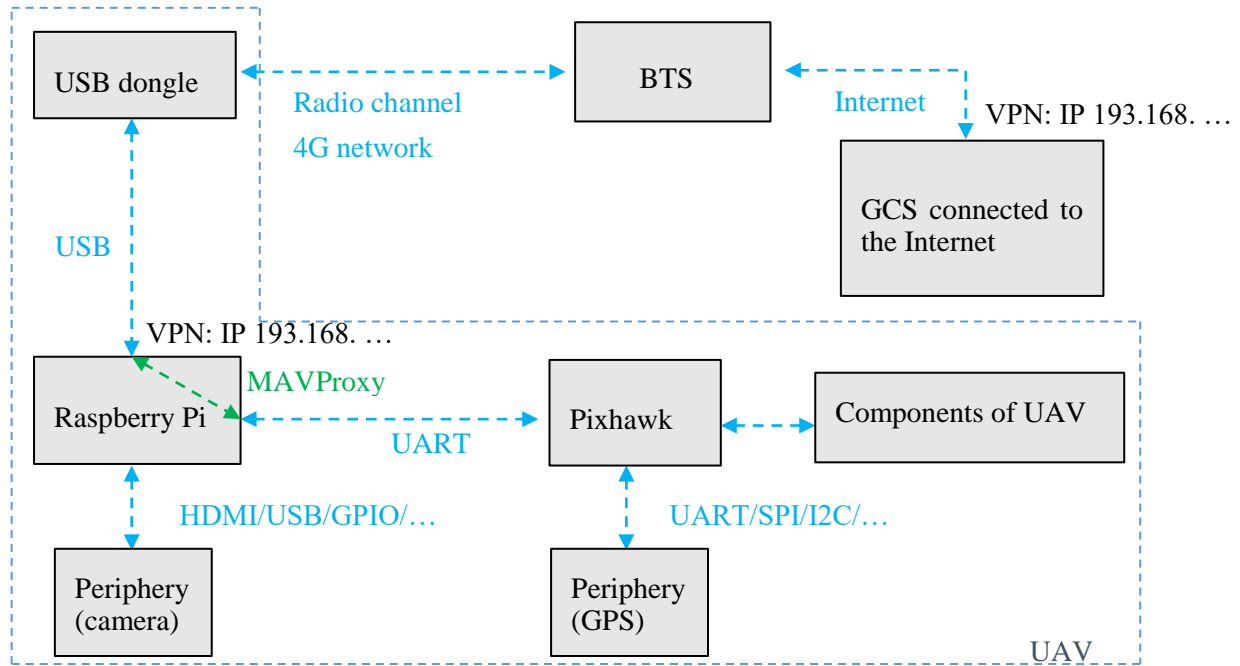
**Figure 1:** Mission Planner running SITL simulation

Mission Planner communicates with UAVs using a MAVLink protocol, which is an open-source protocol used not only by UAVs, but also by other unmanned vehicles, such as boats, rovers, submarines, and others. MAVLink provides methods for detection of lost packets, damaged packets, and authentication. MAVLink is optimized for hardware with limited RAM and flash memory, such as microcontrollers [4].

ArduPilot is an open-source autopilot software used in flight controllers, which is the brain of the UAV. ArduPilot receives and processes tasks transmitted by MAVLink from the GCS. ArduPilot provides advanced functionality including a real-time communication with the GCS [5].

#### 4. COMMUNICATION NETWORK

Proposed communication network consists of a swarm of UAVs, which is connected with a GCS over the Internet. The topology of this network is the star topology since every UAV communicates with the GCS. The wireless connection between the GCS and the swarm is implemented by a 4G network. Current 5G networks operate in non-standalone (NSA) mode, so they still use 4G cores in their architecture, which limits their maximal data rate, so it is still better to use 4G networks because of their coverage of area and availability of modules. The time of use of 5G networks is yet to come, but when it comes, an upgrade means simply just replacing the 4G module with a 5G one. The software infrastructure is the same. 5G will offer low latency (in order of ones of milliseconds for end-to-end connection), data rate of Gb/s, high reliability and some useful features for UAVs (e. g. remote identification of UAVs) since 3GPP (The 3rd Generation Partnership Project) includes UAVs among other use-cases of 5G networks. Figure 2 shows a block diagram of a single UAV communication network [6][7].



**Figure 2:** Block diagram of the network

In order to connect the flight controller of the UAV with the 4G network, a companion computer is needed. The companion computer is the Raspberry Pi 4 and for 4G signal reception will be used a USB 4G LTE dongle. Within the Internet, the network will be implemented by the virtual private network (VPN). Within the VPN, the GCS and companion computers have their unique IP, which is used when creating connections. The companion computer is connected with the flight controller via the UART interface. Companion computer also enables a peripheral connection of many other peripheral devices, which flight controllers cannot handle, such as a camera or other sensors producing large amount of data. The companion computer can also perform on-board tasks, such as signal processing or decision-making processes. Decision-making allows transmission of a lower amount of data and thus the GCS can handle more UAVs in the swarm.

Within the companion computer it is necessary to forward telemetric data between the serial port (in our case the UART port), connected to the flight controller, and the remote GCS, which is represented by its VPN IP address. For this purpose, an open-source software MAVProxy is used. MAVProxy also can emulate an UAV by its own SITL, which can be used during tests of communication infrastructure and offers the possibility of fully controlled tests [8].

#### 5. CONCLUSION

This paper aimed to create a communication network for a swarm of UAVs, which enables BVLOS communication between UAVs and a GCS, and also provides sufficient data rate for not only telemetry

data, but also for application data. The 4G network showed as the most suitable solution for our UAV communication network because of its coverage of an area, high data rate and sufficiently low latency. In combination with a companion computer, our network is very flexible and can be used to control a numerous swarm of UAVs in use-cases such as delivery, surveillance, infrastructure control, first aid, and others. 4G network provides a well secured and reliable network for our swarm.

Another motivation for the use of the cellular network is its perspective future. The use of cellular networks by UAVs is counted by 3GPP. So, UAVs are included among other use-cases of 5G networks during ongoing standardization and 5G networks will provide some useful features for UAVs. 5G networks will also provide a great enhancement in data rate, latency, reliability, connection density and much more.

It is possible to build on this work by many enhancements. The designed network is suitable for tasks requiring high data rate and computing power. An example of an enhancement can be UAVs with their own subordinated UAVs or UAVs serving as a mobile access point.

## REFERENCES

- [1] Y. Zeng, Q. Wu, and R. Zhang, “Accessing From the Sky: A Tutorial on UAV Communications for 5G and Beyond,” *Proceedings of the IEEE*, vol. 107, no. 12, pp. 2327–2375, Dec. 2019. [Online]. Available: doi: 10.1109/jproc.2019.2952892. [Accessed 25-Mar-2022]
- [2] C. (C.I.) Cox, *An introduction to LTE: LTE, LTE-advanced, SAE, and 4G mobile communications*. Hoboken: Wiley, 2012. [Accessed 25-Mar-2022]
- [3] ardupilot.org. “Mission Planner Home — Mission Planner documentation”. [Online]. Available: <https://ardupilot.org/planner/>. [Accessed 25-Mar-2022]
- [4] Mavlink.io. “Introduction · MAVLink Developer Guide,”, 2009. [Online]. Available: <https://mavlink.io/en/>. [Accessed 25-Mar-2022]
- [5] ardupilot.org. “ArduPilot Documentation — ArduPilot documentation,”. [Online]. Available: <https://ardupilot.org/ardupilot/>. [Accessed 25-Mar-2022]
- [6] W. Xiang, K. Zheng, and X. (S.) Shen, *5G Mobile Communications*. Cham: Springer International Publishing, 2016. [Accessed 25-Mar-2022]
- [7] 3GPP. “UAS - UAV,”. 2019. [Online]. Available: <https://www.3gpp.org/uas-uav> [Accessed 25-Mar-2022].
- [8] ardupilot.org. “MAVProxy — MAVProxy documentation,”. [Online]. Available: <https://ardupilot.org/mavproxy/>. [Accessed 25-Mar-2022]

# Secure Two-Party Computation for weak Boneh-Boyen Signature

M. Sečkář<sup>1</sup>, and S. Ricci<sup>1</sup>

<sup>1</sup>Department of Telecommunications, Brno University of Technology, Technická 12, Brno, 616 00, Czech Republic

E-mail: [xsecka04@vut.cz](mailto:xsecka04@vut.cz), [ricci@vut.cz](mailto:ricci@vut.cz)

**Abstract**—Secure two-party computation allows two entities to securely calculate a common result keeping their private inputs secret. By applying this to the weak Boneh-Boyen signature, a trusted third party is able to sign the user’s message (or a secret key) without knowing its content (or value). In this article, we present a C library that implements a two-party computation algorithm for generating a user’s secret key that can be used in a group signature scheme. The library provides a structured output ready to be serialized and sent over a network. We also show the computational benchmarks of the implemented algorithms. The computations on the sender’s side are relatively fast, which broadens the possibilities of deployment on constrained devices.

**Keywords**— Secure Multi-Party Computation, Paillier Cryptosystem, Group Signature, C Language, OpenSSL, GMP

## 1. INTRODUCTION

A digital signature is an algorithm used to verify the integrity of the message and the authenticity of the sender. In addition to these traditional requirements, European Union (EU) regulations and United States (U.S.) strategic plans request new privacy protection features. One way to cover the new demands is to pass to group signatures. For example, an employee (i.e., sender) within a large company needs to sign a message on the behalves of the company without revealing its identity to the external receiver. However, the receiver needs to verify that the sender is an employee of the company and re-identify it in case of malicious behavior. Therefore, we would need that the sender can sign while remaining anonymous (Anonymity) and that the receiver in collaboration with a third trust party can identify malicious senders (Traceability). The signature proposed by Hajny et al. [1] is developed to be efficient on constrained devices and can be turned into a group signature by applying a secure multi-party computation algorithm on the key generation as proposed by Ricci et al. [2].

To the best of our knowledge, we present the first implementation of the secure multi-party computation scheme proposed by Belenkiy et al. [3]. Our implementation has been developed in C language using GNU Multiple Precision Arithmetic Library (GMP) and partly OpenSSL Library. We also show the computational benchmarks of the implemented algorithms.

## 2. IMPLEMENTATION

In this section, we present our proof-of-concept implementation that is focusing on the secure multi-party computation scheme [3], leaving the proof of knowledge for the second stage of development. The library is located in a public GitHub repository <sup>1</sup>. We also discuss how the parameters should be chosen. The key generation algorithm is run by two entities, Signature Group Manager (SGM) and Senders (Ss). The SGM and each S need to agree on a key that will be used for signing. Belenkiy et al. scheme [3] allows S and SGM to jointly compute the value  $\sigma = g^{1/(sk_i+sk_m)}$  of the sender’s secret key  $sk_i$  and the manager’s secret key  $sk_m$  without revealing both values. The scheme can be seen as a modified Paillier cryptosystem that provides proof of knowledge of both secrets  $sk_i$  and  $sk_m$ . Then  $\sigma$  can be used to generate a weak Boneh-Boyen (wBB) signature on a given message. The implementation can be split into two phases: Setup\_SGM algorithm and Join algorithm that are depicted in Algorithms 1 and 2, respectively.

The implementation requires fast computations on large integers. For that reason, C language and specifically GMP library are chosen as a basis for the implementation. The outputs of Setup\_SGM and

<sup>1</sup>[https://github.com/xsecka04/Paillier\\_NIZKPK](https://github.com/xsecka04/Paillier_NIZKPK)

Join algorithms are implemented as structures that can be easily serialized and sent over the network. During the Setup\_SGM phase, the generation of RSA parameters occurs. For this reason generate\_r\_from\_group, generate\_r\_from\_bitlength and generate\_RSA\_SSL were developed. The RSA parameters are generated using OpenSSL Library, in particular using BN\_generate\_prime\_ex2 function. Linux's /dev/urandom source of pseudo-random numbers is used as a random seed for mpz\_urandomm function that is located in generate\_r\_from\_bitlength and generate\_r\_from\_group functions. These functions generate a random number in a given cyclic integer group. Once the RSA parameters are generated, they are encoded in a hexadecimal string from the OpenSSL's BIGNUM type and initialized as mpz\_t integer types to be further used in the GMP functions.

---

**Algorithm 1** Setup\_SGM( $\kappa$ )

---

- 1: Consider  $q_{EC}$  a prime of the right order
  - 2: Generate an RSA-modulus  $\mathbf{n}$  of size at least  $|2^{3\kappa} q_{EC}^2|$ , where  $\mathbf{n} = pq$ , and  $\phi(\mathbf{n}) = (p-1)(q-1)$
  - 3: Consider  $\mathbf{h} = \mathbf{n} + 1 \in \mathbb{Z}_{\mathbf{n}^2}$
  - 4: Generate  $\mathbf{g}$  of order  $\phi(\mathbf{n})$  in  $\mathbb{Z}_{\mathbf{n}^2}$
  - 5: Generate another RSA-modulus  $\mathbf{n}$ , where  $p_g, q_g$  are big primes,  $\phi(\mathbf{n}) = (p_g - 1)(q_g - 1)$
  - 6: Consider  $\mathbf{h} \leftarrow \mathbb{Z}_{\mathbf{n}}$  and  $\mathbf{g} \leftarrow \langle \mathbf{h} \rangle$
  - 7: **return**  $par = (\mathbf{h}, \mathbf{n}, \mathbf{g}, \mathbf{h}, \mathbf{n}, \mathbf{g}, q_{EC})$
- 

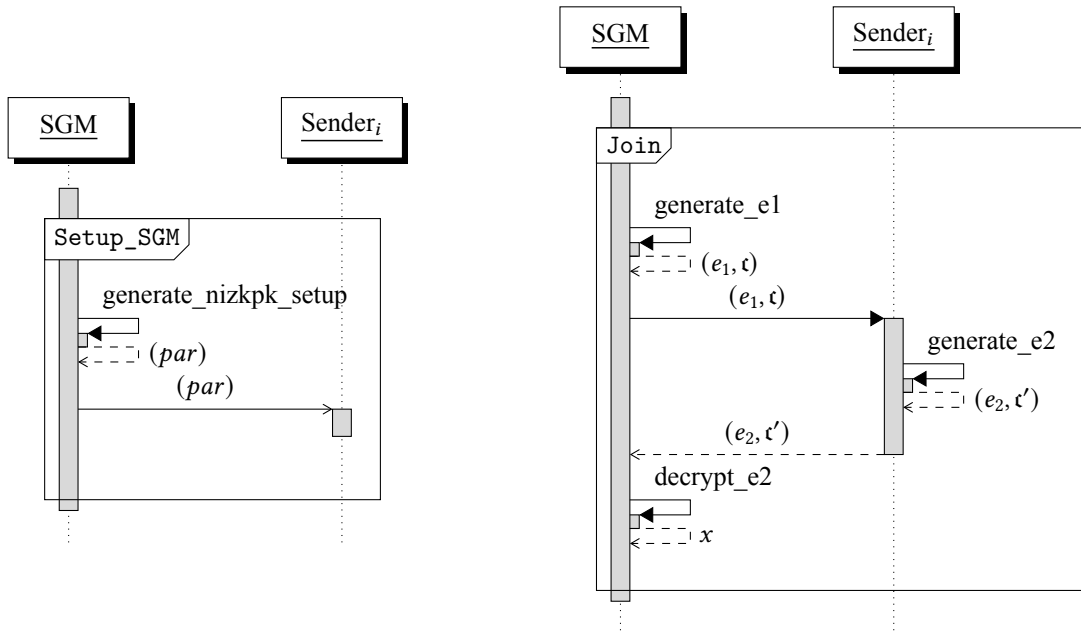
---

**Algorithm 2** Join( $par$ )

---

- 1: SGM computes:
  - 2:  $r \leftarrow \mathbb{Z}_{\phi(\mathbf{n})}, r' \leftarrow \mathbb{Z}_{\phi(\mathbf{n})}$
  - 3:  $e_1 = \mathbf{h}^{\mathbf{n}/2 + sk_m} \mathbf{g}^r \bmod \mathbf{n}^2$ ,
  - 4:  $\mathbf{t} = \mathbf{g}^{sk_m} \mathbf{h}^{r'} \bmod \mathbf{n}$ ,
  - 5: Sender<sub>*i*</sub> computes:
  - 6:  $sk_i, r_1 \leftarrow \mathbb{Z}_{q_{EC}}, r_2 \leftarrow \{0 \dots |2^\kappa q_{EC}|\}$ ,  
 $\bar{r} \leftarrow \{0 \dots |2^\kappa \mathbf{n}|\}$
  - 7:  $e_2 = (e_1 / \mathbf{h}^{\mathbf{n}/2})^{r_1} \mathbf{h}^{\mathbf{n}/2 + sk_i r_1 + r_2 q_{EC}} \mathbf{g}^{\bar{r}} \bmod \mathbf{n}^2$
  - 8:  $\mathbf{t}' = \mathbf{g}^{sk_i} \mathbf{h}^{\bar{r}} \bmod \mathbf{n}$ ,
  - 9: SGM computes:
  - 10:  $x = Dec(e_2) - \mathbf{n}/2$
- 

GMP contains the multiple precision integer mpz\_t data type used for the storage and computations. Multiple precision (or Arbitrary precision) means that the size can grow dynamically and is not restricted as other generic fixed precision data types. For the variable to be properly used, it needs not only to be declared (mpz\_t variable), but also initialized (mpz\_init(variable)) and cleared after use (mpz\_clear(variable)). GMP library provides several integer operation functions such as arithmetic (e.g., mpz\_add()), exponentiation (e.g., mpz\_powm()), modular operations (e.g., mpz\_invert()), and pseudo-random number generators (e.g., mpz\_urandomm()).



**Figure 1:** From left to right, the sequence diagrams of Setup\_SGM and Join algorithms

Figure 1 depicts the sequence diagrams of Setup\_SGM and Join algorithms, respectively. The parameters

sent between SGM and  $\text{Sender}_i$  are serialized into JSON objects as hexadecimal strings using `JSON_serialize_Setup_par`, `JSON_serialize_e1` and `JSON_serialize_e2` functions. On the other end, the objects are being deserialized using their respective deserialization functions, i.e. `JSON_deserialize_Setup_par`. Note that the use of these functions is optional and the end user can choose different formats for the serialization.

Algorithm 2 requires the homomorphic computation of  $\sigma$ . As suggested by Belenkiy et al. [3], we consider the Paillier cryptosystem. In particular, in order to securely compute  $\sigma$ ,  $\mathbf{h}$  has to be taken equal to  $\mathbf{n} + 1$  and  $\lambda = \phi(\mathbf{n})$ . The decryption needs also to be adjusted accordingly. These modifications are directly suggested by Paillier [4].

## 2.1. Parameters Setup

The scheme requires two RSA groups, one for the homomorphic calculations and the other for the commitment parameters. Parameters  $\mathbf{h}, \mathbf{n}, \mathbf{g}$  can be utilized in the non-interactive zero-knowledge proofs of knowledge, alongside the challenges  $e_1, e_2$  and commitments  $\mathbf{c}$  and  $\mathbf{c}'$ . We refer to [3] for more details.

Algorithm 1 shows how the parameters should be chosen. There are several steps needed to be specified in order to implement the depicted `Setup_SGM` scheme. Note that the RSA-modulus  $\mathbf{n}$  has bitlength of at least  $|2^{3\kappa} q_{EC}^2|$ . Following the NIST security standard, we consider  $q_{EC}$  a prime of 256 bitlength and a security parameter  $\kappa$  equal to 1350. This brings the RSA-modulus  $\mathbf{n}$  to be large enough to make the encryption and decryption secure and follow the NIST standards.

In Step 4 of Algorithm 1, we need to generate  $\mathbf{g}$  of order  $\phi(\mathbf{n})$  in  $\mathbb{Z}_{\mathbf{n}^2}$ . Finding an element of order  $\phi(\mathbf{n})$  in a big composite group is not an easy task. The traditional approach for finding a generator on  $\mathbb{Z}_{\mathbf{n}^2}$  requires the selection of a random number  $h$  in the group and the check if  $h$  is of the right order. It is important to notice that a composite group has not guaranteed to possess generators and that the search requires exponentiations in a big modulus that are computationally demanding. We solved this issue by considering how Paillier cryptosystem [4] generates the noise in the encryption. Note that the implemented algorithm is a variant of Paillier scheme. Therefore, we generated a random number  $k \xleftarrow{\$} \mathbb{Z}_{\mathbf{n}^2}$  and we computed  $\mathbf{g} \leftarrow k^{\mathbf{n}} \bmod \mathbf{n}^2$ . In this way,  $\mathbf{g}^{\phi(\mathbf{n})} \equiv 1 \bmod \mathbf{n}^2$ , i.e., has the right order for the computations.

In case of the second RSA-modulus (Step 5), we also considered  $\mathbf{n}$  of size at least  $|2^{3\kappa} q_{EC}^2|$  and we generate  $\mathbf{h} \xleftarrow{\$} \mathbb{Z}_{\mathbf{n}}$ . Moreover,  $\mathbf{g}$  has to be taken from the subgroup generated by  $\mathbf{h}$  (please see Step 6 in Algorithm 1) and, therefore, we generated a random  $t \xleftarrow{\$} \mathbb{Z}_{\phi(\mathbf{n})}$  and we computed  $\mathbf{g} = \mathbf{h}^t$ . Once the parameters are set, Algorithm 2 can be applied to compute  $\sigma = g^{1/(sk_i + sk_m)}$  without revealing  $sk_i$  to SGM and  $sk_m$  to S. If the assertion  $x \equiv ((sk_m + sk_i)r_1 \bmod \mathbf{n}) \bmod q_{EC}$  holds, the exchange has been successful and the parties are ready for a generation of the signature and verification using the wBB signature. Moreover,  $\mathbf{n}/2$  calculation is implemented as  $\lfloor \mathbf{n}/2 \rfloor$  and  $(e_1/\mathbf{h}^{\mathbf{n}/2})$  as  $e_1(\mathbf{h}^{\mathbf{n}/2})^{-1} \bmod \mathbf{n}^2$ , where  $x^{-1}$  denotes a modular multiplicative inverse and  $\lfloor \dots \rfloor$  denotes the floor function. The former change was made due to the fact that 2 has no multiplicative inverse in  $\phi(\mathbf{n}^2)$  and, therefore,  $\mathbf{n}/2 \bmod \phi(\mathbf{n}^2)$  cannot be computed. Table I shows the chosen set of parameters and their respective sizes. Our choices follow the NIST security standard.

**Table I:** Input parameters for `Setup_SGM` algorithm

Parameter	Size [Bits]	Value
$\kappa$ (Security Parameter)	32	1350
$q_{EC}$ (Order of the Elliptic curve)	256	0x2523648240000001ba344d8000 000007ff9f800000000010a100000 00000000d <sup>1</sup>
$\mathbf{n}, \mathbf{n}$ (RSA Moduli)	4572	random
$\mathbf{g}$	9139	random
$\mathbf{h}$	4572	$\mathbf{n} + 1$
$\mathbf{h}, \mathbf{g}$	4572	random

<sup>1</sup>The value is given in the hexadecimal representation.



### 3. EXPERIMENTAL RESULTS

The tests were conducted on two devices: AMD Ryzen 9 5900X CPU with 64GB RAM in a Docker environment built on Debian GNU/Linux 10 (buster) with 5.10.16.3-microsoft-standard-WSL2 kernel and on a Raspberry Pi 4 Model B - 4GB RAM built on Raspberry Pi OS. For the purpose of testing the computational time, `time.h` from the C standard library was utilized as well as a simple script looping over the computations and storing them into arrays of structures. Table II depicts elapsed CPU time from the tests with input parameters shown in Table I, which also depicts the bit-lengths of the RSA moduli. The depicted results in Table II were calculated as a mean over 10 runs. Note that the network overhead is not accounted for in these calculations.

**Table II:** Performance benchmarks

Environment	Number of senders	Elapsed Time of Setup_SGM [ms]	Elapsed Time of Join (SGM) [ms]	Elapsed Time of Join (Sender) [ms]
<b>Docker</b>	1	3330.96	211.13	194.37
	50	2810.68	10341.12	9662.89
	100	3128.50	20655.89	19286.74
<b>Raspberry Pi</b>	1	24489.47	2915.55	2721.11
	50	20472.66	145558.53	135836.86
	100	20974.48	290718.96	271411.49

As shown in Table II, the most time-consuming computations are done on the SGM side and the computing requirements on the sender's side are relatively negligible. The table shows the time complexity independence of Setup\_SGM algorithm and the linear dependence of Join algorithm on the number of senders. If we consider the elapsed time of Setup\_SGM algorithm, the variation between 1 and 100 senders is due to the fact that some of the parameter generators do not take a constant time to compute. Moreover, the serialization of the parameters accounts for around 6 ms in each serialization/deserialization pair. This provides the possibility of deploying the algorithm on microprocessors for a wider array of use cases. Due to the fact that Join algorithm has to be run only once, 2.7 seconds can still be considered feasible on constrained devices.

### 4. CONCLUSION

To the best of our knowledge, we present the first implementation of the secure two-party computation scheme proposed by Belenkiy et al. [3]. To do so, we developed a library that provides a structured output ready to be serialized and sent over a network. Moreover, we presented the performance benchmarks without networking overhead and showed that the computing beared on the sender is minimal.

As future work, we will continue with the design of the PKs and their implementation. Moreover, we will show how the verification phase can be accomplished on elliptic curves using bilinear pairings as well as modeling the network communication.

### REFERENCES

- [1] J. Hajny, P. Dzurenda, L. Malina, and S. Ricci, "Anonymous Data Collection Scheme from Short Group Signatures", *Proceedings of the 15th International Joint Conference on e-Business and Telecommunications*, pp. 200-209, Jul. 2018.
- [2] S. Ricci, P. Dzurenda, J. Hajný, and L. Malina, "Privacy-Enhancing Group Signcryption Scheme", *IEEE Access*, vol. 9, no. 10, pp. 136529-136551, 2021.
- [3] M. Belenkiy, J. Camenisch, M. Chase, M. Kohlweiss, A. Lysyanskaya, and H. Shacham, "Randomizable Proofs and Delegatable Anonymous Credentials", in *Advances in Cryptology - CRYPTO 2009*, vol. 5677, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 108-125.
- [4] P. Paillier, "Public-Key Cryptosystems Based on Composite Degree Residuosity Classes", in *Advances in Cryptology - EUROCRYPT-99*, 1999, vol. 1592, pp. 223-238.

# Measuring Station Firmware for LoRaWAN Wireless Data Collecting

V. Janos<sup>1</sup>, R. Juran<sup>1</sup>

<sup>1</sup>Brno University of Technology, Faculty of Electrical Engineering and Communication,  
Department of Telecommunications, Technická 12, 61600 Brno, Czech Republic

E-mail: [xjanos18@vut.cz](mailto:xjanos18@vut.cz), [xjuran15@vut.cz](mailto:xjuran15@vut.cz)

**Abstract**—This paper describes the software side of the outdoor carbon dioxide flux low-power measuring station prototype. In the following lines, the selected chamber measuring method is described as well as the IoT protocol LoRaWAN, which is used in this project. The principle of measuring software and following data processing is also introduced. Controlled parts are thermometer, CO<sub>2</sub> sensor, and chamber fan. The device is intended to be used in places that have natural sources of carbon dioxide. Measured data are sent through the LoRaWAN network for visualization and further evaluation.

**Keywords**—IoT, LoRaWAN, STM32WL, low power, carbon dioxide

## 1. INTRODUCTION

This paper presents a validation and innovation of development of a wireless field measurement station for natural carbon dioxide monitoring, in particular focused on an optimized system-on-a-chip (SoC) STM32WL platform [9], with focus on its software side to evaluate the performance of this SoC as a possible upgrade to an already existing system presented by authors of [1]. The target application location requires a technology, which allows connectivity over large distances of kilometers while being energy-efficient due to battery power. To meet these requirements, it is appropriate to use a LPWAN solution as this is a typical IoT use-case, of which the LoRaWAN protocol is used in this work, mainly to remain comparable with solution used originally by [1]. The chamber method is chosen for carbon dioxide measuring [1, 8], as the chamber is insulated from the surrounding influences and thus simulates ideal conditions throughout the measurement cycle. Data collection takes place in a fixed interval. Once the measurement is complete, the CO<sub>2</sub> flow is calculated, and the data is sent to the application server, where it can be stored and evaluated in human-readable plots. The prototype is based on the project of authors of [1], yet instead of using discrete MCU and radio module this application focuses on validating and optimizing the design by utilizing STM32WL SoC using the STM development kit NUCLEO-WL55JC, which natively supports the LoRa modulation essential to communicate using the LoRaWAN, and also offers several low-power modes for energy saving. Temperature sensor DS18B20 and MH-Z16 non-dispersive infrared (NDIR) carbon dioxide sensor were both chosen with focus on performance-price ratio and as they are intended for use in the older version of the system [1], as well as the PWM controlled fan is present to ventilate the chamber after each measurement.

## 2. CHAMBER METHOD

The advantages of this solution are the low cost along with ease of deployment and use. The chamber is made of impermeable material, e.g. PVC, and has a volume of liters to tens of liters. To measure the carbon dioxide from the soil, it is advisable to cut out all the flora in the chamber area, as the chamber is shaded and photosynthesis does not take place in the chamber and the CO<sub>2</sub> produced by the plants would distort the results.

$$F_{CO_2} = \frac{PV}{RTA} \frac{dC}{dt} \quad (1)$$

where  $F_{CO_2}$  is the amount of CO<sub>2</sub> per unit area per time s, R is the molar gas concentration [ $\mu\text{mol m}^3 \text{s}^{-2}$ ], P is the atmospheric pressure [Pa], A is the soil area covered by the chamber [ $\text{m}^2$ ], V is the volume of the chamber [ $\text{m}^3$ ], T is the air temperature [K] and  $\frac{dC}{dt}$  is the change in concentration [ $\mu\text{mol/mol/s}$ ] in the chamber over time [1]. The measuring station is placed in the chamber and collects samples for a defined period of time. From these, the parameters are calculated for input to the equation.

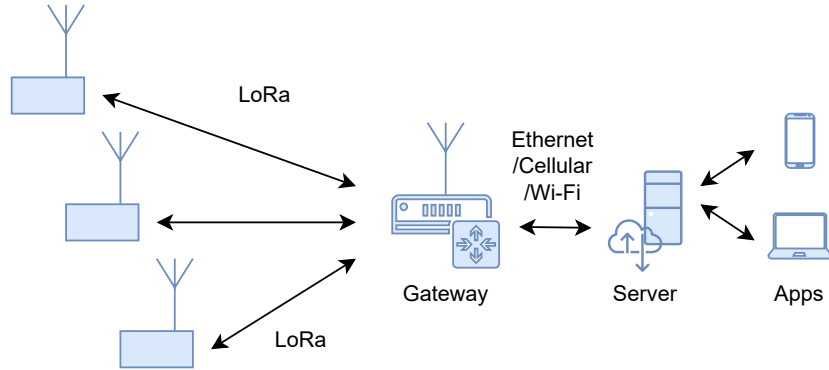
Reference	Use case	Power source
[2]	Outdoor flux measurement	Lead acid battery
[3]	Outdoor concentration measurement	Lithium battery
[4]	Indoor concentration measurement	Grid voltage

**Table I:** Comparison of some devices similar to the project, intended to measure carbon dioxide

Although there are commercially available solutions (for some examples see Table I), presented solution aims on cost reduction and custom design according to specific needs of the target project [1]. Compared to these devices for flux measurement introduced system offers primarily lower cost, also there is no need to pay monthly fees for SIM card, since LoRaWAN has possibility to deploy a private network. On the other hand due to lower cost of device it's purposed to monitor long time trends instead of precise measuring. Another motivation related to low cost is easier replacement in case of theft or damage by third party.

### 3. LORAWAN

LoRaWAN is an LPWA network protocol developed for IoT applications. It supports bidirectional half-duplex communication, end-to-end encryption, and localization. The number of messages per day is unlimited. It also offers the possibility to create a private network. When talking about LoRaWAN, it is also important to mention LoRa modulation, which forms the physical layer for LoRaWAN. Using LoRa itself, it is possible to establish peer-to-peer connections. LoRaWAN extends this and allows multi-point connections to be established, which is particularly applicable to gateways that can serve multiple end devices, as depicted in the architecture diagram in the Figure 2.



**Figure 1:** Architecture of LoRaWAN network

LoRa modulation encodes messages into radio waves using chirp pulses. A chirp is an analog signal that either increases or decreases in frequency over time. Accordingly, there are two basic types of up-chirp and down-chirp. The LoRa framework is composed as follows. First, a preamble of 8 up-chirps is transmitted, followed by two synchronizing down-chirp symbols. Then the data itself and an optional checksum are transmitted.

For further calculations, it is necessary to define several parameters describing the transmitted message. The first is the Spreading Factor, which indicates the number of bits carried by each symbol. In the ISM 868MHz band it takes values 7-12. At a greater distance from the gateway, or in case of high interference, we choose a higher value. However, the increase will result in lower speed and longer Time on Air, which is limited by local regulations. Another parameter is Coding Rate, which expresses the ratio of bits carrying useful data and redundant bits for detecting or correcting errors caused by transmission. Allowable values are from 4/5 to 4/8, again choosing a higher redundancy in case of worse radio conditions.

After defining these parameters we are able to determine equations for calculating Time on Air.

$$t_{preamble} = (n_{preamble} + 4, 25) \cdot T_s \quad (2)$$

Where  $n$  is the length of the preamble and  $T$  is the period of symbol duration.

$$t_{payload} = T_s(8 + \max(\text{ceil}(\frac{8PL - 4SF + 28 + 16CRC - 20H}{4(SF - 2DE)}))(CR + 4), 0) \quad (3)$$

Where  $PL$  represents payload in bytes and  $CR$  represents coding rate. The following values are boolean, representing the activation or deactivation of the parameter –  $CRC$  is the checksum,  $DE$  is Low Data Rate Optimize and  $H$  is the header where the reverse logic is applied, where 0 means enabled.

$$ToA = t_{packet} = t_{preamble} + T_{payload} \quad (4)$$

Where  $t$  are the corresponding preamble and payload times from previous relations.

In the LoRaWAN topology, there are end devices and gateways. End devices are e.g. sensor stations like the one presented here. Gateways are devices that forward LoRaWAN packets to servers on the Internet. The principle of the protocol is as follows. The end device sends a message which is received by all gateways in range. The gateways forward the message to the network server, which handles duplication, message decoding, and forwarding to the application server.

LoRaWAN defines three operational classes. Class A specifies duplex communication with the gateway. The end device can send at any time, after sending it opens two-time windows when it listens for a response from the gateway. It is forbidden to use both, because if a response is received in the first, the second window will not open again. This class is best suited for battery-powered devices, as they spend most of their time asleep. The other two classes extend the functionality of class A. Class B adds the ability to add scheduled reception windows. When using this class, the gateway periodically sends synchronization beacon frames. It is because of the processing of beacon frames that Class B devices typically have less battery life compared to Class A. Class C devices have a constantly open receive window, this causes zero delays in the down-link direction. On the other hand, these devices are constantly active, so they are most often grid powered.

#### 4. MEASURING AND DATA SENDING

The process described above is implemented in software in the microcontroller. The selected sensors are MH-Z16 for reading the concentration of  $CO_2$ . Its range is 0-5000ppm, which is fully sufficient considering standard outdoor conditions, the accuracy reported is 100ppm + 6%. It also offers several readout options such as analog voltage level, PWM signal, and the UART interface used in this project, where there is also an option for possible calibration of the probe. The temperature sensor chosen is the DS18B20 in a waterproof design with an accuracy of 0.2°C and a measurable temperature of -55°C - 125°C. The thermometer communicates over a 1-Wire bus for which there is no built-in support in the STM libraries, so a communication protocol for this interface had to be written. The measurement starts by collecting several samples of temperature and  $CO_2$  concentration, then the program waits and repeats. From the concentrations, a concentration versus time function is calculated using linear interpolation, and from the temperatures, the average temperature throughout the measurement is calculated. In order to the importance of the initial and final measurements, several samples are taken to minimize the effect of the error. All variables in the cycle, such as the number of samples, time between measurements, volume, and area of the chamber, can be configured using macros. Before the actual insertion into 1, its derivative is calculated from the concentration function. After the measurement cycle is completed, the collected data are sent to the server. It sends mainly the calculated flux, but also all the data that was used to calculate it. To provide connectivity, services of The Things Network are currently used, which offers a limited duty cycle of 0.0347% in the basic free model, compared to the maximum allowed 1%. If this duty cycle is not sufficient it is possible to reduce Time On Air by sending only the monitored concentration. Considering the highest possible Spreading Factor = 12 and Coding Rate = 4/8, the Time On Air value for the current 20B payload is 926ms after inserting 2 3 4. With a daily transmission time of 30 seconds per day, up to 32 messages can be sent.

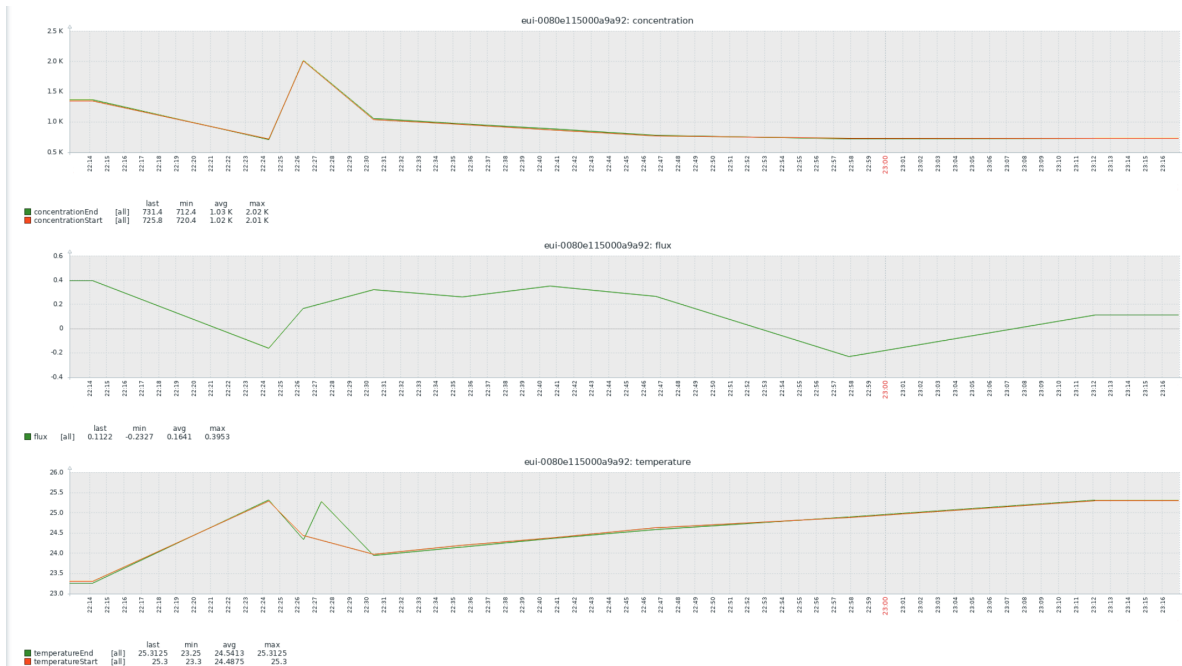
#### 5. DATA PROCESSING AND VISUALIZATION

When a message is received on The Things Network server, a module called broker processes the message first. The broker takes all received duplicates, extracts gateway-specific information from them, and adds it to the final message. After that, we can work with the received message. There are several formatting options available. In this case we will use the JavaScript formatting option. The received data are then

Time	Entity ID	Type	Data preview	Verbose stream	Pause	Clear
↑ 15:48:09	eui-0080e115000a9a92	Forward uplink data message	Payload: { measurements: [...] } 00 02 8D 00 13 FE 3E 09 ... FPort: 2 Data rate: SF12BW125 SNR: -9 RSSI: -113	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="button" value="Clear"/>
↑ 15:48:38	eui-0080e115000a9a92	Forward uplink data message	Payload: { measurements: [...] } 00 02 8C 00 13 FE 3E 09 ... FPort: 2 Data rate: SF12BW125 SNR: -3 RSSI: -111	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="button" value="Clear"/>
↑ 15:48:19	eui-0080e115000a9a92	Forward uplink data message	Payload: { measurements: [...] } 00 02 86 00 13 FE 3E 09 ... FPort: 2 Data rate: SF12BW125 SNR: -8.75 RSSI: -114	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="button" value="Clear"/>
↑ 15:48:09	eui-0080e115000a9a92	Forward uplink data message	Payload: { measurements: [...] } 00 02 8C 00 13 FE 3E 09 ... FPort: 2 Data rate: SF12BW125 SNR: -4 RSSI: -111	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="button" value="Clear"/>
↑ 15:47:38	eui-0080e115000a9a92	Forward uplink data message	Payload: { measurements: [...] } 00 02 9E 00 13 FE 3E 09 ... FPort: 2 Data rate: SF12BW125 SNR: -4.25 RSSI: -111	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="button" value="Clear"/>
↑ 15:47:28	eui-0080e115000a9a92	Forward uplink data message	Payload: { measurements: [...] } 00 02 AB 00 13 FE 3E 09 ... FPort: 2 Data rate: SF12BW125 SNR: -3 RSSI: -111	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="button" value="Clear"/>

**Figure 2:** Received data on The Things Network

decoded into a human-readable form. When it comes to sending message to the application server. Again, there are several predefined options for the integration with the most well-known IoT services. Some of these were tested for data visualization, but even for use with a single device, no option was found for use free of charge. For this reason, it was decided to use own solution, where the received messages are sent to an existing server with a modified open source monitoring tool Zabbix. This was achieved using the universal WebHook integration. The final pitfall to successfully displaying the measured data is the fact that the Zabbix implementation only accepts GET requests with parameters in the URL, while The Things Network sends a JSON file using the POST method. Thus, an additional intermediate step is required. A PHP script is prepared on the monitoring server, which receives the message parses the JSON file, and finally calls the *curl()* method, which sends a GET request to Zabbix with JSON parameters.



**Figure 3:** Zabbix graphs

## 6. CONCLUSION

In the current era of constant surveillance and monitoring requirements, the IoT solutions are proving to be great option for many diverse applications of data acquisition and aggregation. Their biggest advantage is the low energy consumption, which allows to monitor remote locations without the need for intervention for years. The proposed prototype, which focuses mainly on the development of firmware for the STM32WL system-on-a-chip platform on a development kit, has proven to be functional and usable. Hence, after a deeper power analysis and optimization, it will be ready for deployment in the target area, and to undergo a complex comparative analysis against the older system solution driven by a dedicated MCU and a dedicated LoRaWAN radio module, which is the main contribution of this project. An additional benefit is the visualization using plots available in a custom environment, which allows comfortable monitoring of the physical quantities trends over time.

## REFERENCES

- [1] JURÁŇ, Radovan. *Field sensor network for microclimatological measurements*. Brno, 2020, 91 p. Master's Thesis. Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Radio Electronics. Advised by Ing. Aleš Povalač, Ph.D.
- [2] *LI-COR* [online]. 2022 [cit. 2022-03-24].  
Online: [https://www.licor.com/env/products/soil\\_flux/long-term](https://www.licor.com/env/products/soil_flux/long-term)
- [3] *SenseCAP Wireless CO2 Sensor* [online]. 2022 [cit. 2022-03-24]. Online: [https://lora-alliance.org/lora\\_products/sensecap-wireless-co2-sensor-lorawan-us915mhz/](https://lora-alliance.org/lora_products/sensecap-wireless-co2-sensor-lorawan-us915mhz/)
- [4] *Smart CO2 LoRa Sensor* [online]. 2022 [cit. 2022-03-24]. Online: <https://mutelcor.com/lora-co2-digitalmelder/>
- [5] *LoRaWAN 1.0.3 Specification* [online]. In: . Beaverton, OR USA: LoRa Alliance, 2018 [cit. 2021-11-27]. Online: <https://lora-alliance.org/wp-content/uploads/2020/11/lorawan1.0.3.pdf>
- [6] Building wireless applications with STM32WB Series microcontrollers. *STMicroelectronics* [online]. STMicroelectronics, 2021 [cit. 2021-12-13]. Online: [https://www.st.com/resource/en/application\\_note/an5289-building-wireless-applications-with-stm32wb-series-microcontrollers-stmicroelectronics.pdf](https://www.st.com/resource/en/application_note/an5289-building-wireless-applications-with-stm32wb-series-microcontrollers-stmicroelectronics.pdf)
- [7] Network Architecture. *The Things Network* [online]. Amsterdam, Netherlands: The Things Industries, 2016 [cit. 2021-12-11]. Online: <https://www.thethingsnetwork.org/docs/network/architecture/>
- [8] WANG, Miao et al. Comparison of eddy covariance and chamber-based methods for measuring CO2 flux in a temperate mixed forest. *Tree Physiology* [online]. 2010, (30), 149-163 [cit. 2021-12-04]. Online: <https://academic.oup.com/treephys/article/30/1/149/1648371?login=true>
- [9] *STM32WL Series: Long-range wireless STM32WL microcontrollers* [online]. [cit. 2022-03-24]. Online: <https://www.st.com/en/microcontrollers-microprocessors/stm32wl-series.html>



# Role of battery capacity in charging habits of battery electric vehicle users

T. Řehořek

Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech Republic

E-mail: [203333@vut.cz](mailto:203333@vut.cz)

**Abstract**—This paper is focused on the role of battery capacity in charging habits of battery electric vehicle (BEV) users. For this purpose, a simulation that mimics the users' behavior was created. Using this simulation, the differences in charging frequencies in vehicles with different battery capacities were studied. Using the same simulation, a load shape of power demand was derived that showcases how the technological advances in regard to increased battery capacities may affect this power demand.

**Keywords**—Battery electric vehicle, battery capacity, load modeling, electric vehicle charging

## 1. INTRODUCTION

The share of BEV among the vehicle fleet is gradually growing. Consequently, the impact of this rise on the power grid is being explored in many research papers. The main effect to consider is the increased demand created by charging of BEVs. Load shape of such demand depends partially on the behavior of individual drivers, especially regarding daily driven distances, conditions and time of connection to charging, etc. The demand created by charging also depends on the structure and parameters of the vehicle fleet, such as consumption of energy and battery capacity. Charging infrastructure in the place of frequent charging also plays a role. This paper aims to show this demand created by BEV charging with special focus on the role of battery capacity on the charging habits of drivers.

## 2. SIMULATION PARAMETERS

To evaluate the aforementioned effect of rise in battery sizes the simulation aims to mimic the behavior of BEV users to derive the load shape created by the consumption of BEV during their charging over the course of one week. The simulation considers conditions in small sub-urban municipality. The key parameters of the simulation are mentioned in the next paragraphs.

Key parameter is daily energy consumption, that can be calculated by estimating daily driven distance and multiplying it by consumption of energy per km. To estimate the daily driven distance, data from study *Česko v pohybu* were used [1]. On average, one-way trip distance is 15.5 km. Standard deviation of these data is 26.8. According to [2], for estimation of daily driven distances, a log-normal probability distribution function (PDF) can be assumed. Considering that the number of trips in given day is usually 2 (as per data in [1]), the daily driven distance for each BEV is assigned according to log-normal PDF with mean distance of 31 km and standard deviation of 26.8. Logarithmic values calculated from arithmetic values of the associated normal distribution are equal to  $\mu=3.155$  and  $\sigma=0.747$ . In respect to consumption, there are three types of BEV considered in the simulation. Type A with 180 Wh/km consumption and 20 % share in the fleet, type B with 207 Wh/km with 65 % share in the fleet and type C with 263 Wh/km with 15 % share in the fleet.

Another aspect of the simulation is the conditions in which the connection to charging can be assumed. If the mentioned daily driven distances and consumptions are considered, the battery state of charge (SoC) at the end of the day may still remain on a level that does not warrant a plug in for charging. For this reason, a model that does not consider daily charging was created. Specifically, to model driver behavior regarding charging habits a probability of connection is calculated based on SoC at the end of the day. For this calculation an exponential dependence between SoC and probability of connection to charging ( $p$ ) is assumed, based on [2].

$$p = \frac{1}{1 + e^{-k_p \cdot (-SoC + x_p)}}, \quad (1)$$

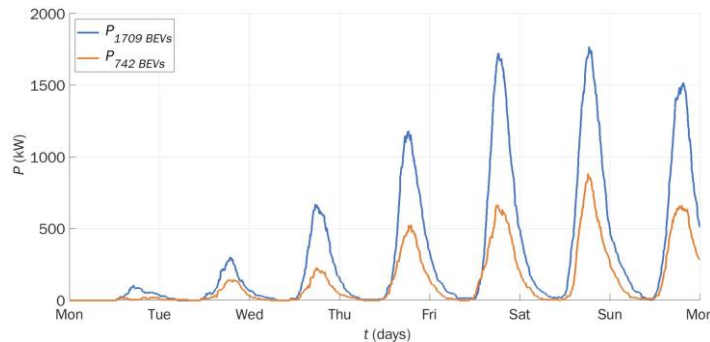
where  $k_p$  represents a coefficient of range anxiety,  $SoC$  is state of charge (%),  $x_p$  represents a SoC in which  $p = 0.5$ . In [2], parameters  $k_p$  and  $x_p$  were set, based on empirical evidence, to  $x_p = 35$  and  $k_p = 0.13$ . Same parameters are assumed in the simulation.

The simulation considers conditions in a small sub-urban municipality, where high share of home charging is to be expected [3] due to higher number of private parking spaces with access to a private charger [3]. The simulation considers three types of chargers, in respect to power. Specifically, 3.6 kW, 6.9 kW and 11 kW chargers are considered. Power delivered from chargers is lowered by expected losses that are defined by the consideration of 95 % efficiency of charging. Another important aspect of the simulation is the time in which the connection to charging can be expected. According to [2], normal PDF can be assumed. The simulation used in this paper also uses normal PDF with mean time of connection set to 16.53 h with standard deviation set to 1.99 h. In the simulation it is assumed that all vehicles get charged to maximum SoC.

### 3. BATTERY SIZE

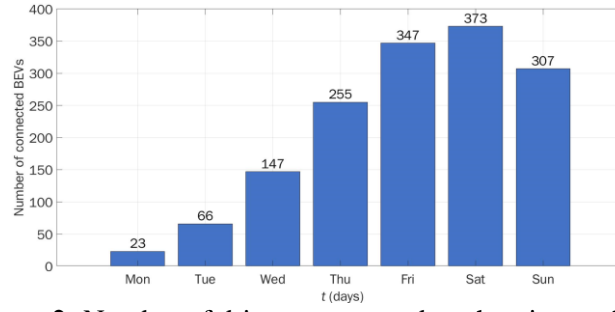
As mentioned in the beginning, the battery size of vehicles can be one of many factors that play a role in drivers charging behavior. It can be expected that with the growing market share of BEVs, the capacities of battery packs will increase in efforts to make the drivable single charge distance longer. The change in battery sizes could lead to change in users behavior in regard to charging habits.

Based on information in [5], the average battery capacity of the most used BEVs in Czech Republic for year 2017 was about 40 kWh. Based on the top 10 most used BEVs in CZ for year 2021 (as per statistic in [6]) the average capacity is around 54 kWh (battery capacities for each BEV were gathered in electric vehicle database [7]). The average capacity of the top 10 newly sold BEVs in 2021 is about 60 kWh, according to information in [7] and [6]. The analysis [5] assumed the rise in battery sizes to be 2 kWh/year and according to another study in this field [8], a rise in battery capacity by 2030 is expected to be 50 %. For baseline simulation a current condition of the battery sizes was considered. This variant of simulation assumes three types of BEVs with their capacity and share in fleet. Type A is a vehicle with 35 kWh capacity and 20 % share in fleet, type B has a 50 kWh capacity and 65 % share, type C has a 80 kWh capacity and 15 % share. The simulation assumes 1709 BEVs, which simulates a condition that would be present at the place of interest, if all the vehicles would be electric. The availability of charging infrastructure in the place of the longest stay is also important. Because of that another simulation assuming that only BEVs parked in garages would be charged in home was carried out. The number of garages in the place of interest is 742. Results of both simulations are in the following figure.



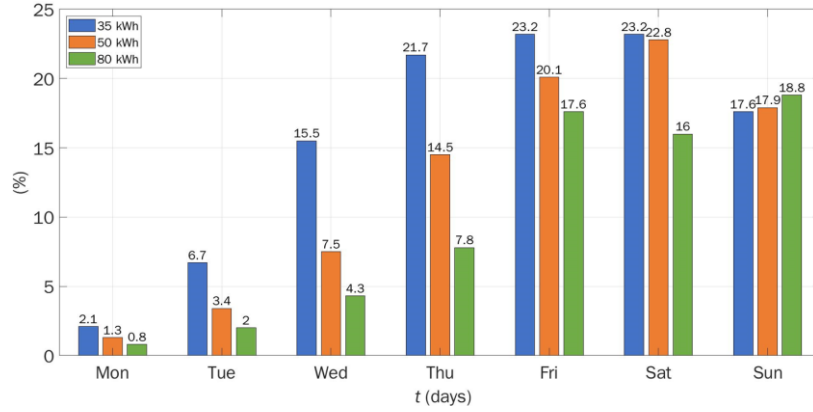
**Figure 1:** Results of baseline simulation

Important note is that the simulation assumes that all BEVs start the given week fully charged. This situation may not be far from reality as it is quite possible that drivers would want to start weekend fully charged. Since only a few users make trips in lengths, that would warrant a charge after just one day, the load level in the beginning of week is quite low. As the week progresses the expended energy cumulates, and the SoC of vehicle decreases resulting in higher peaks in demand for charging in later days of the week. These peaks directly correlate to number of users that charge their vehicle. In the figure below, the number of drivers charging their vehicles is shown (numbers are relevant for simulation considering 1709 BEVs). Compared to Figure 1, the correlation can be seen quite clearly.



**Figure 2:** Number of drivers connected to charging each day

Figure 2 shows that the number of drivers connected to charging each day is a key factor, for the total load demand. The number of drivers that will connect is dependent on the capacity of the vehicles, which can be demonstrated by figure 3. It shows how many vehicles from specified vehicle category are being charged each day of the week.



**Figure 3:** Relative number of drivers (from each vehicle category) that connect to charging

The share of vehicle types is not uniform, so to compare the number of connections across the categories, the number of drivers that connect to charging is shown in values relative to the amount of vehicles in given category. The figure shows that vehicles with the lowest battery capacity (35 kWh) are charged more frequently than the vehicles with higher capacity batteries. Since the role of battery capacity has been established, an analysis of the effect of growing vehicle capacity on the demand created by charging can be examined. The assumed battery capacity growth is 10 kWh over 5-year period. The growth is similar to the one assumed in [5]. Important statistics from the simulation are shown in following table.

**Table I:** Number of charged vehicles per week

Year	Number of vehicles connected to charging per week	Average recharged energy	Maximum recharged energy	Average charging time	Average charging end time
(-)	(-)	(kWh)	(kWh)	(h)	(hh:mm)
2020	1527	35.8	75.6	6.6	23:12
2025	1242	40.7	80.3	7.5	00:12
2030	906	45.4	91.0	8.2	00:54
2035	649	50.0	104.3	9.3	01:36
2040	483	54.5	106.1	9.9	02:24

With growing battery capacities, vehicles stay in the range of SoC that does not warrant charging for longer time, which in turn results in users charging their vehicles less frequently, this leads to lower number of charged vehicles per week. It should be noted that vehicles that do not get charged in shown week will have to be charged in the next week. The simulation was made to simulate just one week. Because vehicles stay longer without charging, they expand more energy between charges, which leads to increased amount of energy, that needs to be recharged. With this the time it takes vehicles to fully

charge increases and subsequently the time when vehicle is fully charged moves more into the following day. As shown earlier vehicles with higher capacities are not charged as often as vehicles with lower capacities. Because the simulation shows only one week, vehicles with high capacities (especially type C) have lower share among vehicles being charged. If a subsequent week was simulated, these vehicles would probably have to be charged, with amount of energy to be recharged higher than shown averages. To show the amount of energy recharged by the vehicles with highest battery capacities among the fleet, the maximum energy recharged is shown. This energy increases with battery capacities as well.

#### 4. CONCLUSION

The main aim of this paper was to show the correlation between BEV user behavior and battery capacity. It was shown that vehicles with smaller battery capacities tend to be charged more frequently. Same effect was shown in simulation that attempted to see how the technological advance leading to increased battery capacities may affect the load demand created by charging. It was shown that growing battery capacities lead to higher amount of energy that needs to be recharged. The charging time for vehicles depends on the amount of energy being recharged and the power of the charger. Increase in recharged energy leads to increase in charging duration and point at which BEV reaches full SoC shifts to later time. If BEV users would want to fully charge their vehicle, increased battery capacities could lead to increased need for higher power chargers, so the vehicles can reach full SoC before their next use (typically overnight). This in turn could lead to increased peaks in power load created by charging. Another possibility is that users may utilize lower power chargers, possibly due to financial reasons and charge their vehicles more frequently, so the amount of recharged energy is not as high. The model shows an ideal situation, where no effects other than battery capacity come to play. It is important to note that there are more factors influence charging behavior. For example, it is possible that in the future some users may utilize devices allowing for charging in times when electricity is cheaper or utilize electricity from photovoltaic panels. These external influences could result in deviation from examined effect of increased battery sizes. The simulation itself is a groundwork for research that will follow.

#### REFERENCES

- [1] "Česko v pohybu [Czechia in motion]," Centrum dopravního výzkumu [Transport research center], 2020. [Online]. Available: <https://www.ceskovpohybu.cz/>.
- [2] O. Lehtinen, S. Pitkäniemi, A. Weckman, M. Aikio, M. Mabano and M. Lehtonen, "Electric Vehicle Charging Loads in Residential Areas of Apartment Houses," 2020 21st International Scientific Conference on Electric Power Engineering (EPE), 2020, pp. 1-6, doi: 10.1109/EPE51172.2020.9269191.
- [3] Z. Fotouhi, M. R. Hashemi, H. Narimani and I. S. Bayram, "A General Model for EV Drivers' Charging Behavior," in IEEE Transactions on Vehicular Technology, vol. 68, no. 8, pp. 7368-7382, Aug. 2019, doi: 10.1109/TVT.2019.2923260.
- [4] "Load modeling and distribution planning in the area of electric mobility," 26th International Conference on electricity distribution, 2021, pp. 21–24.
- [5] "Dílčí studie pro pracovní tým A25 - Predikce vývoje elektromobility v ČR [Partial study for work group A25 - Prediction of electromobility development in the Czech Republic]", Ministerstvo průmyslu a obchodu [Ministry of Industry and Trade], 2018. [Online]. Available: [https://www.mpo.cz/assets/cz/energetika/strategicke-a-koncepcni-dokumenty/narodni-akcni-plan-pro-chytre-site/2019/10/Studie-NAP-SG-A25\\_Elektromobilita.pdf](https://www.mpo.cz/assets/cz/energetika/strategicke-a-koncepcni-dokumenty/narodni-akcni-plan-pro-chytre-site/2019/10/Studie-NAP-SG-A25_Elektromobilita.pdf).
- [6] "Passenger cars M1: CZ Top 10 Models Fleet - Electricity (2021)", European Alternative Fuels Observatory, 2021. [Online]. Available: <https://www.eafo.eu/vehicles-and-fleet/m1#>
- [7] "Electric Vehicle Database," 2022. [Online]. Available: <https://ev-database.org/>
- [8] F. G. Venegas, M. Petit and Y. Perez, "Impact of Non-Systematic Electric Vehicle Charging Behaviour on a Distribution Substation," 2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), 2019, pp. 1-5, doi: 10.1109/ISGTEurope.2019.8905710.

# Study of the ecological impact of production, operation and disposal of various types of cars

M. Los<sup>1</sup>, K. Jaško<sup>1</sup>

<sup>1</sup> Department of Electrical and Electronic Technology, Faculty of Electrical Engineering and Communication, Brno University of Technology, Czechia

E-mail: xlosma00@vut.cz, xjasso00@vut.cz

**Abstract**—The life cycle of a car includes three basic phases: production, operation and end-of-life. All these phases can potentially produce emissions that have a direct impact on the environment. However, different types of cars may produce different amounts of emissions during the life cycle phases, depending not only on the type of propulsion but also on local factors such as the emission factor of the energy mix. The results of the research showed that the potential environmental benefit of lower greenhouse gas (GHG) emissions of battery electric vehicles depends primarily on the emission factor of the energy mix. In countries such as Poland, where the majority of electricity produced is generated by burning coal and lignite, the life cycle environmental performance of battery cars will be higher than that of combustion cars. Conversely, in countries with a favorable energy mix, such as Sweden, the potential to reduce transport emissions by replacing the existing car fleet with electric cars is significant. Practical part of this papers aims to compare the environmental impact from the view of greenhouse gas emissions throughout the life cycle of conventional petrol, diesel, natural gas and electric cars in local conditions.

**Keywords**—Ecology, cars, emissions, life cycle, production, operation

## 1. INTRODUCTION

With the current pressure to minimize the environmental impact and greenhouse gas emissions of transport, alternative forms of vehicle propulsion, primarily battery electric cars, are being promoted. However, the carbon footprint of BEVs (battery electric vehicles) can vary considerably depending on the country in which they are operated, because although they do not produce emissions directly during operation, the emissions from electricity generation depend on the emission factor of the local energy mix.

The life cycle of a car consists of production, operation and end of life. It is then important to further divide the operational phase into primary and secondary emissions. As this is a cradle-to-grave analysis, the life cycle of the car and the resulting emissions are considered from the point of extraction of raw materials, their processing and conversion into parts, assembly, logistics, car operation and fuel production processes to the end of life and associated recycling.

4 types of propulsion are analyzed in this work: petrol, diesel, CNG (compressed natural gas) and BEV.

## 2. MODELLING OF REFERENCE CARS

For the purpose of assessing the cars that are currently in common use in Czechia, 4 reference cars were modelled, 1 for each type of drive mentioned above. The technical specifications of these cars are based on at least 4 representatives of popular C-segment vehicles, i.e., cars of the lower middle class with the intention to represent actual cars actually used in the present on Czech roads. The parameters of the modelled reference vehicles are given in Table I.

The next important parameter to be set is the mileage of the reference car. According to Kazda [1] the average car in Czechia travels 29 km daily and as reported by ACEA (European Automobile Manufacturers' Association) [2], the average car age in Czechia is currently around 15 years. After a

simple calculation with the use of the mentioned data, it emerges that average local car's mileage over its life hovers around the 160 000 km mark, which is a crucial piece of information for further calculations.

**Table I:** Reference car technical parameters

	Unit	Petrol	Diesel	CNG	BEV
Enigne displacement	l	1.5	2.0	1.5	-
Maximum power	kW	116	103	89	98
Curb weight	kg	1254	1378	1309	1522
Curb weight without battery	kg	-	-	-	1208
Battery weight	kg	-	-	-	314
Battery capacity	kWh	-	-	-	46
WLTP emissions	g CO <sub>2</sub> /km	127	116	101.75	0
WLTP combined fuel consumption	l/100km	5.6	4.4	-	-
	m <sup>3</sup> /100km	-	-	5.7	-
	kWh/100km	-	-	-	16

### 3. CALCULATION

The aim of this paper is to evaluate a car's lifecycle environmental impact from the greenhouse gas emissions point of view; therefore, the process of calculation shall lead to a final value of greenhouse gas emissions in tons of CO<sub>2</sub>eq for each propulsion type reference vehicle. The base equation used for the calculation is (1), which is a sum of emission values coming from the individual phases of a lifecycle.

$$E = E_{prod} + E_{op} + E_{EOL} \quad (1)$$

The greenhouse gas emissions emitted during the phase of production were then calculated according to (2), where  $E_{prod}$  stands for the amount of greenhouse gas emissions created during the process of vehicle manufacturing in tons of CO<sub>2</sub>eq,  $m_{curb}$  is the curb weight of the vehicle and  $e_{prod}$  is the emission factor of the production per kilogram of the vehicle's weight in kg CO<sub>2</sub>eq/kg. The same value 5 kg CO<sub>2</sub>eq/kg of  $m_{curb}$  was used for the petrol, diesel and CNG car manufacturing as well as for manufacturing of the BEV glider. Both Hawkins et al. [3], who determined this value in terms of cradle-to-gate emissions by compiling an inventory of generic automotive components (about 140 subcomponents) and inferring their emissions intensity, and Schuller and Stuart [4], who reported an emission factor of 6 kg CO<sub>2</sub>eq/kg for 2017 and assumed 5 kg CO<sub>2</sub>eq/kg for 2030, lean towards this value.

$$E_{prod} = m_{curb} \times e_{prod} \quad (2)$$

The calculation of BEV production emissions requires a slightly different approach, since the production of the car battery has a significant impact on the final amount of greenhouse gas emissions of production. Equation (3) is used to calculate the BEV emissions separately for the vehicle's glider and the battery based on its capacity  $C_{bat}$  and emission factor of battery production resulting from analysis of 50 battery production studies [5].

$$E_{prod} = (m_{glid} \times e_{prod}) + (C_{bat} \times e_{bat}) \quad (3)$$

The phase of operation in this paper is concerned primarily with the well-to-wheel emissions, which represent the fuel cycle from its extraction and treatment, to logistics and its combustion or use in the



target vehicle. Emissions from the maintenance were not considered, because parts like tires are equal to all the reference vehicles and would make no difference. Other maintenance processes would only add an insignificant amount of emissions to the final result. The total well-to-wheel emissions were calculated as a sum of Well-to-Tank and Tank-to-Wheel emissions, as is shown in equation (4).

$$E_{op} = E_{WTT} + E_{TTW} \quad (4)$$

While calculating well-to-tank emissions (5), local pathways of fuel were considered and their emissions factors were taken from the JEC (EU Joint Research Centre, EUCAR, Concawe) Well-to-Tank report v5 [6]. The emission factor  $e_{WTT}$  is then multiplied by the fuel consumption of the reference cars over their operation phase mileage  $V$ . As for the BEV, the emission factor of fuel production represents the emission factor of the local energy mix which, according to European Environment Agency [7], is 436,6 g CO<sub>2</sub>eq/kWh.

$$E_{WTT} = e_{WTT} \times V \quad (5)$$

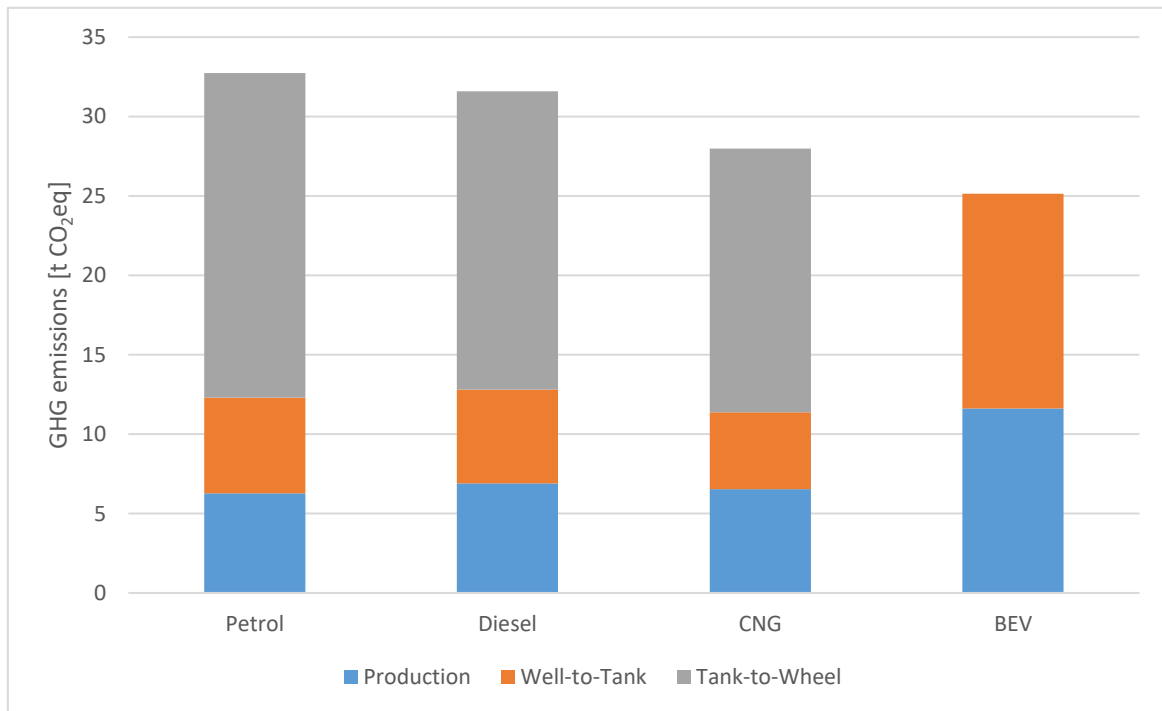
Tank-to-Wheel emissions are then calculated as a sum of emission factors representing WLTP (Worldwide harmonized light-duty vehicles test procedure) CO<sub>2</sub> emissions in addition with CO<sub>2</sub> equivalent of N<sub>2</sub>O and CH<sub>4</sub> calculated as fractions of their emission limits according to EURO 6 standard multiplied by their global warming potential and then multiplying the sum by the reference car life mileage  $d$  (6). For the case of BEV, its Tank-to-Wheel emissions result in naught, since the BEV produces no tailpipe emissions during the operation.

$$E_{TTW} = (e_{WLTP} + e_{N_2O} + e_{CH_4}) \times d \quad (6)$$

The emissions produced in the end-of-life phase of the vehicles have not yet been calculated at the time of writing this paper and are the subject of further work.

#### 4. RESULTS AND CONCLUSION

As of the current progress, the results of greenhouse gas intensity of the selected reference vehicles are visible in Fig. 1.



**Figure 1:** Calculation results

The results show that majority of GHG emissions produced by the petrol, diesel and CNG cars come from the Tank-to-Wheels process i.e., the fuel combustion. Emissions from the production of the fossil fuel vehicles are roughly the same, around 6,5 t CO<sub>2</sub>eq, while production of the BEV is almost twice as GHG intensive. Regarding the energy mix emission factor in Czechia, it is still less emission-intensive, than the whole life cycle of the three considered fossil fuels.

In conclusion, the considered reference BEV emits the least greenhouse gas emissions during its lifecycle among all of the four compared propulsion type vehicles, although its emissions intensity is heavily dependent on the local energy mix, as it varies across the countries and the Well-to-Tank emissions are this vehicle's most notable greenhouse gas emissions source. It is highly probable, that BEVs operated in countries with more environmentally friendly energy sources, such as Sweden or Norway, would account for much lower amount of greenhouse gas emissions over their lifecycle.

It is important to note, that emissions from the end-of-life phase are to be added, although even despite the fact that the recycling process of BEV batteries is more emission intensive, it will, probably, not shift the results.

## REFERENCES

- [1] M. Sedlák, "U elektromobilu stačí měnit pneumatiky a kapalinu do ostříkovačů, říká expert", *Obnovitelně*, c2020.
- [2] ACEA, "Vehicles in use Europe: January 2021," 2021.
- [3] T. R. Hawkins, B. Singh, G. Majeau-Bettez, and A. H. Strømman, "Comparative Environmental Life Cycle Assessment of Conventional and Electric Vehicles", *Journal of Industrial Ecology*, vol. 17, no. 1, pp. 53-64, Aug. 2012.
- [4] A. Schuller and C. Stuart, "“From cradle to grave: e-mobility and the energy transition”: Addendum for Italy, the United Kingdom, Spain and the European Union to “Le véhicule électrique dans la transition écologique en France”", August 2018.
- [5] C. Aichberger and G. Jungmeier, "Environmental Life Cycle Impacts of Automotive Batteries Based on a Literature Review", *Energies*, vol. 13, no. 23, 2020.
- [6] M. Prussi, M. Yugo, L. De Prada, M. Padella, R. Edwards, and L. Lonza, "JEC Well-to-Tank report v5: EUR 30269 EN", 2020.
- [7] "Greenhouse gas emission intensity of electricity generation by country", *European Environment Agency*, 2021.
- [8] J. Mikulka, D. Chalupa, J. Svoboda, M. Filipovič, M. Repko and M. Maxová, "Multimodal and Multiparametric Spatial Segmentation of Spine," 2020 19th International Conference on Mechatronics – Mechatronika (ME), 2020, pp. 1–5, [doi 10.1109/ME49197.2020.9286666](https://doi.org/10.1109/ME49197.2020.9286666).
- [9] P. Horowitz and W. Hill, *The Art of Electronics*. New York: Cambridge University Press, 2021.

# Mobile Robot with Telepresence Control

S. Svědih<sup>1</sup>, L. Žalud<sup>2</sup>

<sup>1</sup>Master Degree Programme (2), Brno University of Technology

<sup>2</sup>Department of Control and Instrumentation, Brno University of Technology

E-mail: [xsvedi01@vutbr.cz](mailto:xsvedi01@vutbr.cz), [zalud@feec.vutbr.cz](mailto:zalud@feec.vutbr.cz)

**Abstract**—This paper focuses on developing a mobile robotic system capable of wireless movement and control. The robot can be controlled immersively via the virtual reality headset, or from any other device like a laptop or a mobile phone connected to the same Wi-Fi network. The robot houses a stereographic camera pair that streams near real-time video allowing the user to pan & tilt using only a head movement. The solution proposed in this paper also focuses on keeping the overall price low while maintaining as much room for future improvements as possible.

**Keywords**—Virtual Reality, Wireless, Robotic System, Stereographic Camera Pair

## 1. INTRODUCTION

Visual telepresence is a field that has experienced a lot of breakthroughs in the recent decades. With the invention of the internet and its speed increasing at an unprecedented rate in the last years, it finally became possible to transmit a high resolution video stream with relatively low latency. Hardware has become so efficient in encoding and decoding, that its now possible to work with such streams on a relatively cheap devices. Also virtual reality is seeing first generation of consumer devices to offer true wireless and fully immersive experience.

Most of the already existing telepresence systems offer very limited motion capabilities, require difficult setup or are just too expensive for individuals to buy. The cost is reasonable once you factor in all the requirements they need to meet in order to function seamlessly in all kinds of environments, carry and use all kinds of tools to complete their predefined tasks etc.

Stripping away most of the cost, requirements and trying to work only with off the shelf parts makes visual telepresence robots easier to get by consumers, maybe even inspiring the masses to build upon those and create new things.

This paper proposes a mobile robot based on the open source 3D printed platform [1] that uses a Nvidia Jetson Nano microcomputer capable of encoding and wirelessly transmitting two FullHD video streams powered by a Li-Po battery. The robot also contains a Raspberry Pi Pico microcontroller board to control 2 stepper motors to allow for robot movement and 2 servos to allow for head tilt & pan. Due to a need for battery voltage regulation and stepper motor drivers implementation a custom PCB has been designed and manufactured. The proposed robot runs a script that creates a RTSP over UDP server with camera streams being provided and to receive a control input stream. This allows for almost any device to connect and control the robot. This paper primarily uses a Oculus Quest 2 VR headset.

## 2. ROBOT ELECTRONICS AND WIRING

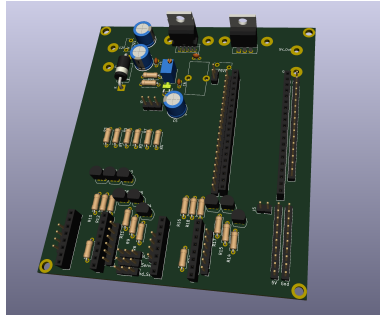
The crucial part behind all the input and video processing is a Nvidia Jetson Nano microcomputer. It is capable of encoding two Full HD resolution streams with hardware acceleration while simultaneously streaming them using GStreamer [2] to multiple other devices. It also has enough processing power to communicate with Raspberry Pi Pico microcontroller to send commands for all the actuators used. Doing all this, it still has plenty of power left for possible future improvements. It has maximum of 25 W power draw, so it doesn't even generate much thermal output to be an issue.

The use of Raspberry Pi Pico microcontroller was a necessity due to a low output current capabilities of Jetson Nano and also for the lack of a precise clock signal needed to drive servo motors. The communication between those chips is handled via the UART protocol.

Custom PCB has been designed and manufactured for easier integration of input voltage regulation that can range anywhere between 8-35 V, stepper motor driver connection (working with either A4988

or DRV8825 stepper motor drivers) and current boost for servo motors. The proposed PCB is about the same size as the Jetson Nano microcomputer, but could be definitely made smaller if SMD parts were to be used.

The robotic platform currently uses two 2MP IMX291 camera modules with a field of view of 120° using USB 2.0 connection. Those modules are capable of delivering 1920x1080 @ 30 FPS video and the also contain a microphone that can be used to enhance the immersiveness even more in the future.



(a) Custom made PCB



(b) Nvidia Jetson Nano

**Figure 1:** Robotic platform control electronics

### 3. ROBOTIC PLATFORM AND ELECTRONICS INTEGRATION

An open source 3D printed platform [1] with minor modifications has been used as a vessel for all electronics. This platform is by no means simple, easy to maintain or even capable of carrying any kind of heavy load. On the other hand, it provides the robot with character which is arguably more important given the aim of this paper discussed in the introduction.

In contrast to the original project, stepper motors have been used instead of DC motors. This adds additional signals to control them making the drive a bit more complex, but also has a potential of providing better feedback if some kind of mapping algorithm were to be used in the future.

All electronics has been fitted into the body of the robot leaving almost no space for expansion. Future versions of the PCB could help to regain some of that space back while also reducing the amount of cables used as they are somewhat prone to disconnect during the operation.

### 4. SOFTWARE ON THE ROBOTIC PLATFORM SIDE

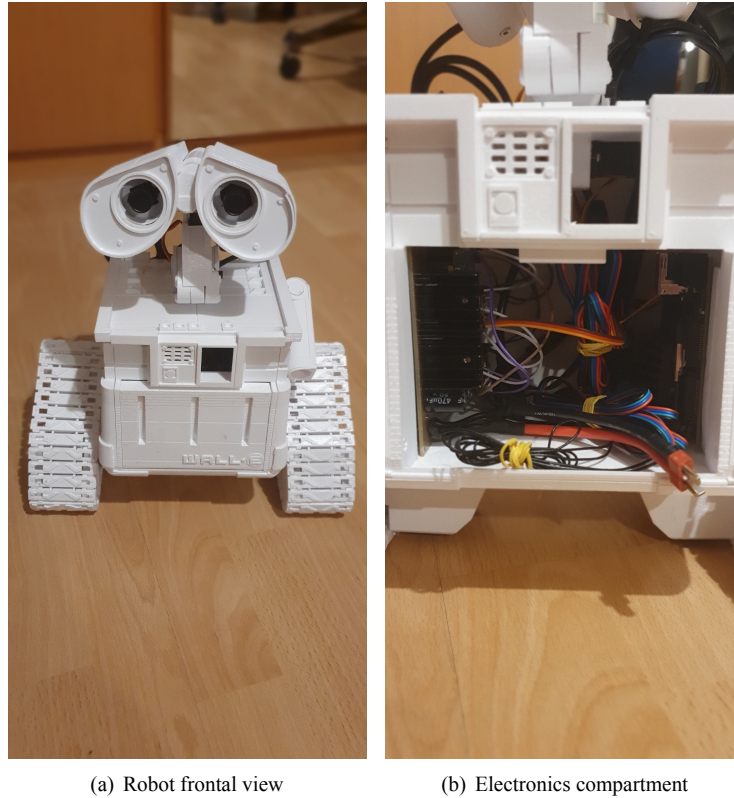
The Jetson Nano microcomputer runs an Ubuntu 18.04 operating system. This provides the robot with a variety of ready to use packages to benefit from. For the video streaming, this paper is using gst-rtsp-server library [3] to build a RTSP server and handle all the incoming connections. Using this library and tuning the GStreamer pipeline for a minimum latency we've been able to achieve around 120 ms of glass-to-glass latency provided the uncongested network. Such latency is noticeable, but definitely not very irritating for shorter sessions.

The microcomputer also listens for a UDP stream containing control input for the actuators. The protocol only contains speeds and directions for the stepper motors and absolute positions for the servo motors at the moment, but could be definitely expanded to encapsulate more data if there would be a need to do so in the future.

The Raspberry Pi Pico microcontroller has a very basic functionality of parsing the control input string and moving the actuators. A timeout for the stepper motors has also been implemented to conserve power when no control input is present.

### 5. VIRTUAL REALITY HEADSET APPLICATION

Oculus Quest 2 virtual reality headset has been chosen for the needs of this paper. Its relatively cheap, but most importantly it is capable of fully wireless operation without any external hardware. Oculus software development kit [4] is intended for its native platform, which is Android, but has integrations even for game engines like Unity or Unreal. Game engines are surely beneficial for quickly creating more advanced scenes, but for the highest flexibility, the native approach is ideal. As this paper only



**Figure 2:** Robotic platform electronics integration

implements basic video rendering and hand controller position querying, the application has been created natively using C and Java programming languages.

To receive from the existing RTSP video stream, pre-built GStreamer binaries has been added to the project. It allows for the same code for video consumption to be shared across all platforms interacting with the proposed robotic platform.

Rendering the received video to the screen is achieved with the OpenGL API and EGL API for interfacing with the native window. The hand controller and head positions are queried via the VrApi which is part of the Oculus mobile software development kit.

Individual tracks of the robotic platform are controlled with respective side joysticks on the hand controllers. The joysticks have analog outputs, so the speed can be set in a very natural way.



**Figure 3:** Quality of the received image

## 6. CONCLUSION

This paper has presented a relatively cheap fully wireless robotic platform with a telepresence control via virtual reality headset. The aim was to try to use only off the shelf parts which has been achieved by

a large part. The robotic system has been designed with a lot of possible future upgrades in mind and leaves some margins for them. The achieved latency of around 120 ms isn't particularly disturbing, but combined with the head movement, the virtual reality can still cause motion sickness to some individuals. One of the ways to combat this would be to tune the regulator that controls the head movement better, or to install a camera that allows for 60 FPS Full HD video, which in turn would double the network usage.

## ACKNOWLEDGMENT

Thanks to the user named Chillibasket for creating and opensourcing the robotic platform 3D models [1]. They have been instrumental for this paper.

## REFERENCES

- [1] CHILLIBASKET. WALL-E Robot Replica. Thingiverse [online]. New York City: MakerBot Industries, c2021 [cit. 2022-03-07]. Available at: <https://www.thingiverse.com/thing:3703555>
- [2] GStreamer. Gstreamer [online]. [cit. 2022-03-07]. Available at: <https://gstreamer.freedesktop.org/>
- [3] Gst-rtsp-server [online]. [cit. 2022-03-07]. Available at: <https://github.com/GStreamer/gst-rtsp-server>
- [4] Oculus SDK [online]. [cit. 2022-03-07]. Available at: <https://developer.oculus.com/downloads/package/oculus-mobile-sdk>



# Eulerian Video Magnification in Surveillance Systems

D. Rícanek<sup>1</sup>, P. Honec<sup>2</sup>

<sup>1</sup>Master's Degree Programme (2), FEEC, Brno University of Technology

<sup>2</sup>Department of Control and Instrumentation, FEEC, Brno University of Technology

E-mail: [xrican02@vutbr.cz](mailto:xrican02@vutbr.cz), [honecp@vutbr.cz](mailto:honecp@vutbr.cz)

**Abstract**—In this article we look at the research behind Eulerian Video Magnification, its beginnings and innovations that lead to its potential use in real time heart rate measurement. The spatial amplitude changes are exaggerated using EVM in conjuncture with face detection using the Viola-Jones algorithm to allow measurement of moving subjects. Fourier transform on the temporal signal of pooled skin pixels then yields heart rate. The algorithm is tested against movement and at different distances.

**Keywords**—Riesz transform, pyramid, phase, Eulerian video magnification, non-contact pulse measurement

## 1. INTRODUCTION

Having a network composed of millions of microscopic sensors, able to cover the surface of an object sounds sensational, yet modern digital cameras could be compared to exactly such a network.

The human eye is not well adapted to recognizing subtle spatio-temporal changes. Even though we're surrounded by an entire invisible world full of useful information. For example, the human head wobbles slightly in rhythm with the beating of the heart. Exaggerating these subtle changes is the domain of a family of algorithms called Eulerian Video Magnification (EVM) invented by Wu et al. in [1] in 2012.

Despite the ten years since its inception, EVM still suffers from major problems with accuracy when large motions occur and as such, isn't suitable to replace traditional monitoring systems e.g. in hospital environments. This work combines the Viola-Jones face detection algorithm and EVM to test its usability in surveillance system for heart rate (HR) detection and measurement in moving subjects.

## 2. PHASE BASED EULERIAN VIDEO MAGNIFICATION

The original linear EVM - which results in the color amplification seen in Fig. 1 - suffers from a low signal to noise ratio (SNR) due to its linear amplification of noise. This can be mitigated by using differences in phases of the image instead of just intensity. This way we can amplify the motion in an image up to four times as much as in the linear method while keeping the noise largely unchanged. Apart from that it also becomes possible to attenuate unwanted small motions such as jittering or movements due to breathing or incorrect stabilization of the measured patch of skin.



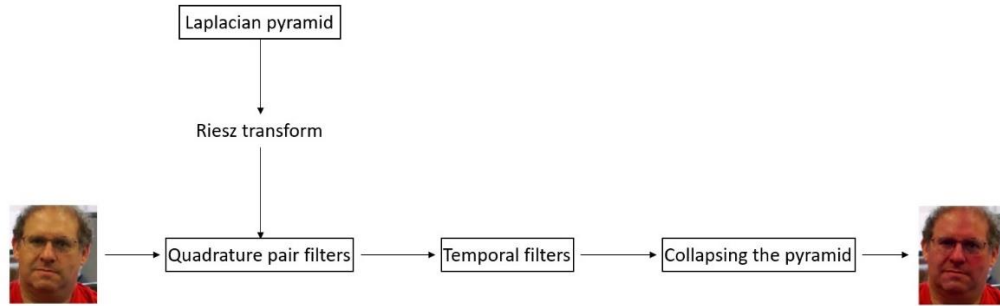
**Figure 1:** Pulsing skin color after exaggeration. Green line shows the section which has been expanded onto the time axis below the picture

However, due to the necessity of calculating quadrature pairs, the overall computational time goes up by a multiple of eight which makes the algorithm unusable in real time applications.

This problem may be dealt with by using a special type of a complex steerable pyramid filter called the Riesz pyramid.

This pyramid is composed of spatial sub bands gained from an invertible version of the Laplacian pyramid used in the original study from Wu et al. in [1], and subsequent approximation of the Riesz transform for all these sub bands.

Since Riesz transform is a special two-dimensional case of the Hilbert transform, its output is a quadrature pair of filters with a phase difference of  $90^\circ$ . The algorithm is visualized in Fig. 2.



**Figure 2:** Individual steps of the Riesz EVM

### 3. SOLUTION

The results from Ramya a Anudev in [2] would suggest that camera choice is largely irrelevant. However, since we're testing the accuracy of measurements on subjects at larger distances, higher resolution had been chosen with Raspberry Pi HQ camera which uses a Sony IMX477 sensor that has a resolution of 12 Mpx. Additionally, the IR filter on this camera can be removed to allow measurements at night.

An oximeter Beurer PO 80 gathering timestamped data was used to obtain ground truth for comparison with HR measured by the algorithm.

The algorithm itself runs in Matlab due to the ease-of-use functions for computer vision. At the beginning the CascadeObjectDetector, which uses Viola-Jones, is applied on the video frame to detect the subjects face. The resulting significant points within the area where a face had been detected are being tracked by the PointTracker method which yields translation and rotation matrices as the frame sequence progresses. This unfortunately takes a significant amount of time and together with separating the frames from the codec makes the algorithm run far from real time.

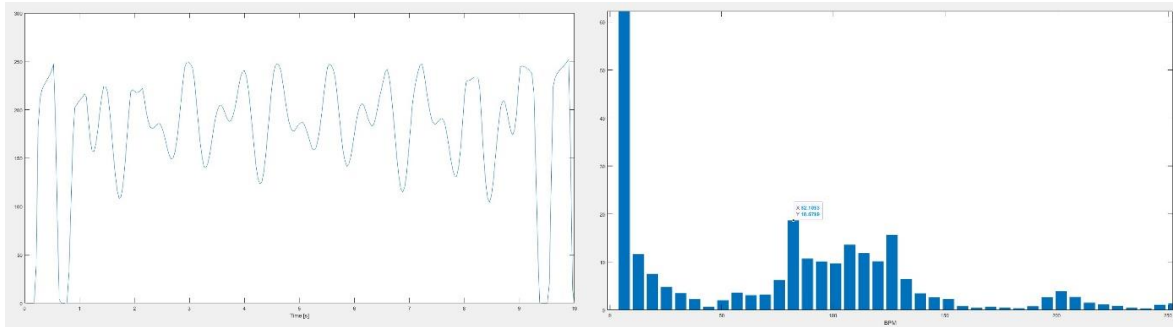
We can then apply these matrices on an arbitrary point on the face – in this case the middle of the forehead – and create a separate video of just this area; effectively stabilizing the video on the area we want to measure as shown in Fig. 3.

Of course, it is nearly impossible to create a perfectly centered video this way due to several reasons. But because the heart's pulse wave travels slow enough we can get away with just applying EVM to this cut out section. The downside is that SNR suffers tremendously as can be seen in Fig. 4 but performing EVM on a smaller video speeds up this part of the algorithm.



**Figure 3:** On the left: Face detected by the Viola-Jones algorithm and the significant points used to

track the face as it moves. On the right: Square area that is to be cut out and composed into another video. The pose of the red square is calculated together with the other significant points



**Figure 4:** On the left: Amplitude of the green channel of the pooled pixels on the subject's forehead over ten seconds. On the right: Fourier transform of this signal with the most significant harmonic being the measured HR

#### 4. RESULTS

The data set was collected under LED lighting powered by a DC power supply to prevent mains frequencies from seeping into the signal via conventional light bulbs. A total of eight measurements had been made to compare the algorithms performance under different frame rates, motion, and different distances. Additionally, two special measurements were made to see if resolution would have any effects and to compare using EVM alone on a mostly still subject.

Unless stated otherwise, the analyzed videos were 640x480 pixels and shot close to the camera (about 1 meter away) at 30 frames per second (fps).

**Table I:** EVM results compared to readings from an oximeter under various conditions

Conditions	Oximeter [BPM]	EVM [BPM]
1280x720 resolution, still subject	75.0	75.8
Still subject, 30 fps	82.0	82.4
Moving subject, 30 fps	81.0	82.7
Still subject, 60 fps	80.0	80.7
Moving subject, 60 fps	68.0	65.0
Still subject, farther away	82.0	63.0
Moving subject, farther away	81.0	65.3
Still subject, EVM applied alone	82	88.7

#### 5. CONCLUSION

Contactless measuring has made a huge step forward since the publication of the pioneering study from MIT in 2012. From experimental biology through defect detection on bridges to monitoring of breathing in sleeping infants, EVM continues to find many use cases.

Thanks to switching from linear EVM to phase based it became possible to exaggerate much smaller movements and with the use of Riesz pyramids this can be done in real time. This has opened the possibility of using EVM in surveillance systems for detecting pathologies and monitoring patients.

Such application of EVM faces some challenges when the subject moves. This article deals with enabling measurements on moving subjects by stabilizing an area of interest on the subject's forehead. This is accomplished by detecting the subject's face using the Viola-Jones object detector in Matlab and passing the extracted, smaller video of just the area on the subject's forehead to EVM for color magnification.

This results in surprisingly accurate readings at close distances even while the subject is moving their head, which wasn't possible when using just EVM on the whole video. In fact, focusing on just a part of the subject's forehead even improves the accuracy and speed of the algorithm when the subject is still.

As was expected there was no difference between the two main frame rates used by today's cameras when measuring HR. At farther distances, the algorithm begins to falter as can be seen in table I. This could be amended by using a higher resolution camera, focusing on skin detection rather than face detection or measuring HR based on the changes in the shape of the subject's head.

As for the real time application. Both EVM and Viola-Jones can be run in real time separately up to a certain resolution. However, at farther distances higher resolution may be required and further testing in this area will be necessary.

We've seen that it's possible to use EVM even on a moving subject, provided they're not too far away. To truly make this algorithm viable in surveillance systems it will be necessary to come up with ways of measuring at farther distances and preventing the signal from being drowned by noise as is almost the case in Fig. 4. Developing a solution for those two problems will be the focus of the future work.

## REFERENCES

- [1] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Frédo Durand, and William Freeman. 2012. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph.* 31, 4, Article 65 (July 2012), 8 pages. doi:<https://doi.org/10.1145/2185520.2185561>
- [2] RAMYA MARIE, S a J ANUDEV. Response Analysis of Eulerian Video Magnification. *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB)*. Cham: Springer International Publishing, 2019, s. 671-678. ISBN 9783030006648. ISSN 2212-9391. doi:[10.1007/978-3-030-00665-5\\_66](https://doi.org/10.1007/978-3-030-00665-5_66)
- [3] Wadhwa, Neal, et al. "Riesz pyramids for fast phase-based video magnification." 2014 IEEE International Conference on Computational Photography (ICCP). IEEE, 2014. <http://dx.doi.org/10.1109/ICCPHOT.2014.6831820>

# Advanced SCARA robot control system

J. Bittner<sup>1</sup>, and T. Beneš<sup>2</sup>

<sup>1</sup>Author FEEC, Brno University of Technology, Czech Republic

<sup>2</sup>Co-author FEEC, Brno University of Technology, Czech Republic

E-mail: [198141@vut.cz](mailto:198141@vut.cz), [151774@vut.cz](mailto:151774@vut.cz)

**Abstract**—This article shows the process of restoring a SCARA type robot, which has had its controller destroyed. The signal cables as well as the power ones were rewired so that it would be able to create new control system for robot arm. New control system is created using Beckhoff devices. For that identification of motors is required. After that PC control is made which will allow further applications for this robot.

**Keywords**—SCARA, Beckhoff, identification, control system

## 1. PREPARING ROBOT ARM

The robot arm is old Epson SCARA robot H554BN with build in robot controller. As the original robot controller was destroyed and repairing it would not be possible, new wiring must be done for the robot to be operable again. Every motor has its own incremental encoder with differential pair signals. Each is powered by 5 V with 2048 counts per revolution. Encoders are connected to Beckhoff servo drives (AX5000 series) and are powered directly from their connectors. Each axis of the root has its own homing sensor which acknowledges its position based on travel time between two parts of the encoder which is read by photomicrosensor (Omron EE-SV3). Those sensors could have been easily destroyed by over-current. Therefore they must have been dismantled and then checked if they are still working properly. By measuring each one of them it is discovered that they need replacement as all of them were fried.

All of the Beckhoff devices as well as power sources are placed in a metallic distribution box that prevents the servo drives from emitting interfering signals all around the robot. The distribution box is designed so that power sources and servo drives are situated in one part of it and PC with input and output cards are in the other part to minimise the interference to between those devices. To easily connect signals from robot to distribution box the DB25 connectors are used. Motor power cords are connected to HART connector which enables disconnecting motors from control box. After the distribution box is prepared as well as connectors on the robot site the system is kind of ready to be programmed.

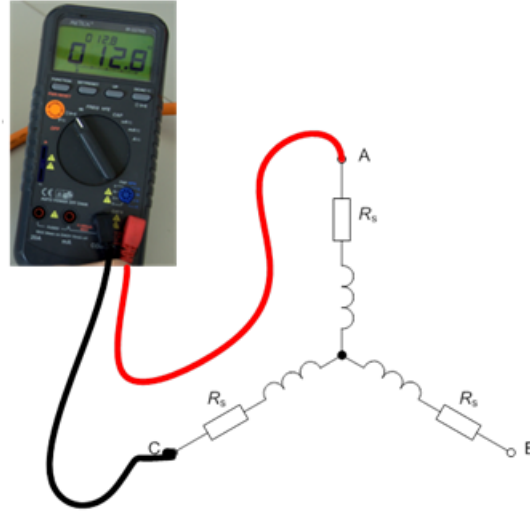
## 2. MOTOR IDENTIFICATION

For SCARA to operate correctly there need to be set the basic parameters of motors. Some of these parameters can be found on the motor itself, those parameters include motors rated current, torque and speed. Other parameters must be measured or estimated. For measurement of parameters the [2] was used as a guideline. In this application note the measurement of the basic motor parameters are described.

First of all the number of motor pole pairs is determined as it defines a ratio between mechanical and electrical quantities. Because a low number of pole pairs is anticipated the easier determination method can be used. This method needs one phase of motor connected to positive potential and the other two phases connected to negative potential of DC power source. The current limit of the power source must be set to maximally 10 % of the rated current of a motor, so that we are able to rotate the shaft manually but the rotor must still be able to align in the stable position. After the rotor is aligned a mark is made on the rotor and then in every stable position of the rotor. Number of rotors stable positions is equal to motors pole pairs.

Next parameter that is determined is stator resistance  $R$ . This parameter is measured with RLC meter between two phases of the motor as is shown on the Figure 1 below. As the measured resistance is between two phases the real resistance of stator is halve of the measured value  $R_m$

$$R = \frac{1}{2} \cdot R_m \quad (1)$$



**Figure 1:** Motor winding measurement

Third parameter that is being measured is inductance of a motors winding. The inductance is measured the same way as resistance, which is series resistance of two motor phases and the real inductance of one phase is half of measured value.

$$L = \frac{1}{2} \cdot L_s \quad (2)$$

Both resistance and inductance measurement are made with RLC meter, which has measuring voltage set to 1000 mV at frequency of 1 kHz. All three possible combinations of motor phases are measured once and the final value which will be used for each parameter is average of those three values.

Last measured parameter is back-EMF constant which gives a ratio between voltage and angular electrical frequency (speed). Three phase measurement is used to determine this constant. This method is used because the neutral point is not accessible. Each of the three motor phases is connected to one voltage probe of oscilloscope and the artificial neutral point is created by connecting the probes together. Then the motor is spin by a hand drill at a constant speed. Spinning of the motor creates voltage at all of the phases and from the peak to peak value of generated voltage  $U_{pp}$  and period of this voltage  $T_{el}$  the back-EMF constant  $k_e$  can be calculate using equation 3.

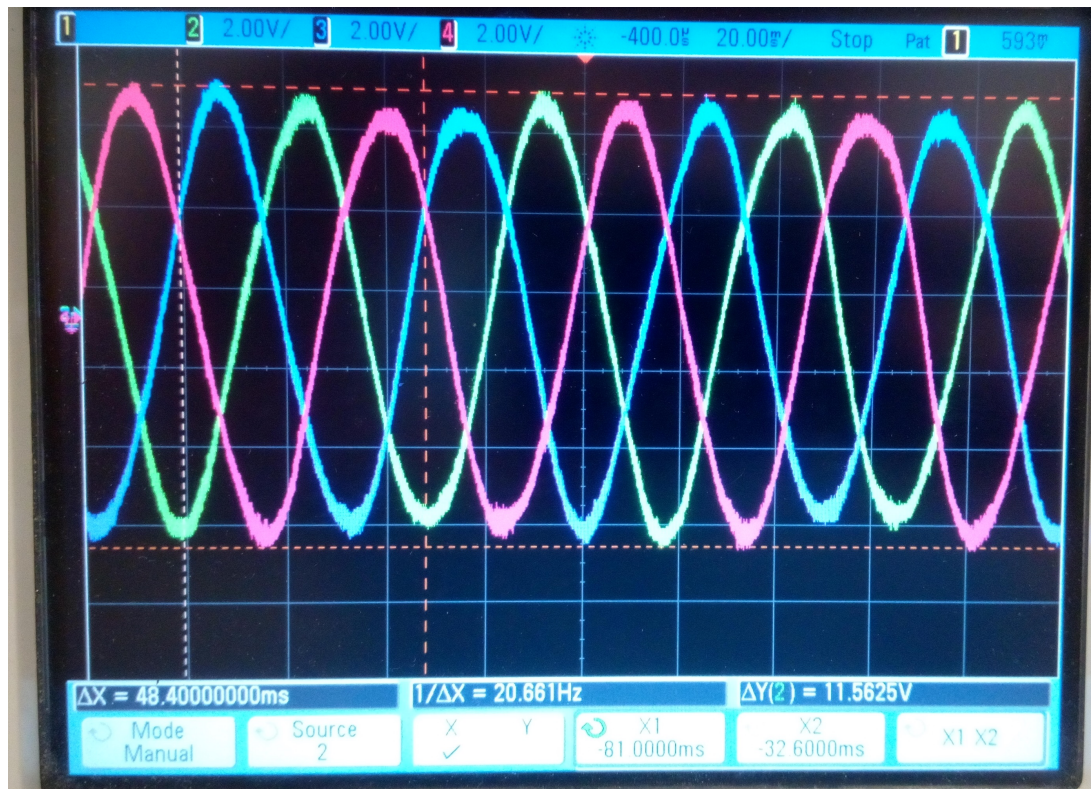
$$k_e = \frac{U_{pp}}{2 \cdot \omega_{el}} = \frac{U_{pp} \cdot T_{el}}{4\pi} \quad (3)$$

From four motors that are responsible for robots movement only two are needed to be identified. That is caused by the fact that three motors are of the same type and therefore should have the same parameters. All of the measured parameters can be found in Table I.

**Table I:** Measured motor parameters

Parameter	First motor	Other motors
$R [\Omega]$	3.35	9.30
$L [\text{mH}]$	7.98	13.14
$T_{el} [\text{ms}]$	58.40	55.20
$U_{pp} [\text{V}]$	11.56	11.10
$k_e \left[ \frac{\text{mV} \cdot \text{s}}{\text{rad}} \right]$	53.72	48.76





**Figure 2:** Back-EMF constant measurement

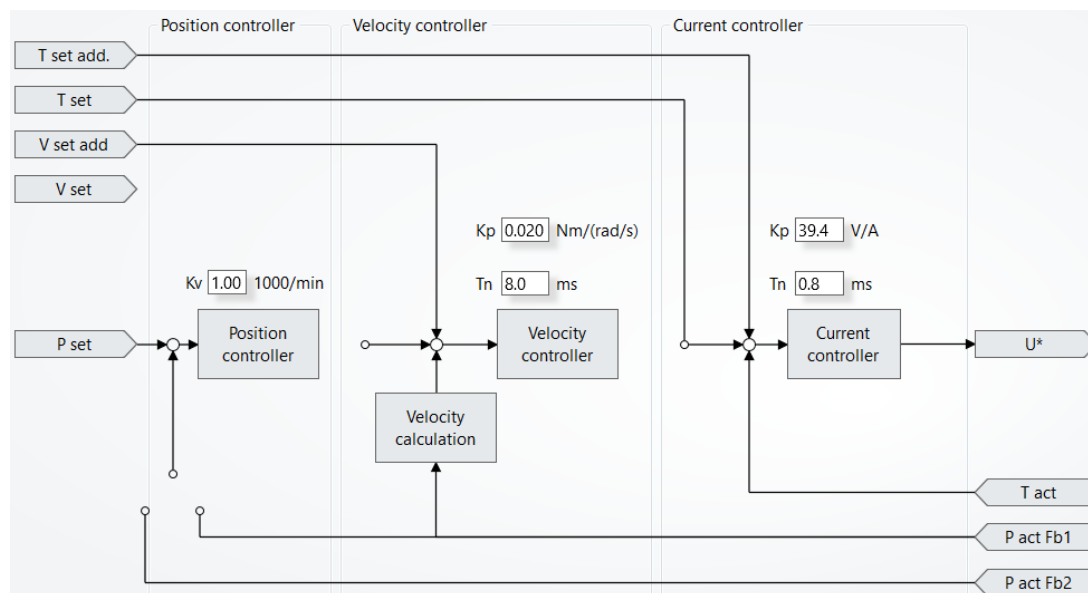
### 3. SETTING MOTORS IN PROJECT

Beckhoff has its own special project for motor configuration that is called *Drive manager 2*. In this project are created parts for each servo drive connected and all of outputs within those drives. In those parts correct power input must be set first, there are predefined variants for European and American power grids, but for motors used in this project this has to be set manually as the drives are powered from 200 V power source. Then for each output channel the right motor must be set. Again there are predefined motors, or the motors can be scanned if they are compatible with this function of drive manager. In our case the motors must be set manually with parameters that were measured and that can be found on motor itself.

As those parameters are static or rated values, they are not enough for drive manager to create characteristic curve of motor. This curve serves as base for control unit to calculate the needed power output for given motor. Therefore further parameters must be estimated. First parameters that are needed are standstill torque and current those are estimated as one third of the rated current and torque. Second parameters are the peak current and torque of the motor. Peak current is set as double the rated current and peak torque as triple the rated torque of motor. Then maximal allowed voltage must be set, this parameter is based on the fact that the motors are running fine on normal grid voltage and therefore are set a little bit higher than that to 250 V and maximum DC-link voltage allowed is set to 350 V. The last two parameters that are needed to set a motor are motor inertia, which was taken from data sheet of a similar motor and set to  $5 \cdot 10^{-5} \text{ kg} \cdot \text{m}^2$  and motor thermal time constant which is set to 10 minutes.

Now that our project knows all it needs about the motors it is time to set motor controller. Controller is divided in two parts as is usual for motor controllers: velocity controller and current controller as is shown on Figure 3. From given parameters the drive manager creates the characteristic curve of motor as is written above and based on those parameters and the curve sets current controllers integration time and gain. The velocity controller is set to default values that are safe for running motor without load. In this controller the integration time and gain are the main parameters. Based on Beckhoff guidelines the controller should be set using Ziegler-Nichols method. This method was tested but the results were not as good as they should be. The motors were overexcited and were really loud. Therefore the velocity regulators are set experimentally by iterating the velocity gain. Integration time is left unchanged on

8 ms. When the motors are running with error not exiting one tenth of requested value, the controller is marked as correctly set for our purpose. Even though the motors are not overexcited from velocity controller at some positions there still is audible squealing. This is the result of current controller trying to set the right current to make the motor stand at given position. To suppress these sounds and lower the current flowing through motor, basic first order low pass filter is set on the input of current controller. This filter does not affect the function of controller it just smooths the output of velocity controller in a way that when the motor is standing still the current is not changing too fast with too low amplitude of input value.



**Figure 3:** Motor controller overview

#### 4. CONCLUSION

The old non-functional SCARA robot was rewired. This was done so that new robot controller can be made with Beckhoff devices and the robot could be reprogrammed. After new electrical wiring was complete, motors must have been identified as there was no available documentation for them. In identification process the basic motor constants and parameters were determined. For full functioning of control system further parameters were estimated. After creating a model of motor in control system motor controller must have been set. This contained two parts velocity and current controller. Current controller was set automatically based on given model of motor. The parameters for velocity controller were determined experimentally. First order low pass filter was set to filter the requested value from velocity controller to current controller. All of those processes and settings were needed to allow the programming of control system for this SCARA robot.

#### ACKNOWLEDGMENT

The completion of this paper was made possible by the grants:

No. FEKT-S-20-6205 - "Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0";

#### REFERENCES

- [1] SEIKO EPSON CORPORATION. *BN TYPE: SCARA ROBOT* Rev. 5
- [2] BOBEK, Viktor. FREESCALE SEMICONDUCTOR. *PMSM Electrical Parameters Measurement* Rev. 0, 02/2013. Dostupné z URL: <<https://www.physicsforums.com/attachments/freescale-an4680-pdf.79167/>>
- [3] BECKHOFF. *System manual AX5200 Servo Drives* Version: 2.7, 31/8/2021

Dostupné z URL: <<https://www.beckhoff.com/cs-cz/products/motion/servo-drives/ax5000-digital-compact-servo-drives/ax5203.html>>

- [4] SPONG, Mark W., Seth HUTCHINSON a M. VIDYASAGAR. *Robot modeling and control*. Hoboken: John Wiley, 2006. ISBN 0-471-64990-2.

# Biometrics using face recognition

M. Koupil<sup>1</sup> and J. Odstrčilík<sup>1</sup>

<sup>1</sup>Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication,  
Brno University of Technology, Czech Republic

E-mail: xkoupi01@vutbr.cz, odstcilik@vut.cz

**Abstract**—This paper is focused on a face detection in a picture and subsequent recognition of the face in its respective database. Face detection had been implemented using Viola-Jones algorithm. To recognize the face afterwards, PCA had been used. Implemented algorithm had been tested on freely accessible biometric databases.

**Keywords**—Biometrics, face localization, Viola-Jones algorithm, face recognition, PCA

## 1. INTRODUCTION

Among biometric systems, ID systems, that are based on working with face features, dominate. Compared to other metrics, they are significantly more popular for various reasons: Human faces are generally unique and universal [1], also willingness of people to provide a picture of themselves is higher compared to other metrics. It finds its use in criminalistics (e.g. database of fugitives) and it is getting more common in commercial applications (e.g. authorization in mobile devices).

Biometrics of face generally consists of two parts: Localization - finding the face within a picture, and Recognition – its identification in given database. There are many various methods for a face localization and recognition, and even though there is a general scheme, individual methods may vary quite significantly.

The goal of this paper is to implement algorithm performing face biometrics using freely accessible picture databases as input data.

## 2. BIOMETRIC DATABASES

For the testing of implemented algorithm for face localization and recognition, images from free picture databases have been used [2],[3]. Algorithm has been tested on five sets of pictures from these databases. For purpose of this algorithm, only pictures that were taken frontally were used. Individual datasets possess high intra-class variability (people on pictures have different facial expressions, wear different hats, have additional accessory...).



**Figure 1:** An example of training pictures from one of the datasets. High intra-class variability is clearly visible.

## 3. METHODS

### 3.1 Viola-Jones Algorithm

For face localization, Viola-Jones Algorithm has been used [4]. By its nature, it's a method based on

appearance. It uses detectors of characteristic features of the face – Haar filters. The filters scan through the picture in various scales and marks features, that could be part of the face. Classification algorithm *AdaBoost* is used to select the best filters. *AdaBoost* is a linear combination of a group of weak classifiers (=simple Haar filters) which are iteratively turned into strong classifier. To avoid applying all the filters on a window, cascade of classifiers is presented. Several filters are put together in different stages of classifiers and scan sub-windows. If a sub-window fails to include part of the face, it is discarded. Sub-window which will pass all stages is considered to be a part of the face.

For implementation of Viola-Jones Algorithm, library *openCV* has been used. For face detection in a picture, pre-learned model [5] was used.

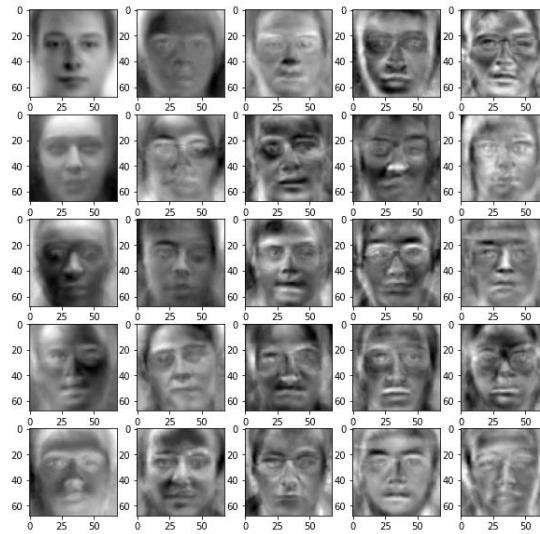
### 3.2. PCA method

For face recognition, principal component analysis [6] has been used. It is a holistic method, meaning it perceives the face as a whole. It converts the dataset to the new space called *eigenspace*, where individual pictures would be ordered by variance from the mean picture. The mean picture is calculated via equation:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i, \quad (1)$$

where  $\bar{X}$  denotes mean picture, which is counted as a sum of input vectors divided by the number of them.

For training, libraries for machine learning *scikit-learn* [7] was used. From the order of pictures in the new space, we can see that the pictures are ordered by importance for given dataset.



**Figure 2:** The most important 25 pictures in eigenspace of one of the datasets.

In terms of dimensionality reduction, it is possible to heuristically determine number of most important pictures for further calculations. If we were to perform reconstruction of any original image, we would lose minor details, with the price being reducing dimensionality, therefore reducing computational expense. To assign input testing image to the training database (=to recognize it), we need to calculate vector of weights for each training image. Weight vector of input testing image will be compared to weight vectors of every individual image of the training database. The most similar one will be declared as a match.

## 4. RESULTS

As mentioned before, the algorithm has been tested on five different sets of pictures. In Dataset 1, there

are 27 white people, each one has 10 training pictures. In dataset 2, there are 10 Asian people, each one has 13 training pictures. In datasets 3,4 and 5 there are 20 people of various races, each one has 15 training pictures. Intra-class variability in the last 3 databases is higher than in the first 2. While in the first 2 databases hairstyles and facial expressions changed, in the latter 3 also additional accessories, such as glasses or hats, appeared.

Given the fact, that face localization, *frontal face* Haar filters have been used, some of the detections have not been performed. One of the reasons of a failed localization was that part of the face was missing on the picture, another reason might be capturing the face from inconvenient angle.



**Figure 3:** Examples of pictures inconvenient for localization

<i>Dataset</i>	<i>Rate of successful localization</i>	<i>Percentual rate</i>
<i>Dataset 1</i>	<i>270/270</i>	<i>100%</i>
<i>Dataset 2</i>	<i>129/130</i>	<i>99,2%</i>
<i>Dataset 3</i>	<i>300/300</i>	<i>100%</i>
<i>Dataset 4</i>	<i>297/300</i>	<i>99%</i>
<i>Dataset 5</i>	<i>298/300</i>	<i>99,3%</i>

**Chart 1:** Rate of successful face localization.

Pictures, at which the localization was not performed correctly, were omitted from the database. Because it was a very small number of pictures, it did not interfere with next steps of the algorithm.

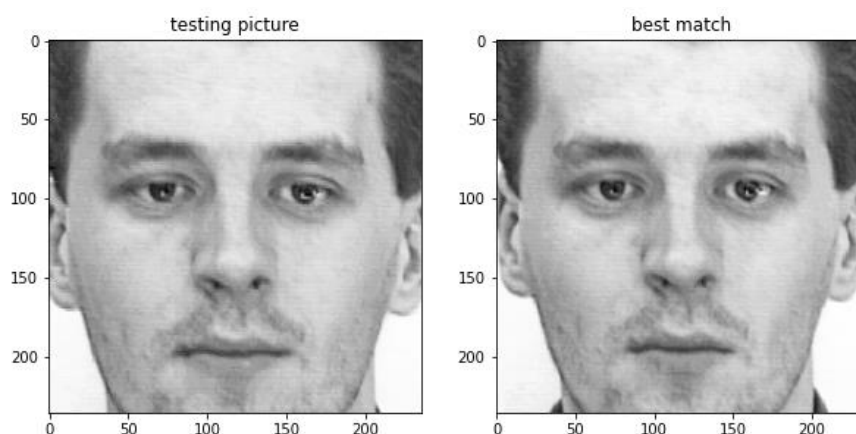
After localization of a face on each picture of the datasets, PCA was performed so the testing pictures can be assigned to their respective datasets. Recognition success of testing pictures in terms of their training datasets is stated in the chart:

<i>Dataset</i>	<i>Rate of successful recognition</i>	<i>Percentual rate</i>
<i>Dataset 1</i>	<i>20/27</i>	<i>74%</i>
<i>Dataset 2</i>	<i>9/10</i>	<i>90%</i>
<i>Dataset 3</i>	<i>20/20</i>	<i>100%</i>
<i>Dataset 4</i>	<i>18/20</i>	<i>90%</i>
<i>Dataset 5</i>	<i>20/20</i>	<i>100%</i>

**Chart 2:** Rate of successful face recognition.

On following picture there is an example of successful recognition:





**Figure 4:** An example of successful face recognition

## 5. CONCLUSION

This paper is focused on biometrics using face recognition. In practice, the biometrics was performed in two steps: face localization and face recognition. Viola-Jones Algorithm was used for face localization and PCA was used for face recognition.

Algorithm was tested on five different picture datasets. At face localization, average success rate was 99,5 % and at face recognition, average success rate was 90,8 %.

Lowest success rate of face recognition was registered when Dataset 1 was tested. There might be more reasons why. One reason could be that there weren't enough training images for the algorithm to learn properly. Another possible reason is that because the individuals shared the same ethnicity, there were more similarities between them, so there is a bigger chance of wrong registration.

Algorithm was also tested on all the databases combined, the PCA method was then applied to pictures of all the datasets. In this case, success rate was 89,6 %.

## REFERENCES

- [1] DRAHANSKÝ, Martin a Filip ORSÁG. Biometrie. Brno, 2011.
- [2] Face Databases [online]. [cit. 2021-12-08]. Dostupné z: [http://web.mit.edu/emeyers/www/face\\_databases.html](http://web.mit.edu/emeyers/www/face_databases.html)
- [3] SPACEK, Libor. Libor Spacek's Facial Images Databases [online]. 2009 [cit. 2021-12-08]. Dostupné z: <https://cmp.felk.cvut.cz/~spacelib/faces/>
- [4] VIOLA, Paul a Michael JONES. Rapid Object Detection using a Boosted Cascade of Simple Features [online]. 2001, 9 [cit. 2021-11-22]. Dostupné z: <https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf>
- [5] OpenCV: Cascade Classifier [online]. [cit. 2021-12-07]. Dostupné z: [https://docs.opencv.org/3.4/db/d28/tutorial\\_cascade\\_classifier.html](https://docs.opencv.org/3.4/db/d28/tutorial_cascade_classifier.html)
- [6] SAXENA, Pawan. Face Recognition Using Eigenfaces (PCA Algorithm) [online]. 2021 [cit. 2021-11-30]. Dostupné z: <https://www.geeksforgeeks.org/ml-face-recognition-using-eigenfaces-pca-algorithm/>
- [7] Pedregosa et al. Scikit learn [online]. 2011 [cit. 2021-12-09]. Dostupné z: <https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>

# Innovation of the laboratory exercises for course Embedded Systems

Matej Pončák<sup>1</sup> and Petr Petyovský<sup>2</sup>

<sup>1</sup>Master Degree Programme (2), Brno University of Technology

<sup>2</sup>Assistant professor, Brno University of Technology

E-mail: [xponca00@vut.cz](mailto:xponca00@vut.cz), [petyovsky@vut.cz](mailto:petyovsky@vut.cz)

**Abstract**—The constant market development of the embedded systems requires an adaptation to this fact. Raspberry Pi has released the Raspberry Pi Pico development board with a 32-bit microcontroller ARM Cortex-M0+, which has the potential to use the board as an innovation in teaching embedded systems. The paper describes this board and its properties, the course Embedded Systems and the possibilities of its innovation, and then describes the proposed HW platform and demonstrates its usage using appropriately selected laboratory exercises.

**Keywords**—Raspberry Pi Pico, RP2040, embedded systems, innovation, education kit, laboratory exercises

## 1. INTRODUCTION

In early 2021, Raspberry Pi Foundation introduced its first 32-bit microcontroller RP2040 along with the Raspberry Pi Pico development board. Its architecture allows use in the simplest applications up to machine learning or video output to a monitor. One of the main goals of the foundation is education in the field of computers, which creates a target group for students, for whom, however, the price of the equipment is often an obstacle. The foundation therefore designs its products at an affordable price. This was also reflected in the price of the Raspberry Pi Pico, which sells for \$4.

At the BUT FEEC in Brno, Czech Republic, the course Embedded Systems and Microprocessors is taught in the bachelor's study programme Automation and Measurement. The purpose of this course is to get acquainted with the basic principles of microprocessors and embedded systems. In the laboratory exercises of this course, we work with an 8-bit microcontroller from the HCS08 family from the NXP company. In order for students to be as adapted to the market as possible after completing this programme, the course requires the innovation of teaching aids, thus creating the potential for the use of the Raspberry Pi Pico development board.

At the beginning, the reader gets acquainted with the microcontroller RP2040. Furthermore, the content of the course Embedded Systems and Microprocessors will be explained and a set of tasks for laboratory exercises will be designed so that their educational character will guide the student through the basic principles of embedded systems and explain the work with the development board.

In order for students to be able to use the development board together with external peripherals, it is necessary to design the concept of the hardware platform. In the following text, this platform will be designed and subsequently tested for selected laboratory exercises.

## 2. MICROCONTROLLER RP2040

The most important part of development board is microcontroller RP2040 from Raspberry Pi. The chip is surprisingly powerful, but also very cheap with a dual ARM Cortex-M0+ architecture processor, which is the most energy-efficient ARM processor.

Key features of microcontroller RP2040 [2]:

- Dual ARM Cortex-M0+ @ 133MHz
- 264kB on-chip SRAM in six independent banks
- Support for up to 16MB of off-chip Flash memory via dedicated QSPI bus

- DMA controller
- Fully-connected AHB crossbar
- Interpolator and integer divider peripherals
- On-chip programmable LDO to generate core voltage
- 2 on-chip PLLs to generate USB and core clocks
- 30 GPIO pins, 4 of which can be used as analogue inputs
- Peripherals:
  - 2x UARTs
  - 2x SPI controllers
  - 2x I<sup>2</sup>C controllers
  - 16 PWM channels
  - USB 1.1 controller and PHY, with host and device support
  - 8 PIO state machines

### 3. INNOVATION OF THE LABORATORY EXERCISES

Currently, the development board TWR-S08LH64 is used in the laboratory exercises of the course. Its core is an 8-bit microcontroller from the HCS08 family from the NXP, specifically the MC9S08LH64 model. Typical aspects for this microcontrollers family are: fast memory access, a small number of registers and a CISC instruction set. There is also support for 32 possible interrupt sources, chip-level address decoding, and a hardware multiplier and divider. Its main modules include an LCD display interface, 16-bit A/D converter, I<sup>2</sup>C, UART, SPI, analog comparator, timer, time of day module, keyboard interrupt module and a total of up to 39 GPIOs [1].

The development board communicates with the PC via the USB interface. The basic interaction with the user is ensured by 4 buttons, 4 LEDs, a piezoelectric speaker and a potentiometer. The board is also equipped with a light sensor in the form of a phototransistor and a three-axis accelerometer with analog outputs. An LCD display that is directly connected to the microcontroller is used as the display unit.

In order for students to be as adapted to the market as possible after completing this programme, the course requires the innovation of teaching aids, and relevant laboratory tasks. The HCS08 family of microcontrollers has an 8-bit architecture but nowadays, multi-core microcontrollers with a 32-bit architecture are commonly used. In addition, they have several peripherals such as DMA, USB or PLL. Also, the ARM architecture, which has a significant market share, is being used more and more. HCS08 microcontrollers find their use in many areas, but due to the mentioned facts, these microcontrollers do not represent the current development in their field, which creates the potential for the use of the Raspberry Pi Pico development board.

#### 3.1. Laboratory exercises for Raspberry Pi Pico development kit

With the use of the development kit in the teaching of the course Embedded Systems and Microprocessors, there is an opportunity to change the assignment of laboratory exercises so that they make the most of the development board's peripherals. In this section, these exercises will be designed. However, these exercises must be in accordance with the content of the course and, in order to be feasible in laboratory exercises, they must be done in one laboratory exercise, which lasts three hours.

**Instruction set:** Students will program finding the minimum and maximum number in the array of numbers using the ARMv6-M instruction set. Function calling and value returning will be implemented in accordance with C language calling conventions.

**GPIOs:** Getting started with the GPIO initialization, such as setting inputs or outputs, enabling the internal pull-up or pull-down resistor, or selecting a current limit for the output. Exercise ends with reading values from the inputs and writing it to outputs.

**Interrupts:** Creating and processing interrupts. Their operation will be demonstrated on the timer

peripheral, which generates an interrupt with a certain period of time. It can control, for example, the logic level of the LED. Connecting a speaker to the output and lowering the timer period will generate a tone of the appropriate frequency.

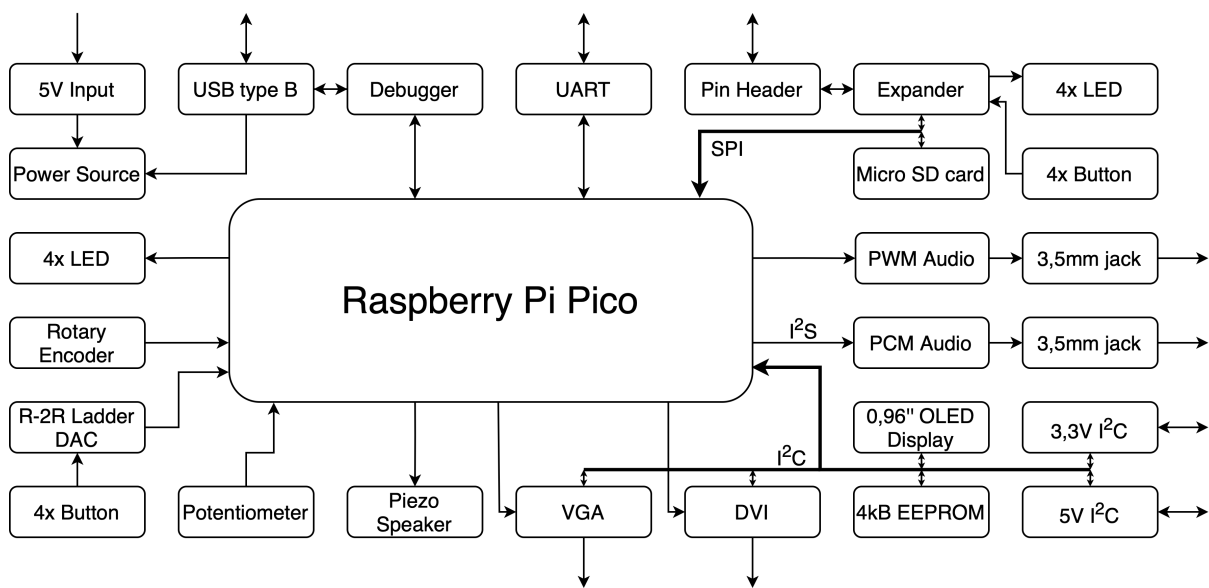
**A/D converter:** In this exercise, students get acquainted with the initialization of the A/D converter and then reading analog values from the input where the potentiometer is connected. The four LEDs will light up one by one according to the input voltage.

**UART communication:** This exercise demonstrates usage of UART. First, the port is properly configured and then receives text messages. All received letters are changed to uppercase by program and they are sent back.

**SPI communication:** In this exercise, students will get acquainted with the basic concept of SPI interface using an I/O expander. They will program the communication with this integrated circuit and create a program that writes the status of the buttons to the LEDs.

#### 4. HW PLATFORM DESIGN

The main purpose of this paper is to create a development kit that will contain the necessary components for the implementation of laboratory exercises. In addition, the development kit will include audio and video output. The kit must include a debugger with support for debugging functions so that the program can be easily debugged. The whole electrical connection is divided into several blocks (Fig. 1), which are described below. The most important part of the design is the Raspberry Pi Pico development board itself, which will connect to the kit modularly via pin headers.



**Figure 1:** HW platform block diagram.

**Debugger:** Another RP2040 microcontroller is used as a debugger. The UART port of the development board is also connected to the debugger to communicate with the board via a serial port via the same USB interface, which is used for power supply, program uploading and debugging. It is necessary to upload the PicoProbe program into the flash memory for debugging.

**Basic user interface:** Four LEDs are used as basic display elements. Four buttons connected to one analog input using an R-2R resistive ladder network are used as input devices, which allows reading several buttons at once. A rotary encoder and a rotary potentiometer are used as additional input devices.

**UART interface:** Communication on the UART interface is signaled by LEDs. For communication with other devices via the UART interface, there is a port on the kit.

**SPI devices:** The I/O expander MCP23S08 is used to increase number of GPIOs. Four buttons and four LEDs are connected to the expander. Furthermore, a micro-SD card operating in 1-bit mode is also connected to the SPI interface.

**I<sup>2</sup>C devices:** The development kit contains several I<sup>2</sup>C devices. A 0.96" OLED display with a resolution of 128x64 pixels is used for basic data display. The built-in SSD1306 driver takes care of display and communication with the I<sup>2</sup>C bus. A 4kB EEPROM memory is also on the development kit to store configuration data. There are two more ports for connecting other devices to the I<sup>2</sup>C bus.

**Audio output:** The simplest audio output is realized by a piezoelectric speaker. The second option generates a stereo sound using a PWM signal, which is connected to a 3.5 mm jack via a filter. The third option is the audio output generated by PCM modulation using the integrated circuit PCM5101A communicating via the I<sup>2</sup>S interface. Its output is also connected to a 3.5 mm jack.

**Video output:** The analog video output to the VGA interface has an 8-bit R3G3B2 color resolution [3]. Digital video output is realized via DVI interface [4]. HDMI connector is used as an output for DVI because HDMI is backward compatible with the DVI interface.

## 5. HW PLATFORM TESTING

The testing of the designed platform was performed on the solderless breadboard. The basic circuitry that allows testing is the debugger, which is made by another Raspberry Pi Pico board with the PicoProbe program. All the necessary components for testing the individual functional units are connected to the board and the appropriate program is loaded into the tested development board. For testing purposes, a development board with basic electronic components was used, which the author had from the previous studies.

## 6. CONCLUSION

In this paper the documentation and possibilities of the Raspberry Pi Pico development board were described. In order to be able to appropriately design new laboratory exercises for the course Embedded Systems and Microprocessors, the course curriculum and the assignments of the current laboratory exercises had to be studied.

The main goal is to design a HW platform that will be used in the laboratory exercises of the mentioned course. In order to make the platform as universal as possible, it goes far beyond being used in laboratory exercises. In addition to the basic user interface in the form of LEDs, buttons, encoder and potentiometer, it also contains a modern OLED display, PWM audio and I<sup>2</sup>S output, video output via analog VGA and digital DVI interface, micro-SD card slot and also provides standard UART, SPI and I<sup>2</sup>C interfaces, which are commonly used in embedded systems. Debugging and uploading the program is provided by another microcontroller RP2040.

The next step is creation of a printed circuit board with mounted components and all the necessary documentation for production. For the created board, assignments of laboratory exercises for teaching will be created, as well as their solutions.

This universally designed board will be able to be used not only in the laboratory exercises of the course Embedded Systems and Microprocessors, but also in another follow-up course Embedded Systems, or even in the course PC Subsystems due to its VGA, DVI and USB interface. In addition, thanks to the availability of all production documentation, anyone will be able to produce the board and then use it for their own needs. The board can also serve as a development kit for assigning semestral projects or other final theses.

## REFERENCES

- [1] J. Mikeska, *Laboratory tutorials for Freescale MC9S08LH microcontrollers*, BUT FEEC Brno, 2013, [ref. 13. 03. 2022]. Available at: <<http://hdl.handle.net/11012/27989>>.
- [2] Raspberry Pi, [online datasheet], *RP2040 Datasheet: A microcontroller by Raspberry Pi*. Rev. 1.7.1, 11/2021, [ref. 13. 03. 2022]. Available at: <<https://datasheets.raspberrypi.com/rp2040/rp2040-datasheet.pdf>>.
- [3] M. Němeček, *PicoQVGA - Minimalistic QVGA Display on Raspberry Pico*, [ref. 13. 03. 2022]. Available at: <[http://www.breatharian.eu/hw/picoqvga/index\\_en.html](http://www.breatharian.eu/hw/picoqvga/index_en.html)>.
- [4] *Bitbanged DVI on the RP2040 Microcontroller*, [ref. 13. 03. 2022]. Available at: <<https://github.com/Wren6991/PicoDVI>>.



# ABB v Brně

## Lepší svět začíná s vámi

Jsme součástí globálního technologického lídra působícího v oblastech elektrotechniky, robotiky, automatizace a pohonů. Z Brna dodáváme rozváděče vysokého i nízkého napětí, modulární a digitální systémy, přístrojové transformátory a senzory. Tyto produkty míří do nejnáročnějších projektů ve více než 100 zemích světa. Najdete je mimo jiné v nejvyšší budově světa Burž Chalífa, v londýnském metru nebo v datovém centru Facebooku.



Přibližně 2000  
zaměstnanců



Tržby 6,5 mld. Kč



Významné centrum  
výzkumu a vývoje



Nejnáročnější  
projekty ve 100  
zemích světa



4 výrobní závody,  
centrum digitálních  
řešení, R&D a servis

### Studentům a absolventům nabízíme:

- Placené brigády
- Odborné praxe
- Stáže
- Trainee program
- Vedení bakalářské nebo diplomové práce
- Hlavní pracovní poměr po ukončení studia

### Kontaktujte naši náborářku:

Kristýna Podobová  
kristyna.podobova@cz.abb.com  
+420 739 552 428

**Virtuální prohlídka**





## Naše jednotky

### Rozváděče vysokého napětí

Výroba rozváděčů vysokého napětí se v našem závodě těší dlouhé tradici a pyšní se statusem největšího závodu svého druhu v Evropě. Co do počtu zaměstnanců, tak i plochy výroby, se jedná o největší jednotku v rámci brněnského závodu. Své produkty dodává do přibližně 100 zemí světa.

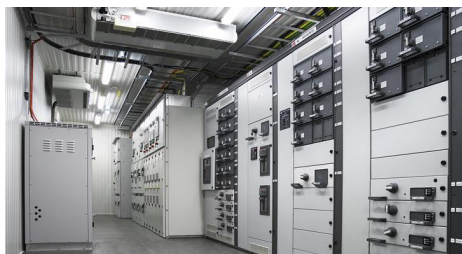


### Rozváděče nízkého napětí

Rozváděče nízkého napětí fungují v režimu tzv. virtuálního závodu a sdílí svou kapacitu se závodem v polské Bielsko-Białe. Opět se jedná o jeden z největších závodů svého druhu v Evropě. Jeho produkty jsou dodávány do přibližně 50 zemí světa.

### Přístrojové transformátory a senzory

Jednotka přístrojových transformátorů a senzorů je 2. největší v rámci brněnského závodu. Zároveň se jedná o největší závod svého druhu na celém světě. Vůbec první transformátor byl v závodě vyroben již v roce 1919 a jejich výroba je tak založena na více než stoletém know-how.

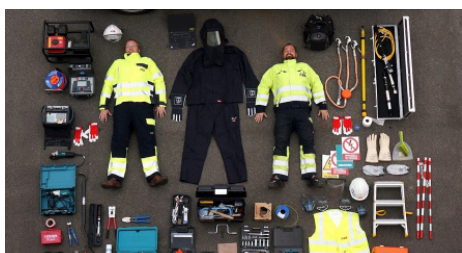
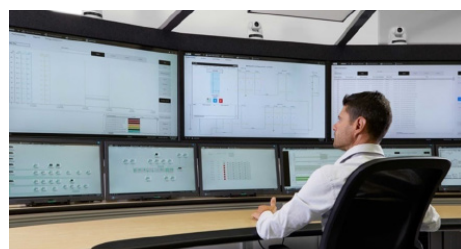


### Modulární systémy

Přestože se jedná o mladou jednotku, má už za sebou několik významných milníků. Své modulární rozvodny, dodávané obvykle v kontejnerovém řešení, dodává do projektů napříč odvětvími. Nejvýznamnější z nich jsou data centra a projekty v oblasti těžby a distribuce ropy a zemního plynu.

### Centrum digitálních řešení

Cílem jednotky Digital Solution Center je realizace inteligentních řešení v oblasti monitoringu, sběru a archivace dat a hlavně vzdáleného řízení distribuce energie. Zákazníkům nabízí řešení na míru, flexibilitu a vysokou přidanou hodnotu.



### Servis přístrojů vysokého a nízkého napětí

Jednotka servisu spolupracuje se zákazníky v rámci oprav, údržby a vylepšování jejich zařízení po celou dobu jejich životního cyklu. U svých zákazníků v přibližně 100 zemích světa stráví technici asi 40 tisíc hodin ročně. Jednotka využívá nejmodernější nástroje včetně tréninkového centra pro zákazníky nebo rozšířené reality.

### Technologické centrum

Technologické centrum se zabývá výzkumem a vývojem produktů. V rámci ABB se celosvětově jedná o velmi významnou výzkumnou jednotku. K dispozici má špičkově vybavenou laboratoř umožňující provádění nejrůznějších simulací a výrobu a testování prototypů.



# Analysis of critical battery parameters for space applications

Michal Šimek<sup>1</sup>

<sup>1</sup>Brno university of Technology, Czech Republic

E-mail: [xsimek33@vutbr.cz](mailto:xsimek33@vutbr.cz)

**Abstract** – This thesis deals with the analysis of lithium-ion batteries from the point of view of temperature. The aim is to find out at which temperature and C value the lithium plating process starts in the battery by means of non-destructive methods. This process is undesirable for the functioning of lithium-ion batteries as it causes capacity reduction and cell degradation. In this study, commercial cells from GOMspace were selected as the test batteries. GOMspace will be able to use the results of the research to optimize its devices while avoiding lithium plating, which will increase battery life.

**Key words** – Li-ion, batteries, temperature, space, lithium plating, degradation

## 1. INTRODUCTION

In today's modern world, batteries are an essential element of most electronic devices, which are key to most industries. The most obvious is the group of mobile phones and smart accessories, which nowadays sets the pace for the development of both batteries and other technologies. Batteries are also used in the burgeoning field of electromobility, for example, or in the vast renewable energy and energy storage sector. Last but not least, batteries are used for space applications. For such batteries, it is very much their parameters to ensure that even in extreme conditions there are no problems in their functioning.

In all these sectors, there is a need to develop and test new types of batteries all the time that are more environmentally friendly and safe, but at the same time have a high capacity, high lifetime and low weight, and in the best case low cost. The development of batteries can bring us a greener society, simplicity in the use of electronic devices and a better understanding of the universe.

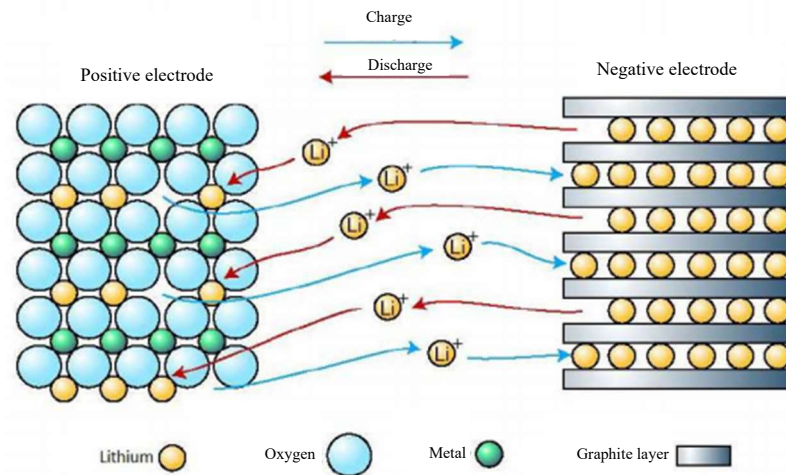
My thesis deals with lithium-ion batteries for space applications from the perspective of temperature and the possible occurrence of lithium plating. The data obtained will be able to be used to optimize devices in space applications in general.

In this paper, lithium-ion batteries are briefly described, a part of the paper is devoted to lithium plating, and in the last section I describe the specific methods I use in my work to detect lithium plating.

## 2. LITHIUM ION BATTERIES

Lithium-ion batteries are nowadays one of the most well-known types of secondary cells. Unlike primary lithium cells, these batteries do not contain lithium in the solid state. The positive and negative electrodes are composed of intercalating lithium compounds. During charging and discharging, lithium ions are transferred between the positive and negative electrodes. Due to this effect, we refer to this type of battery as a rocking chair battery as the lithium ions travel between the electrodes during charging and discharging. The material for the positive electrode is usually a metal oxide with a layered structure, such as lithium cobalt oxide. The negative electrode is most often made of graphite [1].

The chemical process in lithium-ion cells works on the principle of reversible incorporation of lithium in an intercalation process. It is an intercalation reaction in which lithium ions are reversibly intercalated (inserted) and deintercalated (exported) from the host lattice without significant structural change. The positive electrode is made of metal oxide and lithium, which has either a layered or tunnelled structure. The negative electrode is usually made of graphitic carbon with a layered structure. The metal oxide and graphite act as hosts and reversibly incorporate lithium ions into their lattice. This process can be seen graphically in Figure 1. In this reaction, the electrolyte serves only as a medium in which the ions travel from one electrode to the other [1].



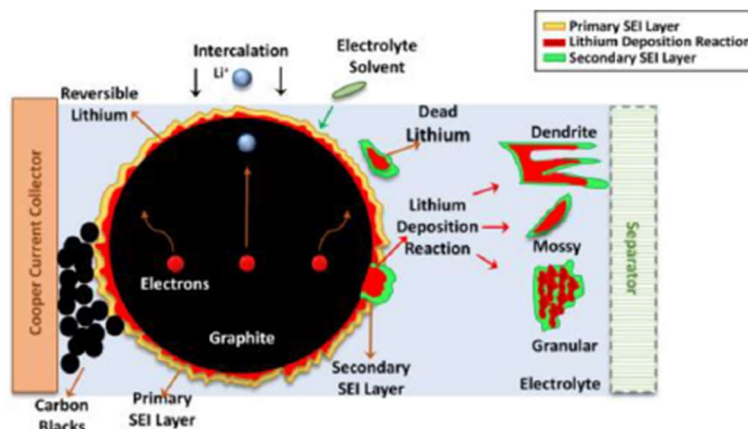
**Picture 1:** The principle of intercalation in a Li-ion battery [2]

### 3. LITHIUM PLATING

Lithium plating is a parasitic process that takes place together with the intercalation of lithium into the carbon structure when charging a Li-ion battery. During charging, two types of charging currents are distinguished, namely the "intercalation current" or intercalation current and the "lithium plating current" or the current that causes lithium plating. These two processes work against each other. As the charging process of the battery continues, the vacancies in the graphite structure begin to decrease and thus the intercalation current decreases, while the current that causes the lithium plating increases. When the anode potential drops below 0 V, the plating rate exceeds the intercalation rate, resulting in more massive plating that may already be problematic.

Lithium plating can have three different consequences. These are dead lithium, reversible lithium and SEI film (solid electrolyte interphase film). [3][4] We can illustrate these states in Figure 2.

Dead lithium can cause a reduction in the possible area for intercalation. The SEI film is the result of the reaction between the solvent and the electrolyte. These two states are irreversible and lead to a reduction in capacity during battery use. The reversible lithium intercalates back into graphite during the resting period. Based on the charging conditions, the morphologies of solid lithium can be divided into three types: sponge, particle and dendritic. The morphology is determined by the charging rate. The sponge and particle forms form at low charging currents, while dendrites form at high charging currents. Dendrite formation is probably the most dangerous phenomenon in which internal battery short-circuiting can occur.[3][4]



**Picture 2:** Diagram of lithium action on graphite anode [3]

#### 4. LITHIUM PLATING DETECTION

As I have already outlined in previous chapters, lithium plating is one of the main reasons for Li-ion battery ageing, so it is important to prevent or reliably detect this process. In this chapter, I present a method to reliably detect lithium plating. These methods can be divided into electrochemical and physical methods. The analysis itself can be approached by opening the battery using special instruments and examining it under microscopes or spectrometers. On the other hand, it is possible to use a non-destructive method such as electrochemical impedance spectroscopy. [3][4]

For my work I have chosen the coulombic efficiency method described by Umamaheswari Janakiraman et al. [5] in their paper on lithium plating detection methods. This method consists of measuring the coulombic efficiency over the cycling time of the batteries. In our case, this will be done according to the proposed measurement table, see Table 2. The coulombic efficiency (CE) describes the ratio between discharge and charge during one cycle and can be calculated according to the following formula [5]:

$$CE = \frac{Q_{discharge}}{Q_{charge}} \quad (1)$$

This is a typical formula for calculating Coulombic efficiency. Umamaheswari Janakiraman et al [5] recommend using a modified formula at low temperatures [5]:

$$CE = \frac{Q_{discharge} - Q_{irr. plating}}{Q_{charge}} \quad (2)$$

$$Q_{irr. plating} = Q_{charge} - Q_{discharge} \quad (3)$$

Based on the assumption that when lithium plating occurs, the coulombic efficiency decreases, it is possible to monitor changes in coulombic efficiency and evaluate the magnitude of lithium plating on that basis. The accuracy of this method is therefore dependent on how accurately we can measure the coulombic efficiency.

This paper also describes the coulombic inefficiency (1.0-CE), which can be normalized by time by dividing it by the cycle time. In doing so, A. J. Smith et al [6] write in their paper that with this normalization we can see that the charge loss at a given temperature depends mainly on the time per cycle, not on the number of cycles or the cycling rate. [5] [6]

For my observations, it will be useful to create plots of coulombic efficiency versus number of cycles for different temperatures and discharge currents. On these graphs it will be important to observe the decreases in Coulombic efficiency. Another suitable tool for detecting lithium plating I would choose to plot a graph of Coulombic inefficiency versus individual C values for each temperature. This graph shows at what values of C at specific temperatures lithium plating begins to occur. This method is demonstrated in the paper by Umamaheswari Janakiraman et al. [5].

## 5. BATTERIES UNDER TEST – MEASUREMENT DESIGN

The task of my master thesis is to measure batteries from GOMspace and specifically a Li-ion battery with 18650 case for space and aerospace products. The technical parameters can be seen in Table1[7].

**Table 1:**Table of battery parameters from GOMspace [7]

Parametr	Conditions	Min.	Nominal	Max.
Nominal Capacity (mAh)	2,0 V cut-off 2,95 V cut-off		3000 2750	
Voltage (V)	Safe	2	3,6	4.2
	Recommended	2,95		
Charge current (mAh)	0-50 °C		1500	4000
Discharge current (mAh)	-20+75 °C		1500	6000
Storage temperature(°C)	3 months	-20		45
	12 months	-20		20
Operating temperature (°C)	Charge	0		45
	Discharge	-20		60
Impedance (mΩ)				30

The actual cycling will take place according to Table 2, with 50 cycles planned for each position.

**Table 2:** Measurement design

Temperature	0,2C	0,4C	0,6C	0,4C
20 °C	2 batteries 3.0-4.0 V	2batteries 3.0-4.0 V	2batteries 3.0-4.0 V	2batteries 3.0-4.1 V
5 °C	2 batteries 3.0-4.0 V	2 batteries 3.0-4.0 V	2 batteries 3.0-4.0 V	2 batteries 3.0-4.1 V
0 °C	2 batteries 3.0-4.0 V	2 batteries 3.0-4.0 V	2 batteries 3.0-4.0 V	2 batteries 3.0-4.1 V
-5 °C	2 batteries 3.0-4.0 V	2 batteries 3.0-4.0 V	2 batteries 3.0-4.0 V	2 batteries 3.0-4.1 V
-10 °C	2 batteries 3.0-4.0 V	2 batteries 3.0-4.0 V	2 batteries 3.0-4.0 V	2 batteries 3.0-4.1 V
-15 °C	2 batteries 3.0-4.0 V		2 batteries 3.0-4.0 V	
Total	12	10	12	10



## 6. CONCLUSION

In this paper, I present my thesis research, which focuses on the detection of lithium plating in lithium-ion batteries. In the paper I discuss the principle of lithium-ion battery function, the negative effects that reduced temperatures can have, lithium plating and its effect on batteries. In the last chapter I introduce the batteries from GOMspace and describe the strategy by which the actual measurement is performed. Within the thesis I discuss the actual measurements and the processing of the data from each measurement. The data obtained will be able to be used for optimization of devices in space applications in general.

## LINKS

- [1] REDDY, Thomas. Linden's Handbook of Batteries, 4th Edition 4th Edition. Electronic book text, May 15th, 2010. ISBN 9780071624190.
- [2] PETR, J. Nové materiály pro Li-iontové baterie pracující na principu konverze. Brno, 2013. 71 s. Vedoucí diplomové práce doc. Ing. Marie Sedlářiková, CSc. FEKT VUT
- [3] LIN, Xianke, et al. Lithium Plating Mechanism, Detection, and Mitigation in Lithium-Ion Batteries. Progress in Energy and Combustion Science, 2021, 87: 100953.
- [4] LIU, Qianqian, et al. Understanding undesirable anode lithium plating issues in lithium-ion batteries. RSC advances, 2016, 6.91: 88683-88700.
- [5] JANAKIRAMAN, Umamaheswari; GARRICK, Taylor R.; FORTIER, Mary E. Lithium Plating Detection Methods in Li-Ion Batteries. Journal of the Electrochemical Society, 2020, 167.16: 160552.
- [6] SMITH, A. J.; BURNS, J. C.; DAHN, J. R. A high precision study of the Coulombic efficiency of Li-ion batteries. Electrochemical and Solid-State Letters, 2010, 13.12: A177.
- [7] GomSpace Battery: Datasheet [online]. 6 April 2018 [cit. 2021-11-29]. Dostupné z: <https://gomspace.com/UserFiles/Subsystems/datasheet/gs-ds-nanopower-battery-17.pdf>



# Saving area and power consumption in 65 nm digital standard cell library

Vojtěch Král

Brno University of Technology, Czech Republic

E-mail: xkralv03@vutbr.cz

**Abstract**—This study aims to investigate multi-bit pulsed latches in comparison with multi-bit flip flops as one of the low-power solutions in 65 nm technology process. Topologies of pulse generators and multi-bit pulsed latches were investigated to find out which can be more suitable. The pulse generator was chosen because of its low power and a small area in comparison with other options. The pulse generator is made of a simple AND logical gate and a double-stacked inverter. The pulsed latch was also chosen because of its low power, small area, and reliability of the circuit. The chosen topology is modified PPCLA. Simulations of the chosen topology had shown that multi-bit flip flops could be replaced with more effective multi-bit pulsed latches.

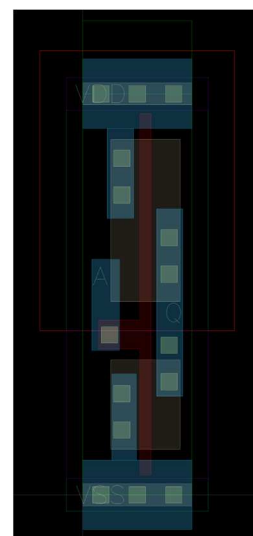
**Keywords**—digital standard library, integrated circuits, chip development, low power methods, pulse generator, multi-bit pulsed latch

## 1. INTRODUCTION

Automotive markets include chips that are developed for automotive-specific application. The problem nowadays is that these chips have high power density which can cause trouble. Some methods can reduce power consumption across the chip. This study will be focused on digital standard cells that are irreplaceable when it comes to integrated circuits. Digital standard cells have significant power consumption in comparison with the whole design, but these digital standard libraries can be enriched with special standard cells that can reduce power consumption to minimum. These special standard cells are also known as low power or ultra-low power solution [2]. One of the low power techniques is the multi-bit pulsed latch which can be used as a replacement for multi-bit flip flops. Multi-bit pulsed latches in comparison with multi-bit flip flops should have smaller area and power consumption [4].

## 2. DIGITAL STANDARD CELLS IN INTEGRATED CIRCUITS

As it was mentioned before, digital standard cells are an irreplaceable part of automotive-specific applications. Digital standard cells are well-defined cells that can be used in design as a building block [1]. These cells have different views that can represent a different function like schematic, layout, symbol, verilog, liberty, and others [2]. These building blocks are also pre-characterized to save time when they are used in a bigger design. These data of the cells for the whole library; like maximal load, input capacitance, and others are stored in liberty to make more complex simulation faster. Many types of digital standard libraries can be found in the technology; it depends on application or requirements from customers. The performance of the libraries is given by the height of the cells. The digital standard library has all cells with the same height because these cells can be easily connected to power rails and can be easily used with the automatic place and route tool [1]. Digital libraries can be often developed as high-density (HD) or high-speed (HS). High-speed libraries have greater height, larger speed, and bigger power consumption than high-density libraries. The width of the cells is given by the complexity of the cell, but the width of the cell should be minimized to achieve a maximal density of the design. These libraries can be also designed with different models of the transistor that have different threshold voltages [2]. The simplest cell is shown in Figure 1. It represents a simple layout inverter that is composed of two



**Figure 1:** Layout of inverter in 65 nm technology

transistors. The height of these transistors in Figure 1 is not the same because the PMOS transistor is slower than the NMOS transistor with the same height. On the other hand, the PMOS transistor can be used differently because the PMOS transistor is more immune to noise and has smaller leakage than the NMOS transistor. The height ratio of both transistors is given by the technology. These libraries can be composed of many types of common cells. The list of often designed cells is shown in Table 1. These libraries can be also enriched with special digital standard cells that can have better power consumption, area, special purpose, or even performance. Following parts can be considered as digital standard cells: level shifter, power-down cell, isolation cell, state retention flip flop, dual or multi-bit flip flops, dual edge-triggered flip flop, low swing dual edge-triggered flip flop, and a multi-bit pulsed latch. This study is primarily focused on multi-bit pulsed latches [2].

**Table 1:** List of common cells in the digital standard library

Name of cell	Description of the cell
BUF, INV, AND, OR, NAND, NOR XOR, XNOR	Simple logical functions with multi-inputs and different output strengths
HALF/FULL ADDER	2-bit Half and full adder with different output strengths
MUX / DEMUX	Multiplexor or demultiplexer with different output strength
ECO CELLS	Universal cells can be used in case of need
AOI or OAI	Multi-input and/or or or/and logical combination
FLIP FLOPS or SCAN FLIP FLOPS	Flip flops with reset/set and different output strength, Scan flip flop can be used as a shifter
LATCHES	Flip flop controlled with level
FILLCAP / FILLER	Decoupling capacitance/cell can connect power rails
DELAYS	Used for compensation of STA violations

### 3. MULTI-BIT PULSED LATCH AS A SPECIAL DIGITAL STANDARD CELL

Multi-bit pulsed latches are a new solution for low-power libraries. The pulsed latch is a simple latch that can be driven by a short pulse instead of a standard square clock source. Multi-bit means that there is more than one latch that can be driven by a short pulse. The function of the pulsed latch is the same as the flip flop only if the latch is driven by a short pulse generated from a standard square clock signal. To put it simply, the multi-bit pulsed latch is composed of two parts; the pulsed latch and a pulse generator which generates a short pulse from a standard square clock source [3].

**Table 2:** Types of negating delay with AND logical gate

Type of delay	Power [nW]	Predicted width of cell [nm]
Inverter	834.0	973
Stacked inverter	848.4	992
Double-stacked inverter	826.7	930

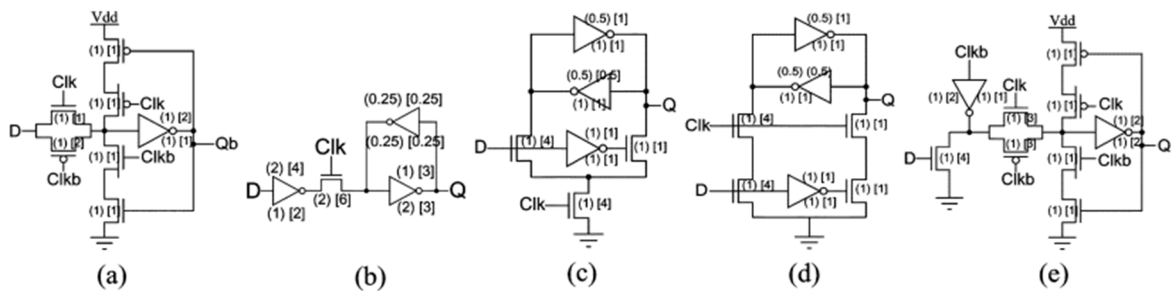
The pulse generator can be designed as local or as global. A local generator describes a generator that is designed inside of the multi-bit pulsed latch cell. The global generator is designed separately from the multi-bit pulsed latch cell. The local generator is more popular because the load of the generator is known, and the generator is custom designed. That means the parasitic routing is limited to the minimum [3]. The pulse generator is a simple logical combination and a negating delay block which can make enough wide pulse for the pulsed latch. The width of the pulse is dependent on negating the delay block. The size of the cell should be small and that is the reason why only the simplest standard logical gates like AND, OR and XOR can be considered. It can be used as a logical combination as it

was mentioned before. The main difference between these logical gates is that the pulse is generated on different events. AND gate makes the pulse on the rising edge, OR gate makes the pulse on the falling edge and XOR makes the pulse on both edges. The output should be changed on the rising edge. That is why the AND gate was chosen. The delay part can be also done differently. The simplest possible way how to make a negating delay block is to make an inverter with a bigger length. However, it can be replaced with a more saving solution like a stacked inverter or a double-stacked inverter. All simulated inverting delay parts are shown in Table 2. The best solution for the delay part is a double-stacked inverter because of the smallest total power and the smallest predicted width of the cell.

Choosing the pulsed latch is more complicated than choosing a pulse generator because the pulsed latch must have scan input. It means that a multi-bit pulsed latch can be connected as a shifter. Standardly, inputs and outputs are independent and are connected separately. The switch is made by the multiplexer and in this case, pass gate multiplexer is used to minimize the number of transistors. The multi-bit pulsed latch is limited by the minimal width of the pulse which can be used for a maximal functional circuit. However, the main problem of multi-bit pulsed latch in scan mode is that the width of the pulse is limited even from the top. If the pulse is too wide, the first latch will be working well but the second latch in the chain will be working on the same pulse. That means the multi-bit pulsed latch will shift the data to the second latch and not only to the first latch. The point to the next latch is often taken from the internal node which represents output data of the first latch. In all circuits that were simulated, there is no internal node that can reach maximal functionality because the propagation delay between the pulse and path to the next latch stage is too short. It means that the simulated circuit can work only in fast or slow corner. There are two solutions to this problem. The first one is an additional delay between the internal node and the input of the next latch. This kind of solution would have a bigger power consumption and bigger size of the area which makes this solution unacceptable. The second solution is to take the point for the next latch from the output stage where the Q is. The path will be long enough to prevent the stability of the circuit in all corners, both fast and slow. Those simulated multi-bit latches are shown in Table 3 and their schematic in Figure 2. The chosen topology PPCLA modified has a great minimal usable width of the pulse and clock to output propagation delay. One of the topologies is working in both cross corners. The total power and number of used transistors are also low. The modified version is different in feedback where on  $Q_b$  another inverter was added and from this point a pass gate is taken to the first inverter. When the topology of the pulsed latch is chosen, multi-bit versions of the chosen circuit can be simulated. These multi-bit pulsed latches are typically made in 2, 4, and 8 bits.

**Table 3:** Types of usable pulsed latches in the digital standard library

Name of pulsed latch	Power [nW]	No. of transistors [-]	Minimal width of pulse [ps]	CLK to Q propagation delay [ps]
PTLA	493.8	10	94	157.3
SSALA	371.6	11	53.5	120.4
SSA2LA	371.7	12	55	113.9
CPNLA	550.5	13	56.5	139.5
PPCLA	313.2	12	41	91.4
PPCLA MOD.	320.7	12	35	112.2



**Figure 2:** PPCLA (a), PTLA (b), SSALA (c), SSA2LA (d), CPNLA (e) [4]

#### 4. COMPARATION OF THE MULTI-BIT PULSED LATCH AND MULTI-BIT FLIP FLOP IN A SIMPLE SIMULATION

This chapter is comparing designed multi-bit pulsed latches with multi-bit flip flops because it is the first step before stating that multi-bit pulsed latches can be used as a replacement multi-bit flip flop. The comparison is based on power consumption and the predicted area which are the most important specifications of digital standard libraries. The compared circuits are shown in Table 4. One of the advantages of the multi-bit pulsed latch is that with an increasing number of bits, the savings are more significant in comparison with multi-bit flip flops. The generator is becoming less significant in comparison with the whole circuit. A 2-bit version of the pulsed latch has power consumption slightly higher than in the case of the 2-bit flip flop. The difference is about 7.8 % (normal) and 9.4 % (scan). The area can compensate for this difference because it is about 20 % smaller. The cell can be used as an advantage if the designer needs to save only area because it is still useful if the designer needs to save area regardless of power consumption. In the following multi-bit versions of pulsed latches, the power consumption is even better than in the case of multi-bit flip flops. In the 4-bit version, the power savings are about -6.3 % (normal) and -4.3 % (scan). The area is about -31.7 % better. In 8-bit version the power savings are about -14.6 % (normal) and -13.4 % (scan). The area is saved about -38.3 %.

**Table 4:** Comparison of multi-bit pulsed latches and multi-bit flip flops

	Mode	Pulsed latch			Flip flop		
Number of bits	-	2	4	8	2	4	8
Power consumption [μW]	Normal	2.512	4.272	7.689	2.330	4.561	9.003
	Scan	2.535	4.342	7.721	2.318	4.538	8.912
Predicted area [μm <sup>2</sup> ]	-	24.96	42.64	76.96	31.20	62.40	124.80

#### 5. CONCLUSION

The designed multi-bit pulsed latches in 65 nm technology show that the power consumption and the predicted area are better than in the case of multi-bit flip flops. The worst designed cell is a 2-bit pulsed latch which has a slightly bigger power consumption than the 2-bit flip flop. However, this designed cell can still be used because the predicted area is about 20 % smaller than a 2-bit flip flop. The designer can use this cell to reduce area regardless of power consumption. Other designed multi-bit pulsed latches like 4-bit and 8-bit versions have smaller power consumption and smaller area which implies that multi-bit pulsed latches can be used as saving digital standard cells in comparison with multi-bit flip flops. The following step is to layout these designed cells and compare them to have accurate results.

#### ACKNOWLEDGMENT

This work was supported by the Internal Grant Agency of Brno University of Technology, project no. FEKT-S-20-6526.

#### REFERENCES

- [1] BAKER, R. Jacob. *CMOS circuit design, layout, and simulation*. 3rd ed. Hoboken: Wiley-IEEE Press, 2010. IEEE series on microelectronics systems. ISBN 978-0-470-88132-3.
- [2] WESTE, Neil H.E a David Money HARRIS. *CMOS VLSI design: a circuits and systems perspective*. 4th ed. Boston: Addison Wesley, 2010. ISBN 0-321-54774-8.
- [3] J. Pliva et al., "Design of a custom standard-cell library for mixed-signal applications in 28 nm CMOS," 2017 IEEE International Workshop of Electronics, Control, Measurement, Signals and their Application to Mechatronics (ECMSM), 2017, pp. 1-6, doi: 10.1109/ECMSM.2017.7945867.
- [4] S. Heo, R. Krashinsky and K. Asanovic, "Activity-Sensitive Flip-Flop and Latch Selection for Reduced Energy," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 15, no. 9, pp. 1060-1064, Sept. 2007, doi: 10.1109/TVLSI.2007.902211.

# Negative electrode for sodium-ion batteries

A. Šimek<sup>1</sup>

<sup>1</sup>A Department of Electrical and Electronic Technology, Faculty of Electrical Engineering and Communication, Brno University of Technology, 616 00 Brno, Czech Republic

E-mail: [xsimek31@vutbr.cz](mailto:xsimek31@vutbr.cz)

**Abstract**—This work is primarily focused on the development of negative electrodes for sodium-ion batteries. This type of battery does not use lithium for its operation and could therefore be a promising successor to lithium-ion batteries in selected areas. Particular attention is paid to sodium titanates, which serve as an active material in the negative electrode mass.

**Keywords**—negative electrode, sodium titanate, sodium-ion, energy storage, renewable energy, theoretical capacity

## 1. INTRODUCTION

Lithium-ion batteries are one of the most common types of batteries used today, mainly due to their suitable properties, such as large capacity, high gravimetric and volumetric energy density, relatively long life and more. They are used not only in a number of portable applications such as mobile phones, notebooks, tools and others, but also as a stationary storage used for backup of buildings and equipment. Today, lithium-ion batteries are also widely used to store energy produced by renewable energy sources, thus increasing the overall usability of renewable sources. In recent years, there has also been a significant expansion of electromobility, where lithium-ion batteries are widely used due to the above-mentioned properties.

However, the ever-increasing interest in lithium-ion batteries poses a problem, which is the limited amount of lithium in the earth's crust. For this reason, the subject of today's research is the development of post-lithium storage systems that do not use lithium for their operation. These systems include, for example, sodium-ion batteries, on which this work is focused. Sodium-ion batteries are based on lithium-ion batteries technology, but have a number of disadvantages over them, such as lower gravimetric and volumetric energy densities. This type of technology could replace lithium-ion batteries in the future, especially in the areas of renewable energy storage. This step could save lithium to areas where their properties are more needed, such as for use in mobile phones and electric vehicles.

## 2. SODIUM-ION BATTERIES

The development of sodium-ion batteries (Na-ion) began in the 1970s together with lithium-ion (Li-ion) batteries. However, due to poorer results and demanding production processes, the development of Na-ion batteries was almost interrupted, as Li-ion batteries had significantly better parameters and relatively simple production technology. Interest in the technology of Na-ion batteries then increased significantly in 2012, when there was a significant increase in papers dealing with development of Na-ion batteries. [1]

Since the technology of both mentioned types of batteries is very similar, the basic components of Na-ion batteries are also very similar to those that use Li-ion batteries. However, there is a fundamental difference in the negative electrodes, as graphite cannot be used for Na-ion batteries. This is due to the significantly larger atomic radius, which is 190 pm for sodium and 167 pm for lithium. This difference of about 23 pm causes, that the sodium atoms or ions are not able to intercalate/deintercalate into/from the graphite structure and the Na-ion batteries with the graphite negative electrode reach a very small capacity. [1]

Despite a number of disadvantages, the technology of Na-ion batteries has a significant advantage, which is the abundant presence of sodium in the earth's crust and the associated significantly lower price compared to lithium. Due to this advantage, Na-ion batteries could be a cost-effective storage of energy

from renewable sources, where, as mentioned in the introduction, high energy density is not the primary requirement. The future development of this technology could significantly increase overall efficiency of renewable energy sources. [1]

### 3. PREPARATION OF SODIUM TITANATES

Sodium titanates were prepared from a colloidal suspension from which an oxide material was formed during high temperature synthesis. The starting colloidal suspension was composed of ethanol ( $\text{CH}_3\text{CH}_2\text{OH}$ ), Titanium(IV) ISO-Propoxide ( $\text{C}_{12}\text{H}_{28}\text{O}_4\text{Ti}$ ; or shortly Ti-ISO) and sodium carbonate ( $\text{Na}_2\text{CO}_3$ ). During the preparation, a different ratio of these basic substances was chosen in order to achieve the largest possible phase content of  $\text{Na}_2\text{Ti}_3\text{O}_7$  and overall phase purity. This phase of sodium titanate according to theoretical calculations reaches a theoretical capacity of  $177 \text{ mA.h.g}^{-1}$ . The prepared sodium titanate samples were then submitted for X-ray powder diffraction (XRD) to determine their phase composition. Table I below shows the amount of starting materials used and their time of synthesis. All prepared samples were synthesized in an elevator laboratory oven at  $800^\circ\text{C}$ . The table also shows the phase representation of titanates occurring in individual samples obtained from XRD. The samples also contained a small percentage of titanium dioxide ( $\text{TiO}_2$ ) in the form of rutile and anatase, as well as other forms of sodium titanates, the representation of which can be neglected.

**Table I:** Precursors of individual samples and their XRD results

Sample number	Ethanol [ml]	Ti-ISO [ml]	Sodium carbonate [g]	Time of synthesis [h]	$\text{Na}_2\text{Ti}_6\text{O}_{13}$ [%]	$\text{Na}_2\text{Ti}_3\text{O}_7$ [%]
01	30	9.9	1.98	5	74.1	12.1
02	30	9.9	1.98	2×5	67.6	26.5
03	30	9.9	1.98	10	51.1	45.0
04	30	19.8	1.98	10	× <sup>a</sup>	× <sup>a</sup>
05	30	19.8	1.98	20	79.6	20.4
06	30	9.9	0.99	20	69.3	30.1

<sup>a</sup> This sample was not sufficiently synthesized so XRD analysis was not performed.

From the phase representation of individual samples it is evident that the dominant phase is  $\text{Na}_2\text{Ti}_6\text{O}_{13}$ , which reaches a lower theoretical capacity. However, this phase can still achieve suitable properties, such as cycling stability.

### 4. ELECTRODE PREPARATION

The electrode mass was prepared from a reference sample from the previous research, whose phase representation was 51.1 %  $\text{Na}_2\text{Ti}_3\text{O}_7$  and 46.6 %  $\text{Na}_8\text{Ti}_5\text{O}_{14}$ . The reason for selecting this sample was its sufficient amount, so it was possible to perform grinding in a planetary mill. Only a small amount of other samples has been prepared so far, so the preparation of electrodes from other samples has not yet been possible and thus remains a task for follow-up research.

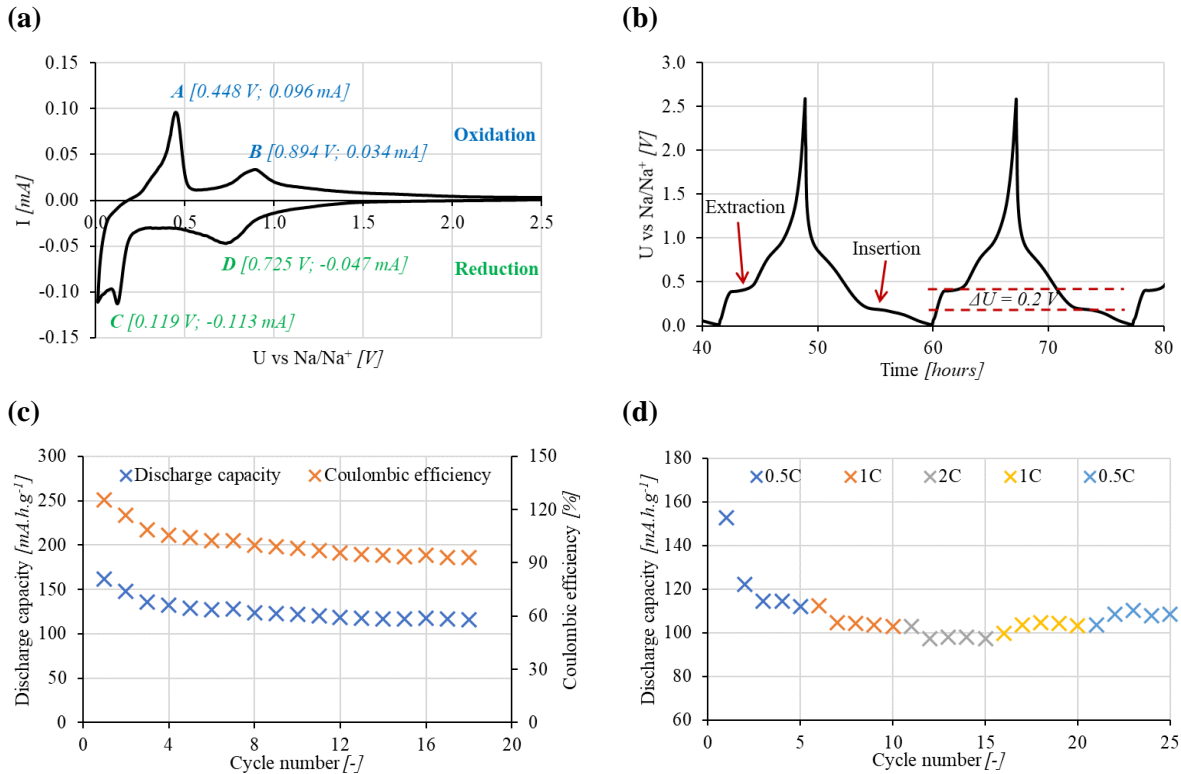
After grinding sodium titanate in a planetary mill, an electrode mass was prepared. A total of 0.5 g of electrode mass was prepared. In agreement with the supervisor, the composition was chosen where 10 wt.% was Polyvinylidene Fluoride (PVDF) as binder, 20 wt.% Super C65 as conductive additive, and 70 wt.% Sodium titanate as active material.

In addition, N-Methyl-2-Pyrrolidone (NMP) was added as a solvent as needed. The electrode mass was then applied to a copper foil with a thickness of about  $15 \mu\text{m}$ . The layer thickness was chosen to be  $80 \mu\text{m}$  (K-hand coater – brown rod). The applied layer was then dried at  $60^\circ\text{C}$  for about 3 days. After drying, electrodes with a diameter of 16 mm were cut out. The cut out electrodes were then pressed at  $1200 \text{ kg.cm}^{-2}$  and dried in a vacuum re-dryer at  $50^\circ\text{C}$  for about 2 days. The prepared negative electrodes were then assembled together with a sodium disc (source of sodium ions), porous separator (cellulose paper with a diameter of 18 mm) and electrolyte (170 ml of one molar  $\text{NaPF}_6$  in EC/DEC=50/50, where EC is ethylene carbonate and DEC is diethyl carbonate) into electrochemical measuring cells. The assembly has been carried out in a glove box in the presence of an argon atmosphere. Then the electrochemical characterization of the electrodes took place.



## 5. MEASURED RESULTS

Electrochemical characterization of the prepared electrodes was performed on a multichannel BioLogic potentiostat. Figures 1a to 1d below show the results of the measurement procedures performed.



**Figure 1:** Measured characteristics of prepared electrodes: **(a)** Cyclic voltammetry, **(b)** and **(c)** Galvanostatic cycling with potential limitation, **(d)** Rate capability

First, the open circuit voltage (OCV) was measured, which stabilized in the range of 2.4 V to 2.6 V for the individual electrodes. This was followed by the measurement of cyclic voltammetry (CV), where the current response to the applied potential was monitored. The potential range was chosen from 0 V to 2.6 V with a scan rate of 50 mV.s<sup>-1</sup>. Fig. 1c plots the second cycle of cyclic voltammetry. The first cycle was significantly different from the second one, as a passivation layer SEI (Solid Electrolyte Interphase) was formed. The marked peaks A and B correspond to the oxidation of the electrode, and peaks C and D indicate the reduction of the electrode. After comparing the measured results with the literature [2], it was further found that the larger peaks A and C correspond to the oxidation and reduction of the Na<sub>2</sub>Ti<sub>3</sub>O<sub>7</sub> phase and the smaller peaks B and D then the oxidation and reduction of the Na<sub>8</sub>Ti<sub>5</sub>O<sub>13</sub> phase.

Another measurement procedure was galvanic cycling with potential limitation (GCPL). The measurement was performed by cyclically charging and discharging the electrode with a current of 0.1C. A total of 20 measuring cycles were performed. Fig. 1b plots the electrode potential time dependence for the first two measurement cycles. In this figure, it is possible to observe two plateaus, which correspond to the extraction/insertion of ions from/into the structure of the negative electrode. The potential difference between these two plateaus is approximately 0.2 V and it is evidence of reversible electrochemical processes in the electrode structure. Fig. 1c shows the decrease in discharge capacity and Coulombic efficiency in individual measuring cycles. The discharge capacity decreased from 162.21 mA.h.g<sup>-1</sup> to 116.09 mA.h.g<sup>-1</sup> during 17 cycles, which corresponds to a decrease of approximately 28.5 %, or a decrease of 1.67 % on average per cycle. Initially, the decrease in capacity is significantly larger, which may again be due to the formation of the SEI layer. Coulombic efficiency is more than 100 % in the first few cycles, which is due to the slow kinetics of the sodium ions transition, so this

result is only indicative and serves mainly to determine the declining trend. Another reason for such a high efficiency may be the fact that a CV was performed before the measurement itself, which could have caused that higher efficiency, as the measurement was not performed from a standstill state.

Finally, rate capability (RC) measurements were performed. The following loads were selected for this measurement:  $0.5C \rightarrow 1C \rightarrow 2C \rightarrow 1C \rightarrow 0.5C$ . Each load was performed a total of five times, with the charging being constant at  $0.1C$  for all loads. The measured results are plotted in Fig. 1d. From the obtained results it is possible to observe a decrease in the discharge capacity with increasing load. When the load is reduced again, the discharge capacity increases again. However, this re-increase is not to the original values, but to lower values, which roughly corresponds to the overall decrease in discharge capacity and Coulomb efficiency in Fig. 1c. Overall, this decrease may be due to gradual degassing of the electrodes.

## 6. CONCLUSION

The aim of this work was to prepare sodium titanate, use it as an active material in the electrode mass and perform basic electrochemical characterization. According to theoretical calculations, sodium titanate  $\text{Na}_2\text{Ti}_3\text{O}_7$  reaches the largest theoretical capacity and is thus a suitable candidate for use in the negative electrodes of sodium-ion batteries. The result of this work is a relatively easy and inexpensive process for the synthesis of sodium titanate, which, however, still needs to be optimized to achieve the highest possible phase purity. The prepared negative electrodes show reversible electrochemical processes with a relatively high discharge capacity. Furthermore, the electrodes show a significant decrease in the discharge capacity during cycling due to the gradual degradation of the electrodes and the slow ion transfer kinetics. The overlap of this work is then to perform more measurements on existing samples in order to reveal long-term results. Furthermore, the preparation of negative electrodes from sample No. 05, which was almost 80 % represented by the  $\text{Na}_2\text{Ti}_6\text{O}_{13}$  phase. Although this phase is calculated to achieve less theoretical capacity, it is necessary to verify other properties, such as long-term stability during cycling.

## ACKNOWLEDGMENT

This work was supported by specific graduate research project of the Brno University of Technology: Materiály a technologie pro elektrotechniku IV., reg. no. FEKT-S-20-6206. Thanks for the supervision of Ing. Jiří Libich, Ph.D.

## REFERENCES

- [1] J. Libich, J. Minda, M. Sedlářiková, et al., “Sodium-ion batteries: Electrochemical properties of sodium titanate as negative electrode”, *Journal of Energy Storage*, vol. 27, no. 27, 2020.
- [2] T. L. Kulova, Y. O. Kudryashova, A. A. Kuz'mina, et al., “Study of degradation of  $\text{Na}_2\text{Ti}_3\text{O}_7$ -based electrode during cycling”, *Journal of Solid State Electrochemistry*, vol. 23, no. 2, pp. 455-463, 2019.

# POSTQUANTUM CRYPTOGRAPHY ON FPGA

A. Györi<sup>1</sup>, D. Smékal<sup>1</sup>

<sup>1</sup>Brno University of Technology, Czech republic

<sup>2</sup>Brno University of Technology, Czech republic

E-mail: [187286@vut.cz](mailto:187286@vut.cz), [smekald@vut.cz](mailto:smekald@vut.cz)

**Abstract**—This work describes the post-quantum algorithm FrodoKEM, its hardware implementation in VHDL and software simulation of implementation. The issue of postquantum cryptography and the VHDL programming language used to describe the functionality of the hardware was studied. The acquired knowledge was transformed into a functional simulation of all parts of the algorithm. All these parts have already been implemented separately, so that the functionality of every single part can be separately approached. These parts are key generation, encapsulation and decapsulation. After successful simulation. These parts will be synthesised and implemented to FPGA board NEXYS A7.

**Keywords**—Post-quantum cryptography, lattices, Frodo, FPGA, Testbench, LWE, hardware, VHDL

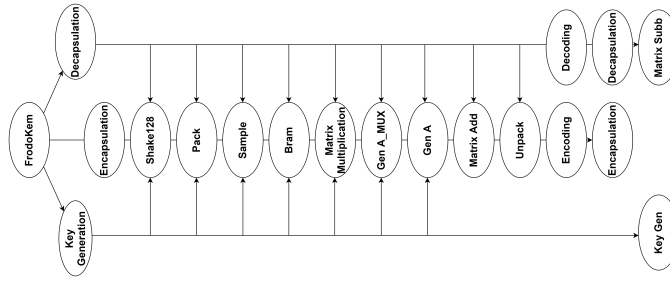
## 1. INTRODUCTION

Once quantum computers become available, it will be possible to solve previously difficult mathematical problems of factorization and discrete logarithms. These two issues underlie many currently used asymmetric cryptographic schemes. As a result, National Institute of Standards and Technology launched a 2017 call for tenders, which are required to submit an asymmetric encryption and signature scheme that can withstand classic attacks, etc. quantum computer. The submitted candidates use very differently approaches to achieving resistance to quantum computers, such as lattices, codes, hash functions, or multidimensional polynomials. Lattice-based cryptographic [5] schemes can be divided into three classes depending on the problem: encryption schemes such as textsf FrodoKEM [1] or Round5 [2] are based on the standard problem of learning with errors, the result is excellent security due to the absence of any structure and scalability for the price reduced efficiency or speed due to large parameters, keys and ciphertext. In contrast, Ring learning with errors -based schemes such as NewHope [3] or NTRU [4] offer better performance, as well as smaller keys and ciphertext than those based on standard lattices. The lattice-based cryptosystem class is based on the Module-LWE problem and offers a trade-off between the security of unstructured lattices and the performance of fully structured lattices. While a high-level C software implementation exists for all the proposals presented, there is virtually no hardware implementation, at least according to my personal research. Software implementations are performed on a processor and can be easily transferred between different types of processors. In contrast, hardware implementations can be performed using Field programmable gate array or Application-Specific Integrated Circuits. They are usually written in the hardware description language HDL, such as VHDL or Verilog. The task of this work is to design and implement the FrodoKEM algorithm directly in HDL.

## 2. ALGORITHM DESCRIPTION

FrodoKEM is a lattice-based key encapsulation scheme. Because FrodoKEM is based on LWE and not RLWE, it has larger parameters and is slower, but provides greater security than other lattice-based algorithms. The Frodo.Gen algorithm generates a matrix A line by line using either the SHAKE hash function or AES. I chose SHAKE in my implementation. Generation is deterministic based on the so-called short sequence and the index of the line just generated. Using Frodo.SampleMatrix, matrices are sampled by elements from a discrete Gaussian distribution. Frodo.Pack transforms the matrix into a bit string, while Fordo.Unpack does the exact opposite operation. Frodo.Encode encodes the bit string into an integer matrix, and Frodo.decode transforms the matrix into a bit string as the name implies.

### 3. ARCHITECTURE OF ALGORITHM IMPLEMENTATION



**Figure 1:** Architecture of FrodoKEM VHDL implementation

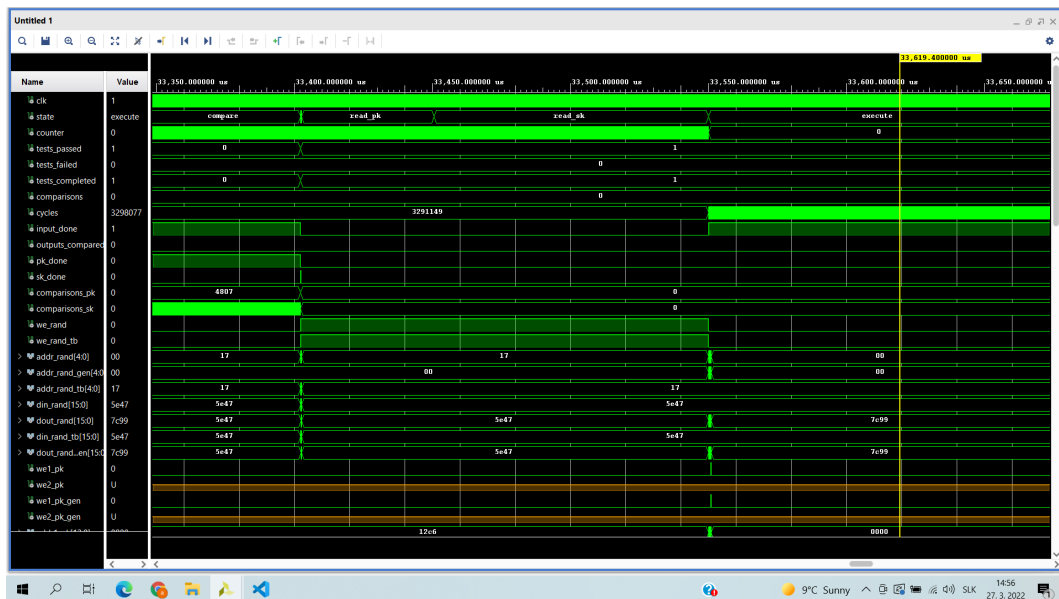
A scheme describes the way, algorithm has been divided to various components. As there is to be seen, some components are shared by more than one part of algorithm. The way a component works is as followed. Component gets an input, provides computation and sends output to following component. This way a component can be shared across more parts of the algorithm as the functionality is still the same only input is what makes the difference. For the implementation, the strongest and most robust Frodo-1344 parameter set was used.

	Frodo-640	Frodo-976	Frodo-1344
$D$	15	16	16
$q$	32768	65536	65536
$n$	640	976	1344
$B$	2	3	4
$\text{len}_\mu = l$	128	192	256
$\text{len}_{\text{seed}_{\text{SE}}}$	128	192	256
$\text{len}_s$	128	192	256
$\text{len}_k$	128	192	256
$\text{len}_{\text{pkh}}$	128	192	256
$\text{len}_{\text{ss}}$	128	192	256
$\text{len}_\chi$	16	16	16
$\chi$	$\chi_{\text{Frodo-640}}$	$\chi_{\text{Frodo-976}}$	$\chi_{\text{Frodo-1344}}$
SHAKE	SHAKE128	SHAKE256	SHAKE256

**Table I:** FrodoKEM parameter sets

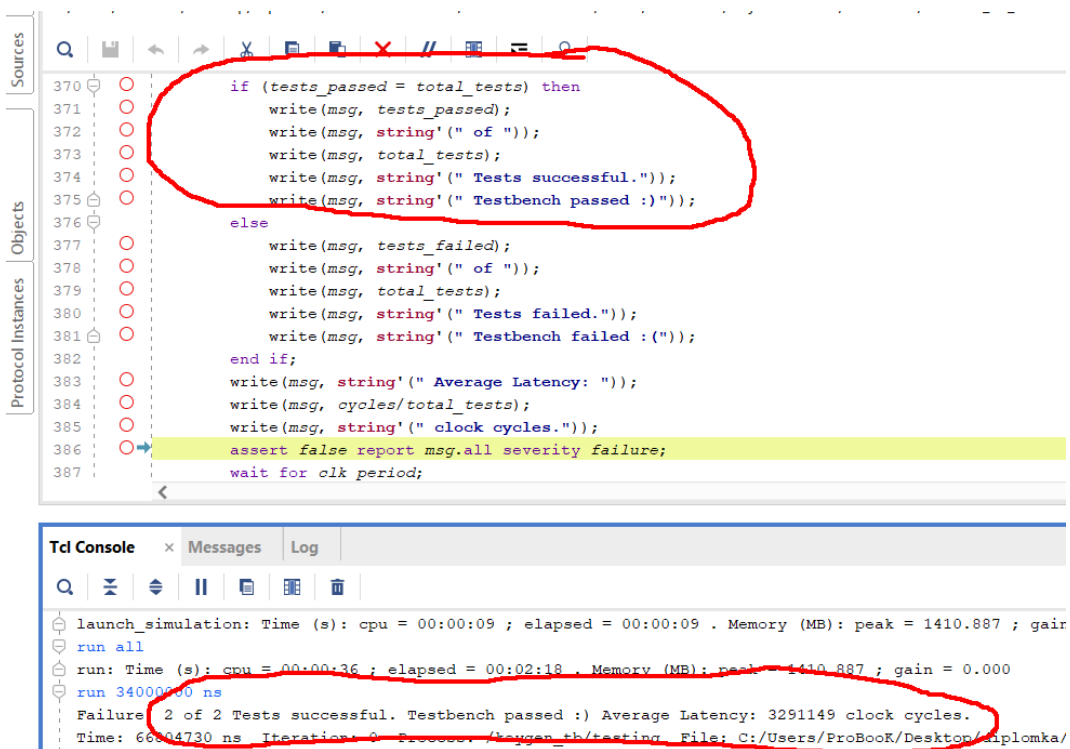
### 4. SIMULATION AND HARDWARE IMPLEMENTATION

All three parts were simulated and subsequently implemented to hardware separately as to show functionality of each one. The way testbench program was designed is as follows. The input of component is a set of hexadecimal values corresponding to certain part of algorithm. For example for key generation the input is initial randomness. These values are read from text file, after that computation processes. Output values are after that compared to values in text form, that are being expected as with the same initial randomness, same key must be generated. Same principles are being used with other parts of algorithm. On the following picture simulation result of key generation is to be seen. The input of this part are hexadecimal values of initial randomness. It is 23 values, that are being read in state read rand. That state is not visible on the screen, because it lasts only 23 clock cycles and is significantly smaller than other states. In state read pk and sk, the private and secret key, that are expected to be result are read.



**Figure 2:** Simulation of algorithm FrodoKEM key generation

Following image shows successfully finished simulation. If all comparisons of computed and initially loaded values were successful, test is evaluated as successful.



**Figure 3:** Result message of simulation

Hardware implementation is currently the only part of work that is yet to be done. After simulation in development environment XILINX Vivado, the three project parts will be synthesised and a bit stream will be generated. The result will be implemented on FPGA board NEXYS A7. The implementation will be run on board. Data will be provided to board via microSD card and the result of computation will be stored on same micro SD card.

## 5. CONCLUSION

The presented paper deals with the problem of implementation of currently proposed quantum security algorithms. Algorithm that was used was FrodoKEM. It was successfully implemented VHDL and simulated. Afterwards the solution will be planted on FPGA board and the functionality will be tested.

## REFERENCES

- [1] Alkim, Erdem and Bos, Joppe and Ducas, Léo and Longa, Patrick and Mironov, Ilya and Naehrig, Michael and Nikolaenko, Valeria and Peikert, Chris and Raghunathan, Ananth and Stebila, Douglas. FrodoKEM - Learning With Errors Key Encapsulation [cit. 14. 11. 2021]. Dostupné z: URL [https://www.designingbuildings.co.uk/wiki/Energy\\_infrastructure?fbclid=IwAR1CMidlcOQs7Rfu-n2HXK9emn5Q30xyrqj\\_eyiIsFNS7jHRdxnbW3e0nM0](https://www.designingbuildings.co.uk/wiki/Energy_infrastructure?fbclid=IwAR1CMidlcOQs7Rfu-n2HXK9emn5Q30xyrqj_eyiIsFNS7jHRdxnbW3e0nM0)
- [2] Baan, Hayo and Bhattacharya, Sauvik and Fluhrer, Scott and Garcia-Morchon, Oscar and Laarhoven, Thijs and Rietman, Ronald and Saarinen, Markku-Juhani O. and Tolhuizen, Ludo and Zhang, Zhenfei (2019) *Post-Quantum Cryptography*, Springer International Publishing, 978-3-030-25510-7.
- [3] Alkim, Erdem and Avanzi, Roberto and Bos, Joppe and Ducas, Léo and de la Piedra, Antonio and Pöppelmann, Thomas and Schwabe, Peter and Stebila, Douglas (2019) *New Hope – Algorithm Specifications and Supporting Documentation*
- [4] Chen, Cong and Danba, Oussama and Hoffstein, Jeffrey and Hülsing, Andreas and Rijneveld, Joost and Schanck, John M. and Schwabe, Peter and Whyte, William and Zhang, Zhenfei (2019) *NTRU – Algorithm Specifications and Supporting Documentation*
- [5] Chris Peikert. *A Decade of Lattice Cryptography* [cit. 17. 2. 2016]. Dostupné z: URL <https://eprint.iacr.org/2015/939.pdf>



# Wireless Battery Management System for Military Unmanned Vehicles

N.N. Pham<sup>1</sup>, J. Leuchter<sup>2</sup>, and H.Q. Dong<sup>3</sup>

<sup>1</sup> Department of Microelectronics, Brno University of Technology, the Czech Republic

<sup>2</sup>Department of Aviation Technology, University of Defence, Brno, the Czech Republic

<sup>3</sup>Viettel Army Telecommunication Industry Corporation, Hanoi, Vietnam

E-mail: [243756@vut.cz](mailto:243756@vut.cz), [jan.leuchter@unob.cz](mailto:jan.leuchter@unob.cz), [quanghuyhvk@gmail.com](mailto:quanghuyhvk@gmail.com)

**Abstract**—This paper introduces a Battery Management System (BMS) for Military Unmanned Vehicles, which uses multi-cell battery packs. The designed BMS is composed of slave modules and a master module. Each slave module is in charge of managing the state of one battery, include flowing current, voltage of battery, voltage of each cell, temperature of battery. From these obtained data, each module can predict the State of Charge (SoC) of connected battery. All slave modules are interconnected to the master module via Wireless Communication Protocol such as Wi-Fi, Bluetooth, or ESP-NOW. The master module will save all the received data to a microSD card, and if required, send these data to a workstation. These data are useful to predict the operation range of the vehicle, which is very important for decision the tactic of using. The designed BMS was verified in various applications in University of Defence in Brno.

**Keywords**—Battery Management System, Military Unmanned Vehicle, Wireless Communication Protocol.

## 1. INTRODUCTION

Since the late 19th century, unmanned vehicles have been used in wars to minimize human casualties. This field has not stopped growing, and the role of these vehicles has been proven to be irreplaceable in recent armed conflicts. The development trends of unmanned vehicles are decreases in size, mass as well as noise, and an increase in scope of operation. However, closely related to these trends is the issue of source parts. It is a fact that combustion engines are too noisy, heavy, and polluted, so nowadays people mainly use electric drivers. These drivers used to be powered by lead-acid batteries, but these batteries were too heavy with too-short life cycles. Therefore, they have been gradually replaced by Nickel batteries and the most modern ones, Lithium batteries[1]–[5]. The outstanding advantages of Lithium batteries is shown in Table I.

A lithium battery is usually a combination of many similar battery cells connected in series and parallel (multi-cell battery pack). Therefore, during usage, Lithium batteries require BMS to ensure safety. The BMS has the function of managing the states of batteries, as well as protecting batteries. However, battery protection function, especially against disbalance (voltage difference between cells), is not mentioned in this paper. The designed system focuses on monitoring the state of used batteries. Communication between the workstation and the vehicles is ensured by the wireless communication protocols.

**Table I:** Advantages of Lithium Batteries

Types of Battery	Voltage/cell [V]	Self-discharge [%/month]	Life cycle [cycle]	Energy density [Wh/kg]
Lead-Acid	1.80 – 2.10	4 – 8	500 – 700	< 40
Ni-Cd	0.85 – 1.35	10 – 15	500 – 1000	< 40
Ni-MH	0.85 – 1.35	15 – 30	500 – 1000	≈ 100
Li-Ion	3.00 – 4.20	≈ 2	> 1000	> 200
Li-Pol	3.00 – 4.20	≈ 2	> 1000	> 200

## 2. ESTIMATE THE STATE OF CHARGE

Currently, there is no absolute extract method to measure the SoC of Lithium batteries. We can only estimate it approximately according to 3 main methods

- Open Circuit Method – OCV
- Battery Internal Resistance – BIR
- Coulomb Counting Method – CCM

For the designed system, CCM method is chosen because of its feasibility and acceptable accuracy. The concept of this method is shown by the equation.

$$SoC = \frac{\sum_{t_0}^t \eta \times i \times \Delta t}{Q_0} + SoC_0$$

where  $SoC$  is the battery's state of charge at time  $t$ ,  $SoC_0$  is the initial battery's state of charge, at time  $t_0$ ,  $i$  is flowing current in period  $\Delta t$ ,  $Q_0$  is total (nominal) battery capacity,  $\eta$  is the efficiency of charging/discharging. Efficiency  $\eta$  depends on both the efficiency of electrochemical reactions at the anode and cathode and the loss in internal resistance. For Lithium batteries, the efficiency can reach up to 94% [6], [7]. In practice, this efficiency is determined by experiment by using the so-called deep discharge rather than theoretical calculation. The battery is charged to a fully charged state, after that it is discharged to a fully discharged state. In this process, we calculate the amount of charge leaving the battery, then compare it with the nominal capacity of the battery, and then estimate the efficiency [1].

## 3. SLAVE MODULES

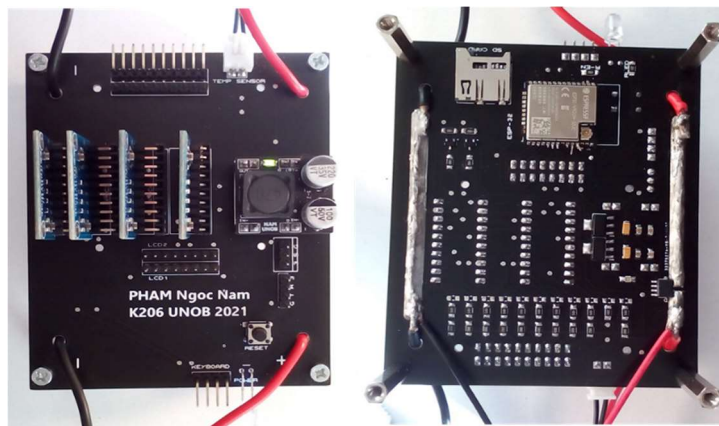
The main part of each slave module is the microcontroller ESP-32. Fig. 1 show a slave module of the designed BMS. The peripherals are in the charge of monitoring states of connected battery, include the following parameters:

- flowing current
- voltage of battery
- voltage of each cell
- temperature of battery

Flowing current is measuring by Hall effect current sensor ACS712. The Hall effect is stated that when a current  $I$  flows through a thin sheet of conductive material that is penetrated by a magnetic flux density  $B$ , a voltage  $U$  is generated perpendicular to both the current and field[1]

$$U = \frac{I \times B}{n \times q \times d}$$

The temperature of battery is measuring by thermistor, which operates based on the dependence on temperature of resistor of special materials. The resistor measurement is realized by using resistor divider or more exactly using Wheatstone bridge.



**Figure 1:** Slave Module

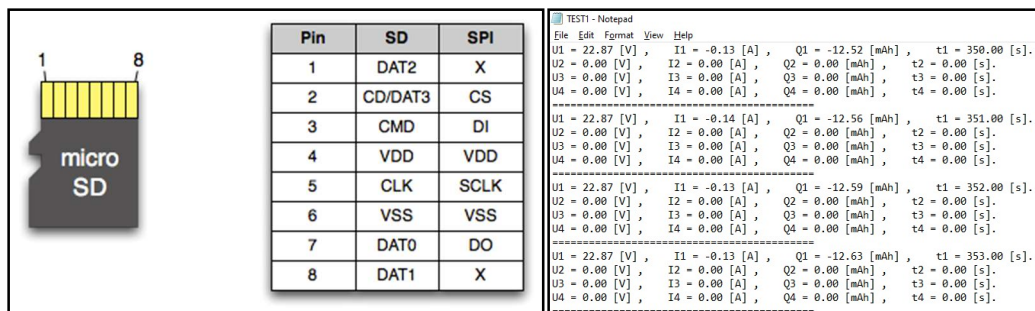
The output voltage of the Hall effect current sensor, the thermistor and the voltage of battery are measured by Analog-to-Digital Converter (ADC). The integrated ADCs on microcontroller ESP-32 have significant problems with nonlinearity, the characteristic of them is not a straight line. Therefore, it is necessary to choose a peripheral dedicated ADC[1]. The chosen one is 16-bit ADC ADS1115. This ADC has operating range 0-5V, resolution 16bit with the differential operation mode for solve the problem of noise. For connection with the microcontroller ESP-32, ADS1115 uses I2C protocols with the programable address from 0x48 to 0x4B

Other parts of these slave modules are power converter block and peripheral parts for displaying measuring data, indicating the state of system, antenna for wider range of communication. Power converter block contains

- switching DC/DC converter for converting the supplying voltage from battery to lower voltage with high frequency
- linear regulator, which generates low noise DC voltage from the output of switching converter for the operation of whole module

#### 4. MASTER MODULE

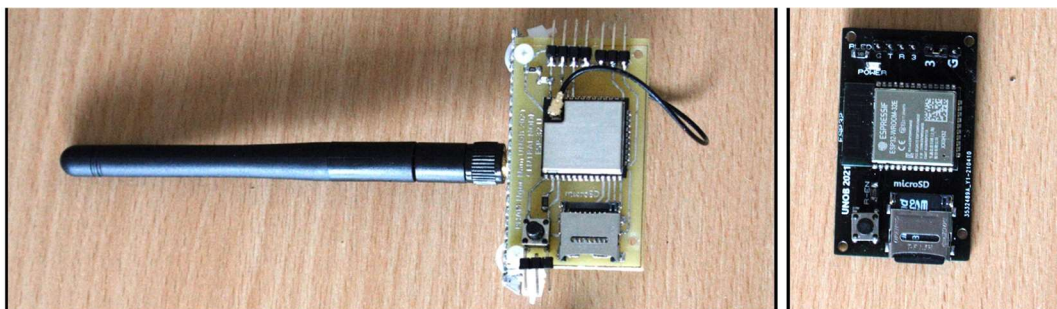
Master module is in charge of collecting measuring data from all the slave modules, saving these data to a microSD card and if necessary, sending data to a workstation. The main part of the master module is also an ESP-32 microcontroller. The microSD card is connected to microcontroller via SPI protocol. The pin mapping of microSD card for connecting to the microcontroller is shown in Fig. 2. An example of recording data to microSD card, in a .txt file is also shown in Fig. 2



**Figure 2:** pin mapping of microSD card and example of recording data to microSD card

Both the interconnection between slave modules and the master module, and the connection between the master module and workstation are based on Wireless Communication Protocols. This requirement is one of the main reasons of choosing ESP-32 microcontroller, which has integrated the function of wireless communications. ESP-32 is able to transfer data wirelessly via the following protocols

- 802.11 Wi-Fi
- Bluetooth 4.2
- Bluetooth Low Energy
- ESP-NOW



**Figure 3:** Master Module – with and without external antenna

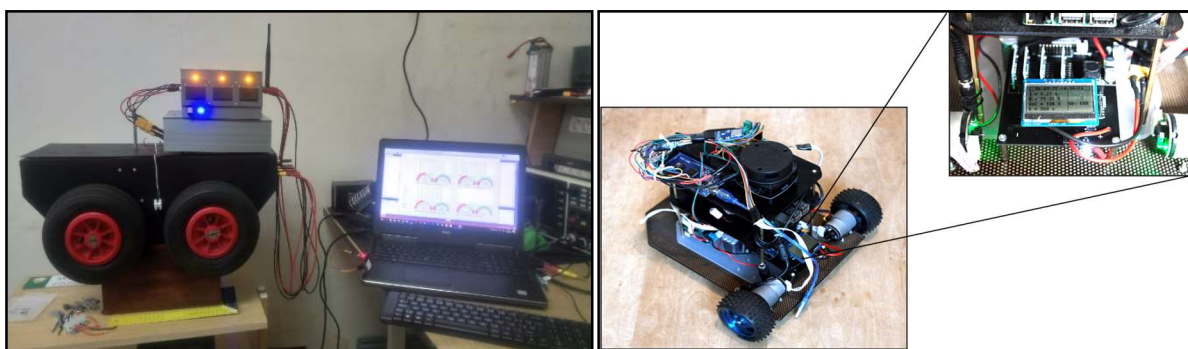
## 5. FUNCTIONAL VERIFICATION

The detailed designs of system, including design of Graphic User Interfaces and flow chart with above protocols are presented in [1]–[3]. Fig. 3 show the design of Master Module.

The functional verification of designed BMS is also presented in [1]–[3], [6]. The verification is realized by using Multimeters Agilent 34401A, Shunt Current Sensor Agilent 34430 30A, and climatic test chamber Vötschtechnik VCL7010.

The practical verifications which are realized in various applications in University of Defence are shown in Fig. 4

- The first one is the mobile interference device of Department of Aviation Technology, to monitor the state of source block using Lithium batteries.
- The second one is the 3-wheel differential drive mobile robot of Department of Air Defence, with the purpose of monitoring and comparing the energy consumption of various controllers.



**Figure 4:** Practical implementation of designed BMS

## 6. CONCLUSION

The paper focuses on the design of a BMS for Military Unmanned Vehicles. The BMS is designed for using in vehicles with multi-cell battery packs. The BMS contains a master module and slave modules. This structure provides a designed system with high modularity where we can insert more slave modules for applications with more battery packs. The connections between modules of BMS and between BMS and workstation is wireless connection, which is suitable for operation in practical tactic environment. The designed BMS was verified in several military vehicles in University of Defence in Brno.

Future work can be development of protecting function and balance function for connected batteries. The designed system can also be used in experimental analysis of various batteries where it measures all the parameters of batteries during the experiments.

Another interesting application of the designed system is an application in terms of unmanned vehicle' cooperation. Each slave module monitors the batteries of an unmanned vehicle and send these data to the control station. Based on obtained data, the cooperation of these unmanned vehicles can be modified to reach the higher efficiency.

## ACKNOWLEDGMENT

The work presented in this paper has been supported by the specific research project of the Faculty of Military Technology SV20-206/2, the University of Defence in Brno.

## REFERENCES

- [1] N. N. Pham, J. Leuchter, L. K. Pham, R. Bystřický, and H. Q. Dong, "Battery Monitoring System using Microcontroller ESP32 and Internet of Things," *ECS Transactions*, vol. 105, no. 1, pp. 517–529, Nov. 2021, doi: 10.1149/10501.0517ecst.

- [2] K. L. Pham, J. Leuchter, N. N. Pham, and V. T. Pham, "Design of Commercial-Off-The-Shelf System for Monitoring UAV's Accumulator," in *2021 International Conference on Military Technologies (ICMT)*, Jun. 2021, pp. 1–7. doi: 10.1109/ICMT52455.2021.9502781.
- [3] K. L. Pham, D. Rozehnal, J. Leuchter, N. N. Pham, R. Bystricky, and E. Blasch, "Test Bench for Regenerative Braking UAVs to Maximize Efficiency," in *2021 IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)*, Oct. 2021, pp. 1–10. doi: 10.1109/DASC52595.2021.9594345.
- [4] D. Valis, K. Hasilova, and J. Leuchter, "Modelling of influence of various operational conditions on Li-ion battery capability," *IEEE International Conference on Industrial Engineering and Engineering Management*, vol. 2016-December, pp. 536–540, Dec. 2016, doi: 10.1109/IEEM.2016.7797933.
- [5] J. Leuchter and A. F. Zobaa, "Batteries investigations of small unmanned aircraft vehicles," *IET Conference Publications*, vol. 2016, no. CP684, 2016, doi: 10.1049/CP.2016.0306.
- [6] K. L. Pham, J. Leuchter, R. Bystricky, V. T. Pham, and N. N. Pham, "Design and Simulation System for Quadrotor UAVs," in *2021 International Conference on Military Technologies (ICMT)*, Jun. 2021, pp. 1–5. doi: 10.1109/ICMT52455.2021.9502775.
- [7] P. Křibský, J. Křivka, L. Valda, and J. Žahour, "Li-Ion state of charge estimation methods," *2014 22nd Telecommunications Forum, TELFOR 2014 - Proceedings of Papers*, pp. 649–651, Feb. 2015, doi: 10.1109/TELFOR.2014.7034492.



# SUPER HOT CHIPS



**ASICentrum**

A COMPANY OF THE **SWATCH GROUP**

Naše srdce planou pro navrhování čipů!

Hledáme další nadšence, co by  
rozšířili naše řady!

★ ★ ★ **JOIN US!** ★ ★ ★

ASICENTRUM PRAHA / BRNO

WE LOVE TO **DESIGN** THEM.

## ČIPY KTERÉ JINDE NENAJDETE

- Navrhujeme ULTRA LOW POWER integrované obvody
- Naše čipy mají vždy něco jedinečného, co konkurence neumí
- Bluetooth, RFID NFC, RFID UHF, Sensor hubs, Optical Sensors
- Na tom všem se můžete podílet

## ČESKÁ FIRMA SE ŠVÝCARSKÝM ZÁZEMÍM

- ASICentrum je středně velká firma s rodinnou atmosférou
- Jsou zde špičkoví odborníci, od kterých se můžete hodně naučit
- Švýcarské zázemí díky mateřské firmě EM Microelectronic
- Máme vlastní továrnu na výrobu integrovaných obvodů

## PŘIJĚ NAVRHOVAT ČIPY

- Nabízíme brigády, diplomky a zaměstnání
- Potřebujeme návrháře, verifikátory i softwarové inženýry
- Studenty naučíme moderní techniky návrhu integrovaných obvodů
- Zaměstnanci získají obrovské zkušenosti s celým procesem vývoje čipů
- Skvělé platové ohodnocení a benefity
- Co očekáváme – zájem, chuť do práce a nezbytné základní znalosti ze školy

## HLEDÁ SE KOLEGA HOMO TECHNICUS

- Digital IC Designer
- Digital IC Verification Engineer
- Embedded Software Engineer
- Digital Backend Engineer

hr@asicentrum.cz

[www.hotchips.cz](http://www.hotchips.cz)



**em microelectronic**

A COMPANY OF THE **SWATCH GROUP**



# Proportional counter measurement system

O. Kolář<sup>1</sup>, M. Kubíček<sup>1</sup>

<sup>1</sup>Brno University of Technology, Czech republic

E-mail: [ondrej.kolar@vut.cz](mailto:ondrej.kolar@vut.cz), [kubicek@vut.cz](mailto:kubicek@vut.cz)

**Abstract**—This paper describes development of a proportional counter measurement system which is used to measure ionizing radiation. The system will allow easy acquisition of data from various detectors for further analysis and research of new processing algorithms. At first, hardware and software of the prototype system is described. Then, a test measurement and its results are mentioned.

**Keywords**—Ionizing radiation, Proportional counter, Acquisition system, Pulse classification, Microdischarge, Red Pitaya, STEMLab, Zynq, FPGA, Linux

## 1. INTRODUCTION

Proportional counters are a type of ionization chamber detector. They are very similar to Geiger-Müller tubes (G-M tube in short) as their mechanical construction is nearly identical. However, the main difference between them is how much information they can provide. G-M tubes are designed to detect ionizing radiation regardless of its properties with the advantage of their strong output signal. Proportional counters, on the other hand, provide only a weak output signal, but on the contrary it can provide much more information about the measured radiation. [1]

Similar type of detectors in terms of use case are scintillation detectors. These work on a very different principle than ionization chambers and are more complex. They are mostly used as spectrometers which means they can detect the energy of a measured particle. In theory, proportional counters can achieve the same goal with less complex equipment. However, processing of their output is more challenging because unlike with scintillation detectors the output signal depends not only on properties of a measured radiation but also on a place of interaction within the active zone of the detector [2].

## 2. EQUIPMENT

The block diagram of the measurement system is on figure 1. The system consists of several devices. The proportional counter LND 270132 is powered by high voltage laboratory power supply SRS PS350 through the preamplifier which feeds amplified signal from the detector into the custom built measuring unit.

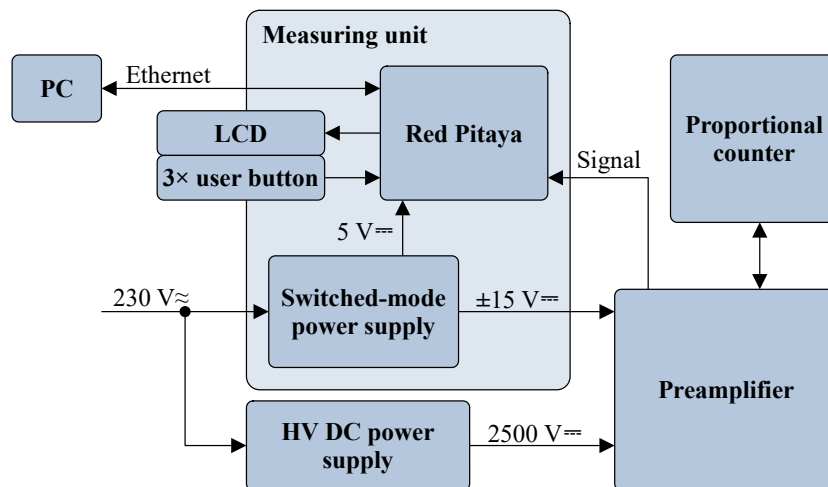


Figure 1: Block diagram of the measurement system

### 2.1. Measuring unit

The main part of the measuring unit, which can be seen in figure 2, is a RedPitaya STEMLab 125-14 board. It is equipped with Xilinx Zynq 7010 which is a system on a chip (SoC) consisting of a dual-core ARM processor and a field-programmable gate array (FPGA). The board also has an LTC2145-14 which is a dual 14bit analog-to-digital converter (ADC) with sampling frequency of 125 Msps. For high speed data transfer to a computer a gigabit Ethernet interface is used.

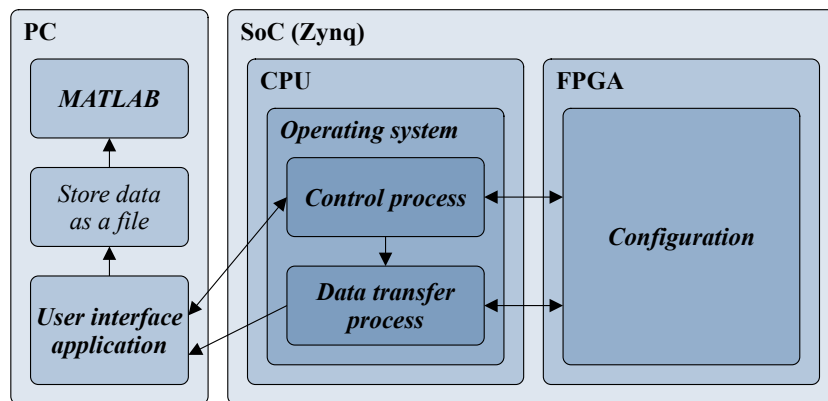
User interface consists of a 16x2 character display (LCD) and three buttons. The whole device is powered by a switched-mode power supply MEAN WELL RT-50C. This power supply also provides power to the preamplifier.



**Figure 2:** Photos of the custom built measuring unit

### 3. SOFTWARE

A complete software chain whose block diagram is in figure 3 had to be developed for the measurement system to work.

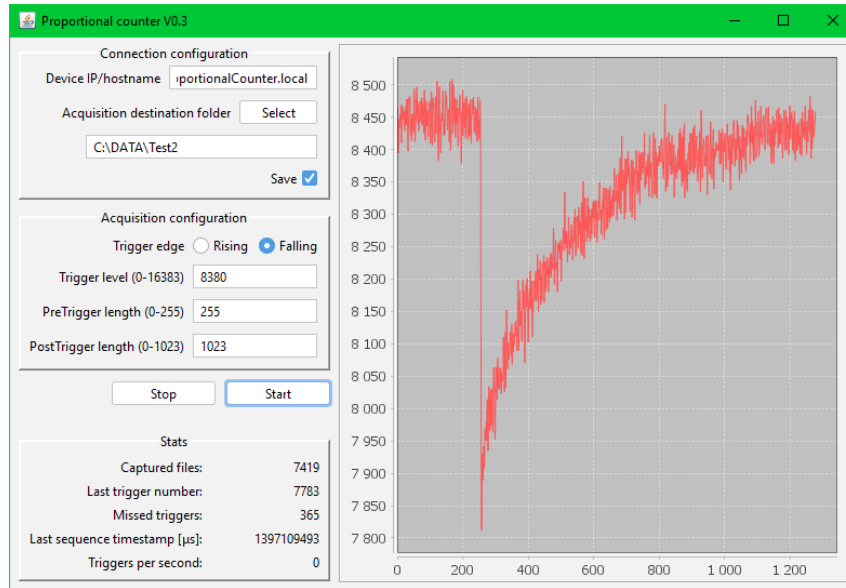


**Figure 3:** Block diagram of a software chain

Design in FPGA part of the Zynq SoC consists of a finite-state machine (FSM) written in VHDL language. The FSM takes parallel input from ADC and based on user-given parameters stores waveforms of measured pulses in a two-port block memory (BRAM). The second port of BRAM is connected to the CPU via AXI BRAM Controller IP. There are also other utility parts such as custom AXI LCD driver IP.

For the CPU a ready-to-use *RedPitaya STEMLab* Linux OS (operating system) is used. Although, the OS is fully functional by default, some tweaks were done to it. All unnecessary services such as a webserver or Jupyter were disabled to free-up the resources. The second major tweak was to change the system configuration to run every single process on the first CPU core only, by default. The reason for that is to reserve all of the computational time of the second core for the data transfer process to minimize its delay. Both the data transfer process and control process are written in C language.

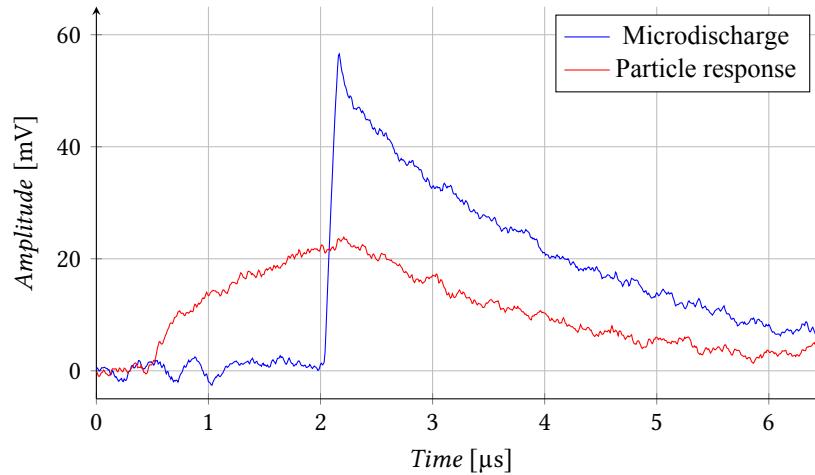
The measuring unit is controlled from a master PC using a Java application with a graphical user interface (GUI). The Java language was chosen for the application to be platform independent. The GUI which can be seen in figure 4 is minimalist to be as user-friendly as possible. The application is also responsible for saving the measured waveforms as binary data files for further processing.



**Figure 4:** GUI of a user interface application

#### 4. TEST MEASUREMENT

With the prototype system working a test measurement was conducted to ensure that the data acquisition is working properly. Data were gathered over the course of approximately 30 minutes. During this time period over 7 400 pulse samples were gathered. Two of them can be seen in figure 5 as an example.



**Figure 5:** Sample waveforms acquired from test measurement

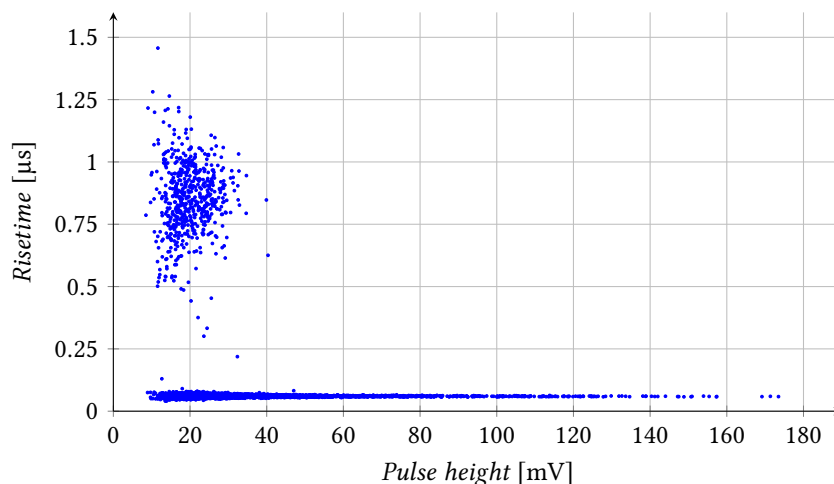
##### 4.1. Microdischarge

The figure 5, mentioned above, shows two different types of impulses. The one with a very steep rising edge is an outcome of a parasitic phenomenon called *microdischarge*. It is caused by material imperfections of the detector under a high voltage stress [3]. There is no simple way to prevent this once the detector is manufactured. Luckily, the shape of these pulses allows them to be easily distinguished and filtered out.

## 4.2. Data processing

Acquired data were processed in MATLAB. As they were a raw output of the ADC they were rather noisy and so they had to be filtered by a moving-average filter which was good enough for further processing as a proof of concept.

The most significant metrics for proportional counter pulse classification are their peak amplitude and rising time. With these two values the data can be plotted in an X-Y chart to distinguish the origin of each pulse [4]. A classification chart of the measured data is in figure 6. Points at the very bottom of the chart are microdischarges mentioned in section 4.1.



**Figure 6:** Data acquired from test measurement (Detector LND 270132 @ 2500 V)

## 5. CONCLUSION

The developed prototype of a proportional counter measurement system was briefly described. Then, a subsequent test measurement with the system provided a small sample of data for further analysis. Even without advanced processing it showed that the used detector suffers from frequent parasitic microdischarges. It also discovered a few bugs in the measurement system that will be fixed before conducting real measurements in the future. Data from these measurements will be used to research new processing algorithms for proportional counters such as genetic algorithm which is promising to give good results.

## ACKNOWLEDGMENT

This work was supported by the Internal Grant Agency of Brno University of Technology, project no. FEKT-S-20-6526.

## REFERENCES

- [1] V. Ullmann, "Detekce a spektrometrie ionizujícího záření," *Astronuklfyzika.cz*. [Online]. Available: <https://astronuklfyzika.cz/DetekceSpektrometrie.htm> [Accessed: 09-Mar-2022].
- [2] K. Berry, "Neutron Lifecycle Series: Helium-3 Gas Detectors at the SNS and HFIR," Oak Ridge National Laboratory, 21-Jul-2016. [Online]. Available: [https://conference.sns.gov/event/56/attachments/64/192/Lecture\\_6a\\_-\\_Neutron\\_Detection\\_with\\_He3\\_-\\_Kevin\\_Berry.pdf](https://conference.sns.gov/event/56/attachments/64/192/Lecture_6a_-_Neutron_Detection_with_He3_-_Kevin_Berry.pdf) [Accessed: 09-Mar-2022].
- [3] K. M. Heeger et al., "High-voltage microdischarge in ultra-low background  $^3\text{He}$  proportional counters," in *IEEE Transactions on Nuclear Science*, vol. 47, no. 6, pp. 1829-1833, Dec 2000, doi: 10.1109/23.914454.
- [4] T. Langford et al., "Event identification in  $^3\text{He}$  proportional counters using risetime discrimination," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 717, pp. 51-57, 2013, doi: 10.1016/j.nima.2013.03.062.

# A Fuzzy Controller for a Wind Tunnel Model

D. Piskoř, O. Mihálik, T. Sýkora, and J. Arm

Department of Control and Instrumentation, Brno University of Technology, 61600 Brno, Czech Republic.

E-mail: [203322@vut.cz](mailto:203322@vut.cz), [173702@vut.cz](mailto:173702@vut.cz), [173753@vut.cz](mailto:173753@vut.cz), [arm@vut.cz](mailto:arm@vut.cz)

**Abstract**—In this paper, an alternative approach to control the model of the wind tunnel is proposed. The aim is to find out if it is possible to achieve better results in regulating a system containing nonlinearities when a fuzzy controller is used instead of common linear controllers. The designed fuzzy controller will be implemented into a PLC via automatic code generation in MATLAB later.

**Keywords**—wind tunnel, fan, rotary encoder, controlling, fuzzy controller, linear controller

## 1. INTRODUCTION

When controlling real systems using conventional linear regulators, the required results may not be achieved. One reason may be the presence of internal nonlinear dependencies in the regulated system. When the output of a system depends nonlinearly on its inputs, sometimes the input-output relationship can be decomposed into two (or more) interconnected elements. In the case of two elements, the dynamics can be represented by a linear transfer function and the nonlinearities can be captured by using nonlinear functions of inputs and outputs of the linear system. This series connection of static nonlinear block with a dynamic linear block is called the Hammerstein model [1]. Hammerstein model applications cover several areas, such as modeling electromechanical systems, audio or speech processing, and predictive control of chemical processes [2].

One of the solutions to achieve better results in controlling (e.g. short transient to setpoint without overshoot or undershoot) a Hammerstein-type system can be to use a fuzzy controller. This work is focused on this solution.

## 2. THE MODEL DESCRIPTION

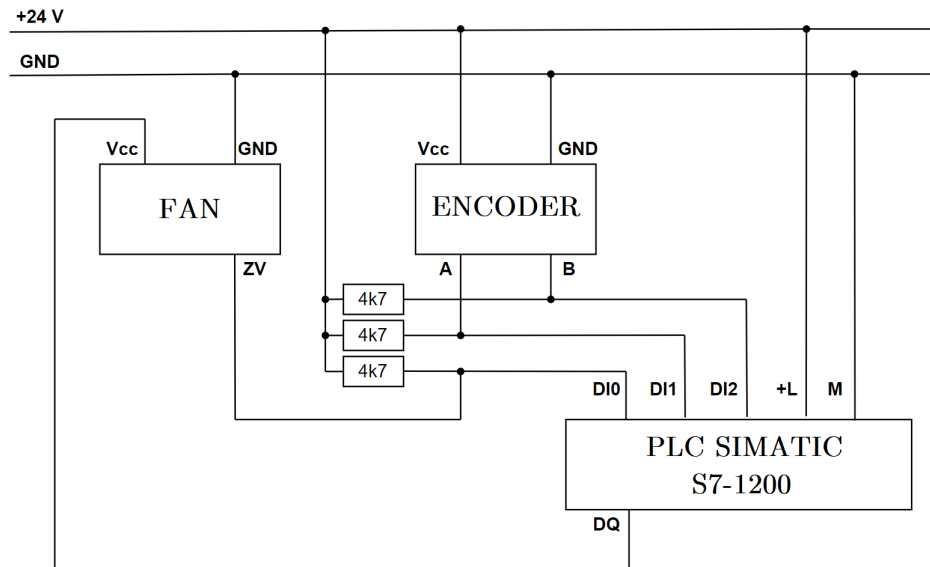
The real system is represented by a wind tunnel model. This model consists of three parts. The body is made up of a 500 mm length polypropylene tube with an inner diameter of 75 mm. As a source of airflow is placed a voltage-controlled fan SUNON MagLev ME80252V3-000U-G99 at the inlet of the tube. Then the air flows through the tube. The airflow is measured by a rotary encoder LPD3806 600BM G5 24C with a vane flow meter at the outlet of the tube. The whole model is shown in Fig. 1.

This wind tunnel model is controlled by a compact PLC SIMATIC S7-1214 DC/DC/DC manufactured by Siemens. The controlled variable is the output encoder frequency. This frequency can be transformed to RPM as well. The model contains a number of nonlinearities that occur in the fan and the airflow measurement. The fan starts rotating at an effective voltage of 12V and stops when the voltage decreases below 9V. The output flow meter is unable to detect an airflow lower than that generated when the fan is supplied by 10V. Due to this situation, the whole system is used in a range of 12-24V.



**Figure 1:** The whole wind tunnel model (left) and the detailed look (right [3])

The wiring diagram of the wind tunnel model with PLC is described in Fig. 2. Three resistors with a nominal electrical resistance of 4,700  $\Omega$  are used to suppress the effects of unwanted noise. Output ZV at the fan is an integrated speed sensor and outputs A and B at the encoder are its output channels.



**Figure 2:** The wiring diagram of the wind tunnel model

### 3. THE DESIGNING OF THE FUZZY CONTROLLER

For designing and subsequent testing of the created fuzzy controller, it was necessary to transfer the wind tunnel model to Simulink. For this, the MATLAB Identification Toolbox was used. According to comparing the response of the real system with the response of the identified system, the accuracy of the identification reached 92%.

The fuzzy controller was designed and tested in MATLAB and Simulink, where Fuzzy Toolbox was used. A Mamdani fuzzy PI controller was chosen because it is sufficient and easier to implement than a fuzzy PID controller. The controller has two inputs (a deviation and a sum of the deviation) and one output. The universe of both inputs was set to a range of -10 to 92 and contains four membership functions called ZO-zero, PS-positive small, PM-positive medium, and PB-positive big. Two membership functions have a form of the  $\Lambda$ -function and the others have  $\Pi$ -function. The positions of the vertices of these membership functions in the universe were chosen based on the shape of the identified nonlinearity of the wind tunnel model. The second input is identical to the first input. The type of implication was chosen as minimum consequent implication, aggregation as maximum, and defuzzification as a centroid.

In a similar way, the membership functions for the fuzzy controller output were created. The difference was the number of membership functions and the scope of the universe. Here were selected 6 membership functions of the type of the  $\Lambda$ -function (ZO-zero, PN-positive negligible, PS-positive small, PM-positive medium, PB-positive big, and PVB-positive very big).

The designed fuzzy controller follows these rules:

		$e$			
		Z	PS	PM	PB
$\Sigma e$	Z	Z	PN	PS	PM
	PS	PN	PS	PM	PB
	PM	PS	PM	PB	PVB
	PB	PM	PB	PVB	PVB

**Figure 3:** The rules of the designed fuzzy controller

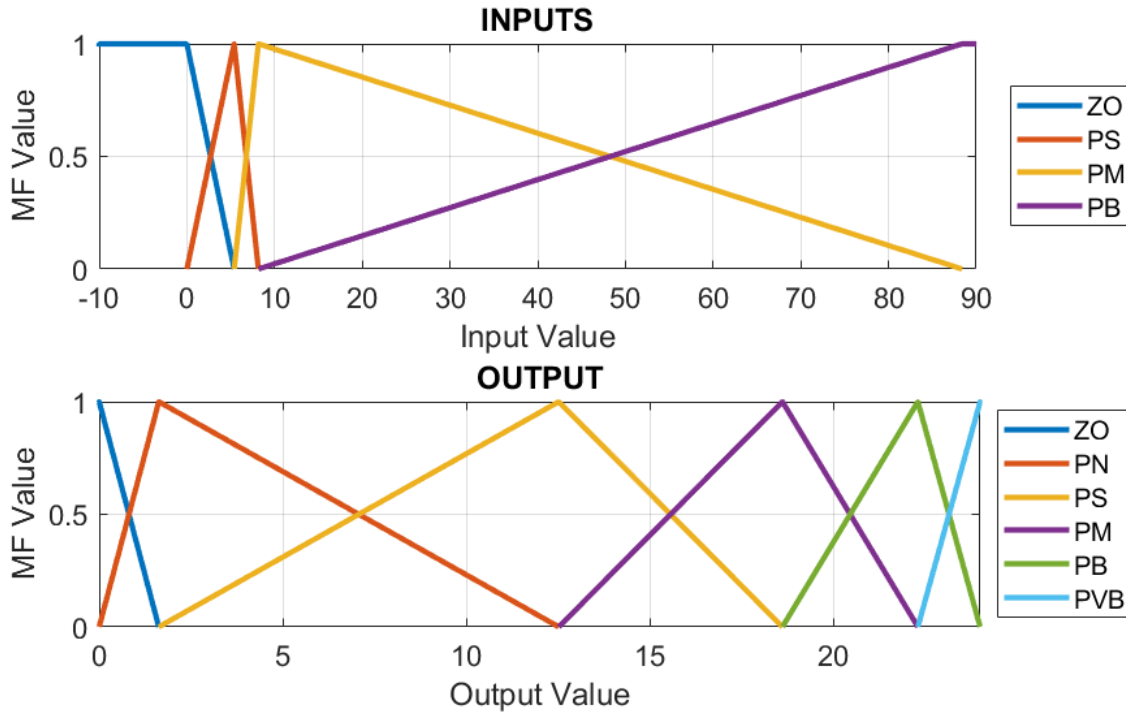


Initial estimation of the proportional gain and the integral time constant of the fuzzy PI controller was obtained by using a block in Simulink called PID Controller.

Subsequently, the created fuzzy PI controller was optimized by using the MATLAB function called *fminsearch*. This function uses the Nelder-Mead method to minimize [4]. The values of proportional gain, integral time constant, and also the positions of all membership functions at the input and the output were optimized. As the minimization criterion was chosen the Integral of time-multiplied Absolute value of Error criterion (ITAE).

$$I = \int_0^{\infty} |e(t)| \cdot t \, dt \quad (1)$$

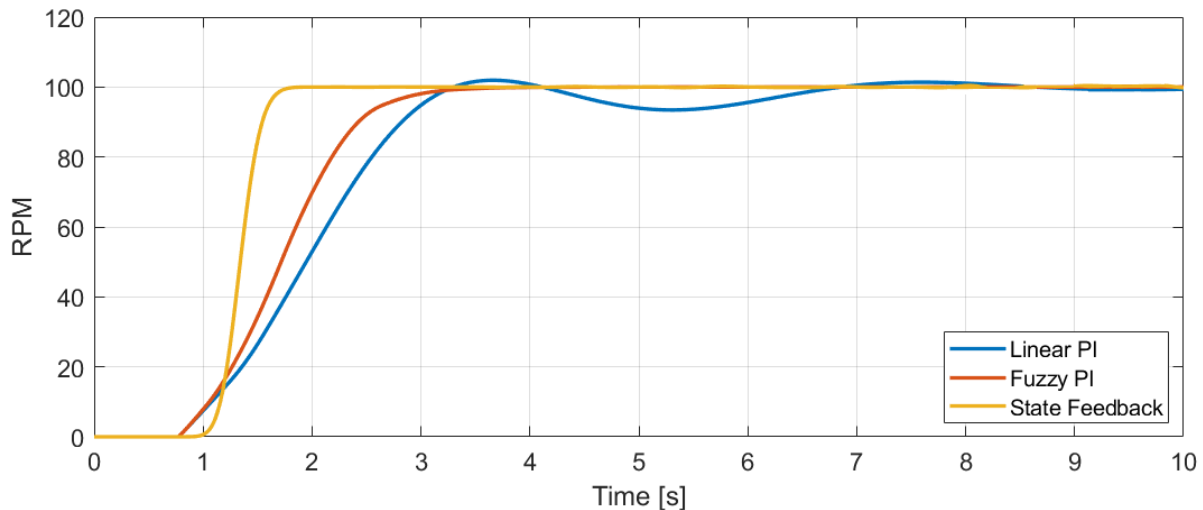
The optimized positions of membership functions on the inputs and the output of the fuzzy PI controller are shown in Fig. 4.



**Figure 4:** The optimized membership functions at the inputs and the output of the controller

#### 4. RESULTS

For a selected setpoint of 100 RPM, the output of the control circuit stabilizes in approximately 4 seconds. During the control process, there is no overshoot or undershoot of the setpoint. The response of the control circuit to the setpoint value 100 RPM is shown in Fig. 5. For comparison, the response of the linear PI controller and state feedback controller are also plotted. The linear controller was designed in Simulink by PID Controller block. The proportional gain and the integral time constant of this linear PI controller were also optimized by *fminsearch* function. As can be seen in the figure, the control circuit with the linear PI controller has a response with damped oscillations. The control circuit with the state feedback controller with optimized pole placement reached the best result. For the fuzzy controller the value of ITAE criteria is 166.81, for the linear controller is 291.69, and for the state feedback controller is 91.78. Here it is important to mention, that the results were obtained on the identified system in Simulink except for the control circuit with the state feedback controller, where the result was obtained on a linearized model of the identified system. This linearized system does not include all nonlinear dependencies and that could be the reason why the control circuit with the state feedback controller has the best result.



**Figure 5:** The comparison of control circuit responses to setpoint step change

## 5. CONCLUSION

In this paper, an alternative approach to control the model of the real system containing nonlinearities was proposed. The designed fuzzy PI controller of the type of Mamdani has four membership functions at the inputs and six at the output. In its design, emphasis was placed on regulation without overshoot. As can be seen in Fig. 5, the designed fuzzy PI controller achieved better results in controlling than the linear PI controller. The linear controller has damped oscillations around the setpoint value. The best result reached the state feedback controller, but this controller was linked into the control circuit with the linearized model of the identified system. This linearized system does not include all nonlinear dependencies which can affect the control process. A robustness analysis has not been performed yet. The future work will be focused on the transformation of the MATLAB code into a PLC via automatic code generation in MATLAB.

## ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-20-6205—“Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0” financially supported by the Internal science fund of Brno University of Technology.

## REFERENCES

- [1] ABONYI, J. Fuzzy model identification for control. Boston: Birkhäuser, 2003, 273 s. ISBN 978-1-4612-6579-5. DOI: 10.1007/978-1-4612-0027-7
- [2] THE MATHWORKS, INC. What are Hammerstein-Wiener Models?. Help Center for MATLAB, Simulink and other MathWorks products [online]. [cit. 2021-12-30]. <https://www.mathworks.com/help/ident/ug/what-are-hammerstein-wiener-models.html>.
- [3] SÝKORA, T., J. STREIT, O. BAŠTÁN and T. BENEŠL, Automatic generation of a PLC controller based on a control system identified model. DE GRUYTER Journal of ELECTRICAL ENGINEERING
- [4] THE MATHWORKS, INC. fminsearch. Help Center for MATLAB, Simulink and other MathWorks products [online]. [cit. 2022-03-11]. <https://www.mathworks.com/help/matlab/ref/fminsearch.html>.

# Using wavelength multiplex and optical circulators for power-supply temperature fiber-optic polarization sensor

Z. Vyležich<sup>1</sup>, and M. Kyselák<sup>1</sup>

<sup>1</sup> University of Defence, Faculty of Military Technology, Department of Electrical Engineering, Kounicova 65, Brno-střed, 662 10, Czech Republic

E-mail: [zdenek.vylezich@unob.cz](mailto:zdenek.vylezich@unob.cz), [martin.kyselak@unob.cz](mailto:martin.kyselak@unob.cz)

**Abstract**—A temperature fiber-optic polarization sensor, which thanks to its non-electricity, inertness to its surroundings, and high sensitivity is suitable for remote sensing in dangerous environments with a high risk of explosion, can be power-supplied via classical single mode fiber. The research investigates the use of a wavelength multiplex of 1550 and 1557,36 nm, and connections of polarization maintaining circulators, optical routes, and couplers to ensure quality power-supply to the sensor and parallel data transmitting. Separately, without the connected sensor, data transmission was successful without problems. With a connected sensor, due to the whole concept of a circuit caused problems, and data transmission was not working. Sensor sensitivity was tested by the application of a container with water at different temperatures. Experiments show that sensor sensitivity is not negatively affected by data transmission, it only affects the outgoing level of power.

**Keywords**—Polarization sensor, Temperature sensor, Using wavelength multiplex for power-supply temperature fiber-optic polarization Sensor

## 1. INTRODUCTION

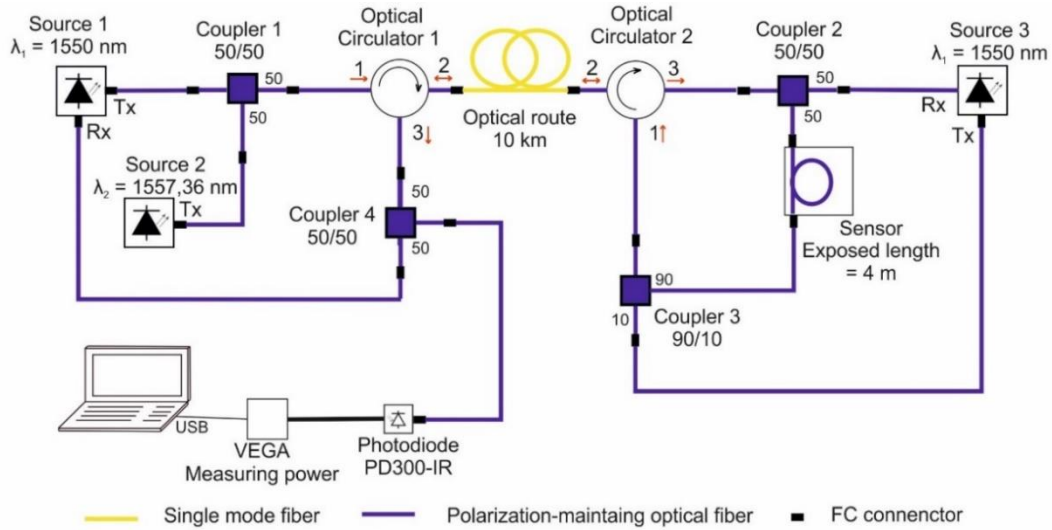
The thesis proposes the utilization of wavelength division multiplexing and optical circulators for simultaneous power supply of a remote fiber optic polarization sensor and data carrying. The main advantage is the usage of an optical circulator, which allow to use the same optical single mode route to power supply the sensor and transmit internet data across single mode fiber. This connection is suitable for remote sensing in a dangerous environment with the possibility of explosion, where it is necessary to have a non-electric sensor that does not affect its surroundings.

The issue of a sensor is the subject of a separate paper [1]. There are written details about physical principles and effects inside fiber. The paper [2] also discusses the sensor, but the focus is on the effect of temperature on light-descriptive quantities. There are also evaluated outcomes from the sensor, but more details about describing and evaluating them can be found in [3]. All the articles concern the sensor, but this thesis is more focused on overall connection and usage in a real environment with all its negative external influences, as paper [4] describes.

## 2. WORKPLACE SCHEME

Small form pluggable modules laser diodes SPS-7380WC, 1.25G SFP-ZXD 1550 nm, dedicated to data carrying, and SFP-PLU-ZR-DWDM-25 10G DWDM ZR 1557,36 nm used as power supply for a sensor are chosen as sources. Power is then coupled from these sources in CP-PS-1X2-1550-50/50-900-5-1M-3x54-FC coupler. It couples the half of incoming power from its both inputs. Then the light goes to a polarization maintaining optical Circulator 1 PMCIR-3-B1550-F-2-1-FC, where it is possible to use only one route for transmitting signal. A 10 km long optical route of single-mode fiber is connected between circulators. At the end of the route there is connected Circulator 2. It used its property to allow pass only y-plane. Coupler 2 is connected as a splitter and equally splits incoming power into two branches. Data reaches Source 3 in the first branch. The splitter does not affect the polarization state, so it is suitable for power supply of the sensor in the second branch. This sensor has in the input +45° rotated FC connector, so the sensor sufficiently reacts on temperature changes in both planes. The output power is then coupled with signal from Source 3, which operates at 1550 nm, in the Coupler 3, 1x2

90:10 PM Narrowband Coupler. The 10% of the power comes from Source 3 and 90% from the sensor. After that, the light travels the same way back to the beginning. Through the first circulator and another in Coupler 4. There is power equally split between the VEGA laser power meter and computer.



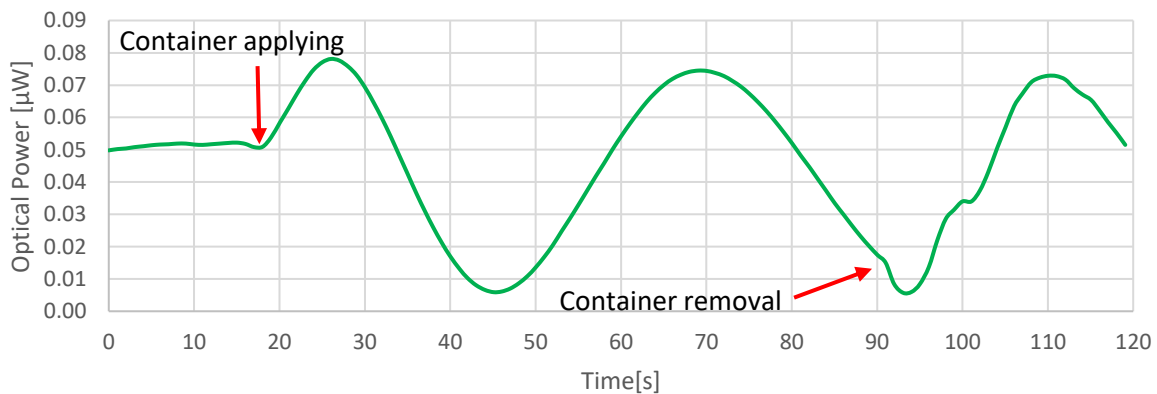
**Figure 1:** Workplace layout

### 3. RESULTS

The method of testing sensor sensitivity is based on applying a container with water of different temperatures. The laboratory has ambient temperature 24 °C and the sensor is tested by 36,5 °C, because it is body temperature, 0 °C container and 95 °C, to prove sufficient reactions.

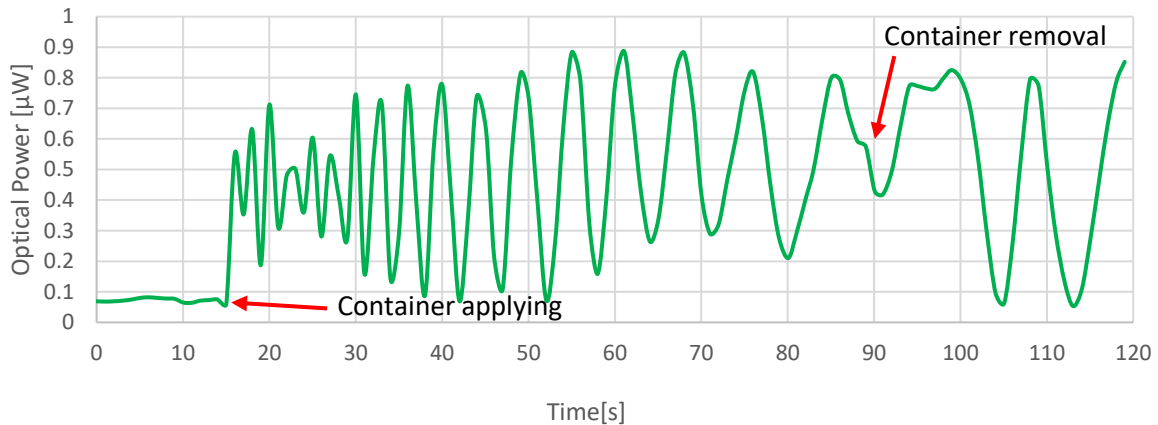
During these tests, optical sources operating at wavelength  $\lambda_1 = 1550$  nm and carrying data are turned off. This solution was chosen to evaluate only the sensor data operating at  $\lambda_2 = 1557,36$  nm. When all sources are connected, optical power has bigger values, but still only the wavelength of 1557,36 nm is changing. Therefore, it is necessary to set the proper range of measuring device. Data carrying is tested separately without a connected sensor. In that case, data carrying is functioning, but the measurements with a connected sensor were not sufficient. Couplers are not able to split wavelengths, so the data from Source 1 and the power from Source 2 go through the sensor and then in Coupler 3 are mixed with data from Source 3. This phenomenon causes non-functioning data carrying. Nevertheless, the sensor is still working properly.

The output power is strongly dependent on an optical route between the optical circulators. There is a noticeable disadvantage to an optical circulator because he lets pass only y-axis polarization, so the power of the other polarization states is dropped. Another observed phenomenon is reflection, which affects incoming power from the sensor and Source 3. It is not a negligible value.



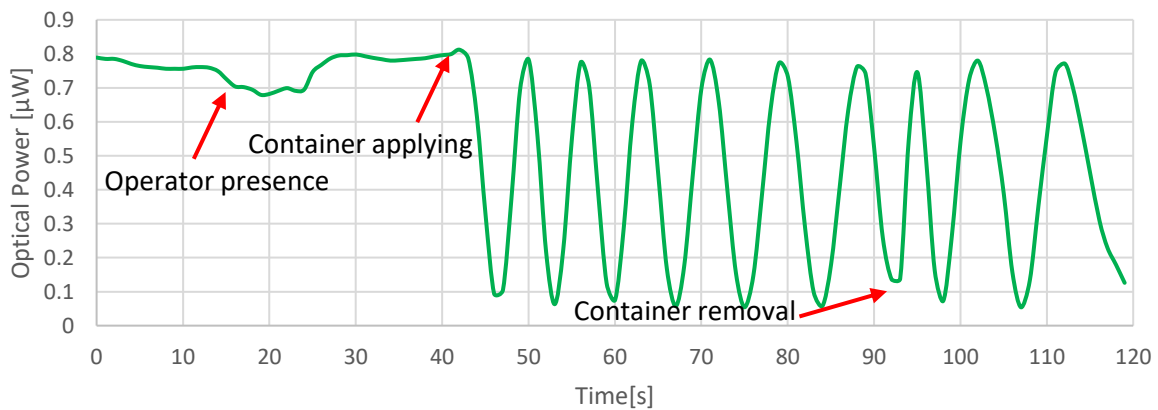
**Figure 2:** The change in optical power caused by the application of a container of water at 36,5 °C

A container with water at 36,5°C was applied in the 12th second. The reaction in Fig. 2 has almost a harmonious course, and the time of application is clearly visible in the reaction of optical power. Container removal is not so obvious, but it can be noticed by violation of harmonious course.



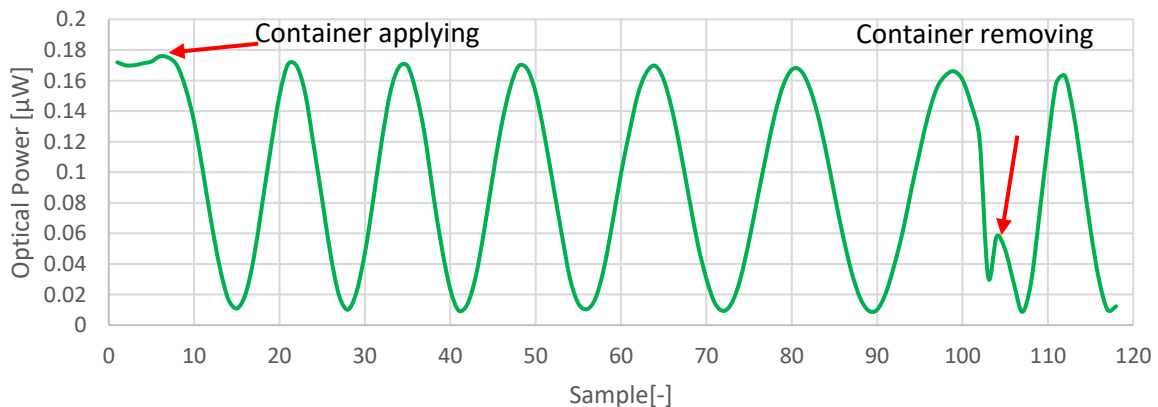
**Figure 3:** The change in optical power caused by the application of a container at 0 °C

Fig. 3 shows the power reaction to an applied container at a temperature of 0 °C. The oscillation is visible, and the maximum value was slowly ascending, as well as the mean value. After container removal, the sensor reacts immediately by the visible change of a course.



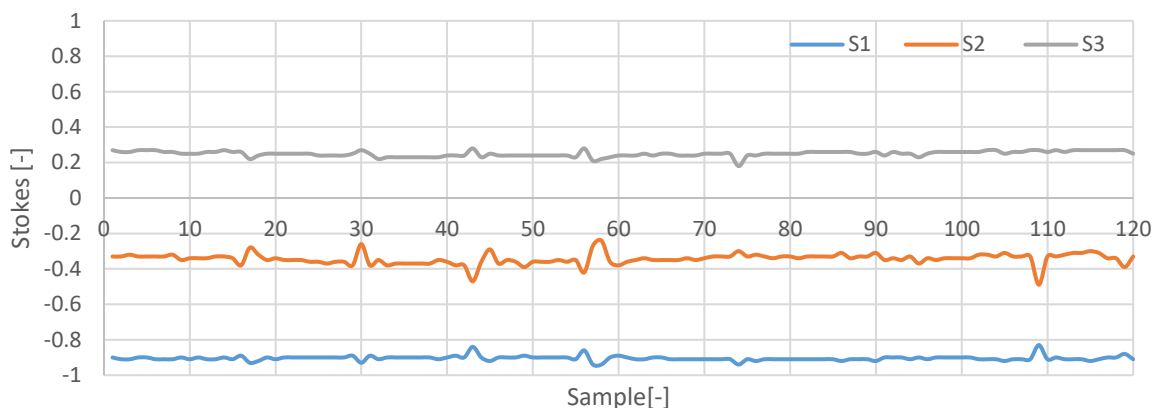
**Figure 4:** The change in optical power caused by the application of a container of water at 95 °C

Fig. 4 shows sufficient reaction in all cases. Firstly, there is a visible reaction to the operator's presence when a balanced state is interrupted. Subsequently the reaction on a container application is shown. Container removal is not so easy to notice, but the next short period confirms it. This measurement took place alongside the connected polarimeter, which captured data are displayed in Fig. 5 and Fig. 6.



**Figure 5:** The change in optical power caused by the application of a container of water at 95 °C

Fig. 5 shows the reaction in more detail, because the sample frequency is 0,5 s. The difference between power levels measured by a laser power meter, displayed in Fig. 4, and a polarimeter, captured in Fig. 5, is caused by a polarimeter, which shows only the power of the main polarization state, which is displayed in Fig. 6. The degree of polarization was around 90 %.



**Figure 6:** The change in Stokes parameters caused by the application of a container with water 95 °C

Fig. 6 captured balanced state in normalized Stokes parameters, that is caused by the property of an optical circulator to only pass y-axis polarization. Differences from the linear vertical polarization can be caused by flaws in components or couplers. Small differences from the constant course are caused by power oscillations.

#### 4. CONCLUSION

The thesis proposes a connection for power-supply to a sensor and ensures data carrying to a remote destination. All the benefits of fiber-optic polarization sensors, such as non-electricity, sufficient sensitivity, and simplicity, should be kept, with the bonus of data carrying. Data carrying is working separately as well as the sensor. Simultaneously, there were physical borders, which did not allow us to use the whole potential of the connection. Better power and polarization management can increase functionality. But the functionality of the sensor is not affected by carried data. That is proved by testing the sensor by 0°C, 36,5°C and 95°C containers. All power changes are sufficient, alongside data carrying. The only difference is in power, so it is necessary to set the receiver to a higher power range. I would like to go on researching and try to solve all imperfection and use it alongside with wavelength division multiplex.

#### ACKNOWLEDGMENT

This work was created with support by project MVČR VI3VS/678.

#### REFERENCES

- [1] M. Kyselák, Č. Vlček, J. Maschke and F. Dvořák, "Optical fibers with high birefringence as a sensor element," *2016 6th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, 2016, pp. 190-193, doi: 10.1109/ICEIEC.2016.7589717
- [2] M. Kyselák, F. Dvořák, J. Maschke and C. Vlcek. (2017). "Phase response of polarization-maintaining optical fiber to temperature changes", *Optica Applicata*, vol. 47. 635-649. 10.5277/oa170412.
- [3] J. Vavra and M. Kyselak, "Analog Signal Processing for Fiber Optic Sensor Detecting Temperature Changes," *2020 27th International Conference on Telecommunications (ICT)*, 2020, pp. 1-4, doi: 10.1109/ICT49546.2020.9239582.
- [4] M. Kyselák, Z. Vyležich, J. Vávra, D. Grenar and K. Slavíček, "The long fiber optic paths to power the thermal field disturbance sensor," *Proc. SPIE 11682, Optical Components and Materials XVIII*, 116821A (5 March 2021); <https://doi.org/10.1117/12.2575832>



# Test Bench for Experimental Analysis of Conductive Composite Materials for Aerospace Applications

N.N. Pham<sup>1</sup>, J. Leuchter<sup>2</sup>, and J. Václavík<sup>3</sup>

<sup>1</sup> Department of Microelectronics, Brno University of Technology, the Czech Republic

<sup>2</sup>Department of Aviation Technology, University of Defence, Brno, the Czech Republic

<sup>3</sup>AERO Vodochody AEROSPACE a.s., Odolena Voda, the Czech Republic

E-mail: [243756@vut.cz](mailto:243756@vut.cz), [jan.leuchter@unob.cz](mailto:jan.leuchter@unob.cz), [jan.vaclavik@aero.cz](mailto:jan.vaclavik@aero.cz)

**Abstract**—The paper presents the materials used in aerospace industry and introduces an automated test bench for investigating these materials' electrical properties. The test bench is also designed to investigate aviation electronic components, especially regarding electromagnetic compatibility (EMC). The test bench contains various measuring instruments which are connected to a computer (workstation). The designed program on the computer supervises identifying connected instruments and creating communication channels to those instruments. This program also controls those instruments to conduct required measurement or testing and obtain the measuring results. Instrument control is realized by using Standard Command for Programmable Instrument (SCPI) commands. The program is designed in MATLAB development environment and add-on application MATLAB App Designer with toolboxes Instrument Control Toolbox and Signal Processing Toolbox.

**Keywords**—Aerospace Industry, Conductive Composite Material, MATLAB, SCPI

## 1. INTRODUCTION

For decades, metals have been used in Aerospace industry as traditional materials. For example, Aluminum was a prominent material from World War I. Steel was also used but at a lower rate. From the 1970s, Titanium and Nickel began to be used [1]. The use of metals instead of wood, fabric and twine was started when engineers wanted to surmount challenges in strength and wind resistance which increased as speeds improved. On the other hand, metals had their own disadvantages. They are costly, heavy, need expensive maintenance and are prone to corrosion [1]. Therefore, composite materials are recently used as alternative materials. Main advantages of composite materials are light-weight, high-strength, high-stiffness, and easily-to-form. Obviously, the use of composite materials then can reduce the weight of aircraft, then reduces the power consumption and increase efficiency.

However, used as cover structures of aircraft, composite materials have significant problems of poor electrical and thermal conductivities. During the operation, aircraft could face problems such as ice accumulation and lightning strikes which are dangerous for the aircraft. Therefore, to protect the aircraft, materials for surface structures must have good electrical and thermal conductivity. In terms of EMC, small electrical resistivity is also required for good electromagnetic interference (EMI) shield [2]–[4].



**Figure 1:** Aero L-39NG light jet trainer project [5]

## 2. CONDUCTIVE COMPOSITE MATERIALS

As mentioned above, the significant problems of composite materials are poor electrical and thermal conductivities. The solution for these problems is called Conductive Composite Material (CCM). A CCM usually consists of conductive fillers in a nonconductive matrix. These fillers manifest a large scale of electrical conductivities needed for intended applications. Apparently, the electrical conductivity of a CCM depends on the volume fraction of conductive filler. When the amount of conductive filler increased, the filler particles begin to move continuously, free electrons in the composite can move easily by contacting each other, then the electrical conductivity will be increased to higher levels [1].

Concerning materials for conductive filler, traditional metals such as Aluminum, Copper, Silver, and Nickel are promising options. Another superior option is using carbon-based materials. In comparison with traditional metals, carbon-based materials have much lower weight, higher electrical conductivity, and higher thermal conductivity (see Table I, II). Therefore, many manufactures have been using CCM with carbon-based conductive fillers for their products. For example, to modernize the new generation L-39NG jet trainer (see Fig. 1) from its L-39s legacy, different parts made of composite materials are used by Aero Vodochody AEROSPACE a.s. to improve efficiency [5]. The cover structures are redesigned using composite materials with carbon-based conductive fillers.

**Table I:** Properties of traditional metals [1], [6]–[8]

<b>Metals</b>	<b>Density</b> <b>[g/cm<sup>3</sup>]</b>	<b>Electrical Conductivity</b> <b>[10<sup>6</sup> S/m]</b>	<b>Thermal Conductivity</b> <b>[W/mK]</b>
Aluminum	2.69	38.2	200
Copper	8.93	59.8	400
Iron	7.87	10.3	80
Silver	10.49	68	450
Nickel	8.9	14.6	100

**Table II:** Properties of carbon-based materials [1], [6]–[8]

<b>Carbon-based materials</b>	<b>Density</b> <b>[g/cm<sup>3</sup>]</b>	<b>Electrical Conductivity</b> <b>[10<sup>6</sup> S/m]</b>	<b>Thermal Conductivity</b> <b>[W/mK]</b>
Carbon Nano Tube	1.4	3.33	2000~6000
Graphene	0.3	100	800
Graphite	2.25	0.073	100~400

## 3. DESIGN OF TEST BENCH

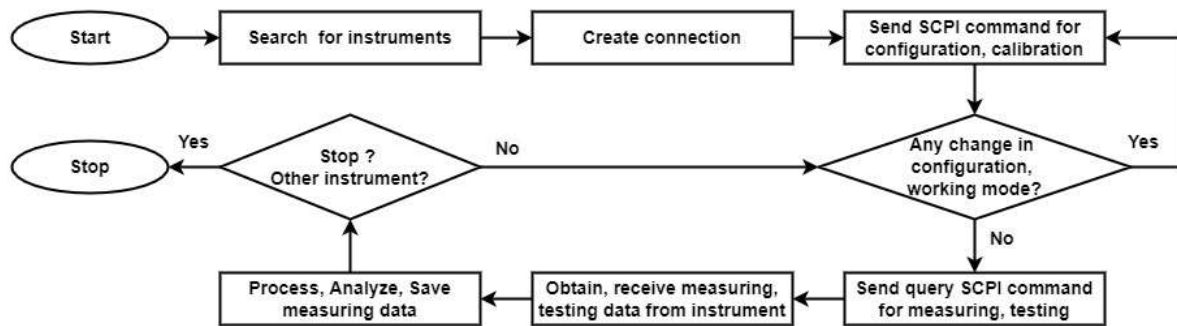
The conductivity of CCM depends not only on the amount of conductive filler as mentioned above, but also on the types, shapes, sizes, dispersions, and distribution of fillers in the composite [1]. Therefore, it is important to experimentally analyze the conductivity of various designs to verify and choose the best one for required application. Furthermore, the EMI protection issues of these materials are also interesting to investigate.

The test bench, containing measuring instruments and computer programs, aims to provide a workplace for experimental analysis of CCM. For experimental analysis, different measuring instruments are required. Manually manipulating these devices and instruments is usually slow and inconvenient. In addition, the data from experiments then must be transferred to a computer/workstation for further analysis. Therefore, it is necessary to design a computer-based system that can control several measuring instruments, obtain the measuring data from these instruments, and analyze these data.

The main part of test bench is a computer program, which must fulfill the above requirements. There are various development environments for purpose of controlling instrument such as LabView, VEE Pro, MATLAB, .... While LabView and VEE Pro are graphical-based environments which are optimized for instrument control, MATLAB is text-based environment that is designed for various engineering and scientist tasks, including instrument control. Therefore, to control instruments fast and

easily, LabView or VEE Pro are the best choices. In contrast, the computation ability of MATLAB is much better than LabView and VEE Pro. The graphical-based property of LabView and VEE Pro is not proper for a large complex project [9], [10]. MATLAB also provides a large number of toolboxes for data analysis, signal analysis and so on. It is typical to use LabView and VEE Pro for instrument control and MATLAB for data computation tasks. However, such a system is quite bulky, and expensive to purchase all the environments.

Our whole designed computer program is developed in MATLAB environment. GUI, which is required to simplify manipulation, is designed by using add-on application MATLAB App Designer, which allows us to create GUI by dragging and dropping visual components and programing their behavior. For instrument control tasks, the Instrument Control Toolbox is used. This toolbox helps the computer to connect to measuring and testing instruments via various communication protocols and standards. For data analyzing tasks, the Signal Processing Toolbox and DSP System Toolbox are used. In some special test benches, for example, test bench with thermal camera, Image Processing Toolbox is also available. The flowchart of MATLAB program is shown in Fig. 2 below.



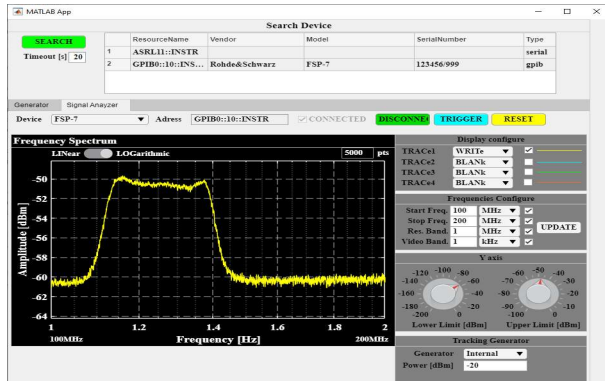
**Figure 2:** Flowchart of MATLAB program

The instrument control is conducted by using Standard Commands for Programmable Instruments (SCPI). In MATLAB program, SCPI commands are “encapsulated” in MATLAB commands, which also defined which instruments will receive which SCPI commands. The data responded by instruments are also obtained by MATLAB. These data are then visualized in GUI and processed, analyzed, or saved to computer for further working according to user’s requirements. An example of MATLAB program with encapsulated SCPI commands for HM8118 LCR Bridge is shown in Fig.3. Designed GUIs for some instruments are introduced in Fig. 4. Results of some data processing done by designed test bench are shown in Fig. 5.

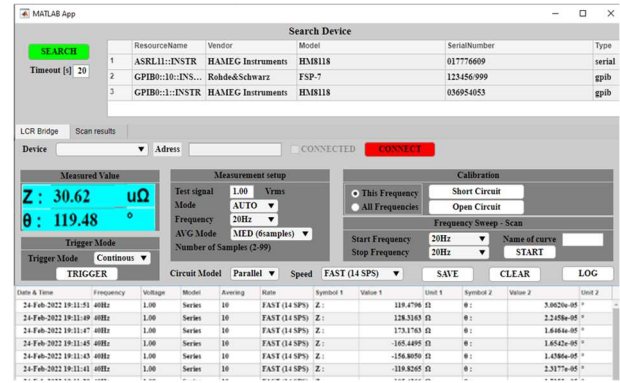
```

% Connect to the instrument
app.LCRBridge = visadev(app.LCR_AddressEditField.Value);
% Waiting for connecting
while isempty(app.LCRBridge)
end
% Configure the terminator
configureTerminator(app.LCRBridge, "CR");
% Change the display
app.LCR_CONNECTButton.Text = 'DISCONNECT';
app.LCR_CONNECTButton.BackgroundColor = 'Green';
app.LCR_CONNECTEDCheckBox.Value = 1;
% Choose the Trigger Mode
[~, Locb] = ismember(app.LCR_TriggerDropDown.Value, app.LCR_TriggerDropDown.Items);
writeline(app.LCRBridge, "MMOD " + num2str(Locb-1))
% Choose the Voltage and Frequency for the tests
app.LCR_Freq = app.LCR_FrequencyDropDown.Value;
app.LCR_Volt = app.LCR_TestsignalEditField.Value;
[~, Locb] = ismember(app.LCR_FrequencyDropDown.Value, app.LCR_FrequencyDropDown.Items);
% Set the measurement
writeline(app.LCRBridge, "FREQ " + app.LCR_Frequencies(Locb));
writeline(app.LCRBridge, "VOLT " + app.LCR_TestsignalEditField.Value);
writeline(app.LCRBridge, "PMOD " + app.LCR_mode);
writeline(app.LCRBridge, "CIRC " + app.LCR_ModelNumber);
writeline(app.LCRBridge, "RATE " + app.LCR_RateNumber);
writeline(app.LCRBridge, "MMOD " + app.LCR_TriggerNumber);
% set measurement signal's frequency
% set measurement signal's amplitude
% set measurement function
% set measurement model
% set measurement speed
% set trigger type
  
```

**Figure 3:** Example of MATLAB and SCPI code

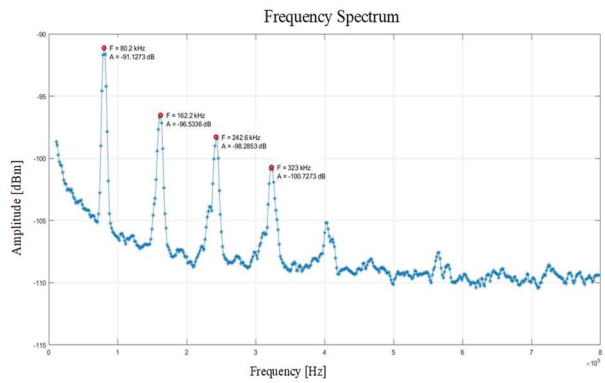


a)

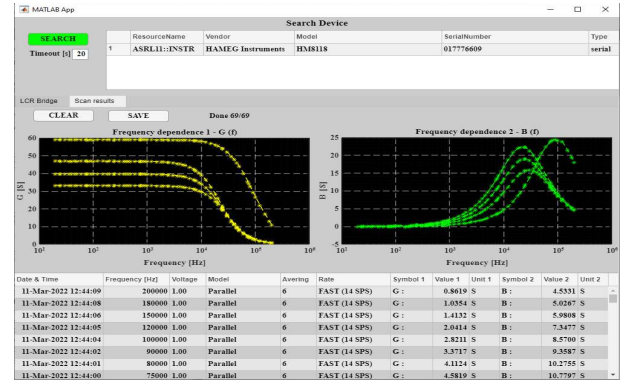


b)

**Figure 4:** MATLAB GUI for a) Signal analyzer and b) LCR Bridge



a)



b)

**Figure 5:** Results of data processing in MATLAB

Fig. 5a shows frequency spectrum of radiated noise from a switching AC/DC converter. The noise was sensed by an antenna which is connected to a signal spectrum analyzer. The signal spectrum analyzer operates under the control of test bench. Test bench obtained the measuring data from signal analyzer and calculated frequency, amplitude of four first harmonics. From the visualized data we can easily conclude that the tested AC/DC switching converter operate at a frequency of 80 kHz. Further analyses can be also realized using MATLAB functions and toolboxes. This test is simple but important regarding EMI shielding.

Fig. 5b shows the impedance of electrical conductivity of different samples of CCM in frequency. The electrical conductivity is measured by HM8118 LCR Bridge in a whole frequency range from 20 Hz to 200 kHz. The GUI visualized the data in a graph, which is more understandable than in a table. The electrical conductivity is approximately constant with frequencies below 2 kHz. Then, it rapidly decreases with the rise of frequency. This behavior of electrical conductivity then is analyzed by manufacture and designer of measured samples to improve the designs

#### 4. CONCLUSION

The paper briefly presents the trend of using Conductive Composite Materials in Aerospace applications. The main part of the paper introduces design of a testbench for experimental investigating these new aerospace materials. The test bench is developed in MATLAB environment, which can control instruments and have a large number of functions and toolboxes for data processing, analyzing, and visualizing. The functions of test bench are also verified in different measurements and testing in University of Defence in Brno, the Czech Republic. The test bench fulfills all the requirements.

The future work will be experimental analysis of aerospace composite materials and aviation components using the designed test bench. The result of experimental analysis will be used to design new aerospace and aviation structures and components with high performance and reliability.

## REFERENCES

- [1] B. Alemour, O. Badran, and M. R. Hassan, "A Review of Using Conductive Composite Materials in Solving Lightning Strike and Ice Accumulation Problems in Aviation," *Journal of Aerospace Technology and Management*, 2019, doi: 10.5028/jatm.v11.1022.
- [2] Institute of Electrical and Electronics Engineers, IEEE Aerospace and Electronic Systems Society, American Institute of Aeronautics and Astronautics, and IEEE Computer Society. Design Automation Technical Committee, *36th DASC, Digital Avionics Systems Conference, 2017 proceedings papers: Sunday-Thursday, September 17-21, St. Petersburg, Florida, USA*.
- [3] S. Yang, K. Lozano, A. Lomeli, H. D. Foltz, and R. Jones, "Electromagnetic interference shielding effectiveness of carbon nanofiber/LCP composites," *Composites Part A: Applied Science and Manufacturing*, vol. 36, no. 5, pp. 691–697, May 2005, doi: 10.1016/j.compositesa.2004.07.009.
- [4] P. Steffan, J. Stehlik, and R. Vrba, "Composite Materials for Electromagnetic Interference Shielding".
- [5] AERO Vodochody AEROSPACE a.s., "L-39NG program remains a key priority to AERO," <http://www.aero.cz/>.
- [6] Z. Han and A. Fina, "Thermal conductivity of carbon nanotubes and their polymer nanocomposites: A review," *Progress in Polymer Science (Oxford)*, vol. 36, no. 7, Elsevier Ltd, pp. 914–944, 2011. doi: 10.1016/j.progpolymsci.2010.11.004.
- [7] T. Kuilla, S. Bhadra, D. Yao, N. H. Kim, S. Bose, and J. H. Lee, "Recent advances in graphene based polymer composites," *Progress in Polymer Science*, vol. 35, pp. 1350–1375, 2010, doi: 10.1016/j.progpolymsci.2010.07.005.
- [8] "Material Properties | Website about Elements and Materials." <https://material-properties.org/> (accessed Mar. 12, 2022).
- [9] T. Tašner, D. Lovrec, F. Tašner, and J. Edler, "of ENGINEERING-HUNEDOARA, ROMANIA 389 1-2."
- [10] C. A. Cansalar, E. Mavis, and C. Kasnakoglu, "Simulation time analysis of MATLAB/Simulink and LabVIEW for control applications," in *Proceedings of the IEEE International Conference on Industrial Technology*, Jun. 2015, vol. 2015-June, no. June, pp. 470–473. doi: 10.1109/ICIT.2015.7125143.

# Horticulture IoT module

M. Šimek, J. Arm

FEEC, Brno University of Technology, Czech Republic

E-mail: [xsimek29@vut.cz](mailto:xsimek29@vut.cz), [arm@vut.cz](mailto:arm@vut.cz)

**Abstract**— This paper deals with the measurement of soil parameters and growing conditions. The processes required to continuously monitor and ensure ideal growing conditions are quite time consuming. Therefore, the aim of this work will be to design and implement a module that eliminates the need for human interaction in monitoring important parameters and can be used as a part of a complex autonomous cultivation system. The main requirements of the proposed module include energy self-sufficiency and wireless communication with the master device.

**Keywords**—soil bulk EC, volumetric water content, ESP-32D, LoRa

## 1. INTRODUCTION

Maintaining ideal growing conditions throughout the entire plant cycle is an extremely challenging process that requires constant human interaction in the absence of supporting automation.

The subject of this work is to design a module that will be able to monitor the necessary parameters of the growing process in a defined period of time. According to these measured values, processes such as fertilization, humidification, ventilation, and irrigation of the area in which the sensor is located can be controlled. The following requirements and parameters that have been determined for the proposed module: energy self-sufficiency, wireless communication with the master system, local access to measured values, measurement of soil parameters (bulk EC, VWC and temperature), measurement of ambient air parameter (temperature and relative humidity) and calibration of bulk EC and VWC sensors.

## 2. SOIL WATER CONTENT

Soil water content (SWC) represents the amount of water contained in a soil sample with certain properties. There are two main basis for determining SWC: gravimetric and volumetric.

Gravimetric water content ( $\theta_g$ , GWC) is defined by mass of water ( $M_w$ ) per mass of dry soil ( $M_s$ ). It can be measured by weighing a soil sample ( $M_t$ ), then removing the water by drying and weighing the dried soil. [1]

$$\theta_g = \frac{M_w}{M_s} = \frac{M_t - M_s}{M_s} [-] \quad (1)$$

Volumetric water content ( $\theta_v$ , VWC) is defined as ratio of the water volume ( $V_w$ ) to the total volume of soil sample ( $V_t$ ). Gravimetric water content can be converted to volumetric by knowing the density of water ( $\rho_w$ ) and bulk density of soil sample ( $\rho_b$ ). [1]

$$\theta_v = \frac{V_w}{V_t} = \theta_g \frac{\rho_b}{\rho_w} [-] \quad (2)$$

Another important variable is water content at saturation ( $\theta_s$ ) which gives us information about how much water content the soil sample can hold. It can be determined according to the following formula.

$$\theta_s = 1 - \left( \frac{\rho_b}{\rho_s} \right) [-], \quad (3)$$

where  $\rho_s$  denotes density of solids (which is typically  $2.65 \text{ g/cm}^3$  for mineral soil [2])



Volumetric water content can be determined indirectly by measuring of the following soil parameters:

- Electrical resistance – inserting the electrodes into the soil and exciting them with a current source creates a voltage drop from which the electrical conductivity or resistance of the soil can be calculated. Increasing water content resolves in decreasing resistance across measuring electrodes. However, an increase in the soil's conductivity results in a decrease in its resistivity even if the water content has not changed at all. These facts make this method unsuitable for accurate measurement of soil water content.
- Dielectric permittivity – the dielectrical constant of the soil ranges from 2–3 depending on the physical, mechanical, and chemical properties of the soil, whereas the dielectrical constant of water is about 80. The resulting permittivity of the soil sample is therefore strongly influenced by the amount of water it contains. To determine the permittivity of the soil sample, mainly FDR (Frequency Domain Reflectometry), TDR (Time Domain Reflectometry) and capacitive methods are used. [2]

### 3. SOIL ELECTRICAL CONDUCTIVITY

Electrical conductivity of the soil gives us an idea of the total amount of dissolved salts in a soil sample. By measuring this parameter, it is possible to prevent over-fertilization or nutrient deficiencies in the substrate and thus guarantee the correct development of the plant throughout its growing cycle.

There are many ways to describe soil EC, but the most common are following [2]:

- Bulk EC ( $\sigma_b$ ) – electrical conductivity of the bulk soil, it's strongly influenced by SWC and total dissolved salts in soil sample. All sensors which are placed directly in a soil sample measure this parameter.
- Pore water EC ( $\sigma_w$ ) – this parameter represents electrical conductivity of water in soil pores and gives us information about what conductivity the plant is experiencing. It can be measured by extracting pore water from the soil sample and measuring its EC.
- Saturation extract EC ( $\sigma_e$ ) – value represented by electrical conductivity of extracted water from saturated soil sample. Indicates the amount of salts contained in the soil sample. It is measured by adding deionised water to the soil sample until it is saturated and then extracting the water and measuring its EC. Depending on the  $\theta_v$ , it is 3 times to 9 times larger than  $\sigma_b$ .

According to the following equation, it is obvious that when the soil sample is at the saturation point, the EC of pore water and the EC of saturation extract are equal. [2]

$$\sigma_e = \sigma_w \left( \frac{\theta_s}{\theta_v} \right) [mS/cm] \quad (4)$$

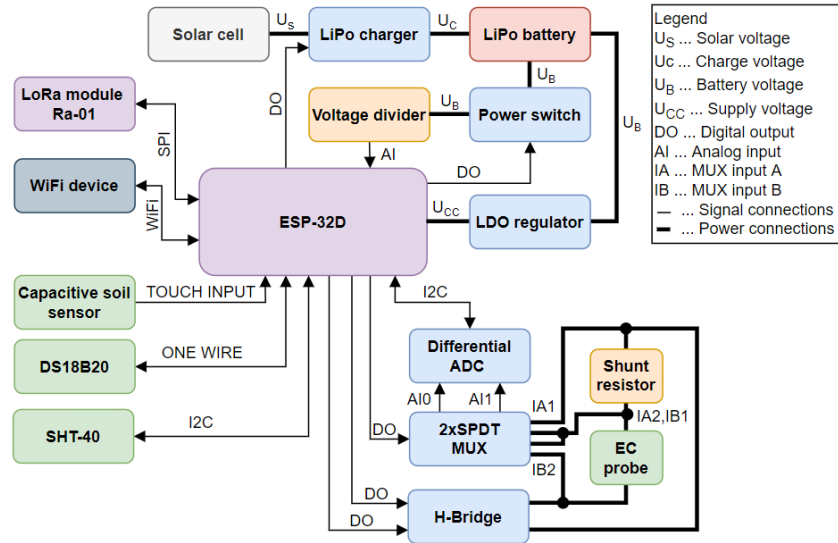
### 4. STEM-01 MODULE

In accordance with the requirements of the introduction and for the need of continuous monitoring of volumetric water content and soil bulk EC, the STEM-01 module (abbreviation for Soil moisture, Temperature and humidity, Electrical conductivity Module) was designed.

The ESP32 microcontroller was chosen as the main controller of the module, mainly due to its sufficient computational power, support for IEEE 802.11 communication standards, and wide support for device libraries. As a part of energy self-sufficiency, the module is powered by a LiPo battery which is recharged by a charger powered by a solar cell, thus ensuring its long-term operability and eliminating module maintenance in the form of battery replacement or manual recharging. For the determination of VWC, the method of measuring the capacitance of an interdigital capacitor placed in a soil sample was chosen. Capacitance measurements can be made directly using the touch inputs of the ESP32 controller. To measure soil conductivity, a circuit consisting of an ADC with differential inputs and programmable gain amplifier, a 2xSPDT multiplexer, an h-bridge, a shunt resistor and the conductivity probe itself was designed. The principle of conductivity measurement is as follows: a positive voltage pulse is generated using the h-bridge, during which the voltage drop across the shunt resistor is first measured, then the ADC input is redirected to probe electrodes using a multiplexer and the voltage drop across it is also measured. The same procedure is repeated for the generated negative voltage pulse. The resulting voltage drop across the shunt resistor is then used to calculate the current through the probe and, thanks to the known voltage drop, its resistance can be calculated. Using the calculated probe resistance, the

electrical conductance of the probe is further determined, and the final electrical conductivity is determined using the cell constant. The SHT-40 sensor is used to measure the temperature and relative humidity of the environment and a DS18B20 temperature sensor located near the EC probe for its calibration and temperature compensation of electrical conductivity.

The STEM-01 module has two modes of operation: low-power mode and local access mode. In local access mode, it is possible to connect via WiFi to the ESP32 on which a web server is running. Through this web server you can perform the initial setup of the module, monitor the current measured values and perform calibration of EC and  $\theta_v$  measurements. In low-power mode, the module wakes up cyclically from deep sleep mode, sends the measured data via LoRa technology to the gateway and go back to sleep. The gateway then sends the received data to the MQTT broker into the preset topic. Block diagram of the module is shown in the figure 1.



**Figure 1:** Block scheme of module STEM-01

A waterproof box was designed in Autodesk Fusion 360 for the individual components and the circuit board. The box consists of three main parts: the lid on which the solar panel is mounted, the body in which the PCB is placed and the stem which contains the EC probe, the DS18B20 sensor and the interdigital capacitor for water content measurement. Designed box and DPS are shown in figure 2.

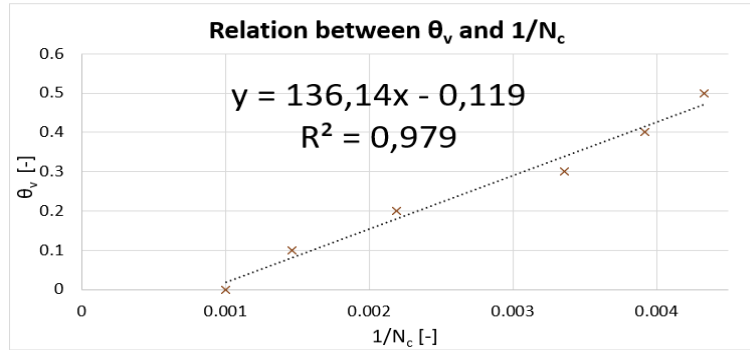


**Figure 2:** 3D enclosure design, real enclosure, PCB

## 5. CALIBRATION PROCEDURES

The water content measurement can be performed in two modes. In the first mode, the sensor can be calibrated in two steps by first measuring the number of pulses of the capacitive input  $N_c$  when the sensor is placed in air (water content equals to 0 %) and then measuring  $N_c$  for the sensor placed in water (water content equals to 100%). Based on these two values, a regression line of the water content on the inverted  $N_c$  value is then calculated. The second mode measures directly the volumetric water content. For this mode to work correctly, it needs to be calibrated as follows: a certain volume of dry soil sample (in our case 300 ml) is placed in the container, then the  $N_c$  corresponding to a  $\theta_v = 0$  is measured, then 30 ml of water is added to the container (10% increase of the  $\theta_v$ ), the soil is mixed properly and the number of

pulses of  $N_c$  is measured again. Adding water, mixing and measuring  $N_c$  is repeated until the soil sample starts to seep water. From the logged values of  $N_c$  and their corresponding  $\theta_v$  linear regression line is calculated. However, the obtained regression is valid only for soils of the same or very similar type. Therefore, if the calibration is performed on peat and the sensor is subsequently moved to a location with clay soil, this calibration process must be performed again. A sample calibration process was performed in the  $\theta_v$  range of 0 - 0.5 for a peat-like sample with  $\rho_b = 0.88 \text{ g/cm}^3$ . This resulted in the following equation illustrated in the following graph (figure 3).



**Figure 3:** Relation between  $\theta_v$  and  $1/N_c$

As already mentioned, the EC sensor measures  $\sigma_b$ . Its calibration therefore consists mainly of placing it in a reference solution of known electrical conductivity  $\sigma_c$ , where the probe conductance  $G_p$  is first measured, and the cell constant  $K$  is determined from the known temperature  $T_0$  of the solution and the  $\sigma_c$  according to the following formula.

$$K = \frac{\sigma_c [1 - 0.019(T_0 - 25)]}{G_p} [cm^{-1}], \quad (5)$$

## 6. CONCLUSION

Maintaining ideal growing conditions guarantees the final quality and quantity of the grown product and the reusability of the soil for the next growing cycles. By using the STEM-01, common cultivation mistakes such as overwatering or underwatering and overfertilization or lack of fertilizers can be avoided. In larger applications, monitoring water and maintaining ideal irrigation will result in less water consumption, higher yields, greater financial savings, and environmental sustainability. In a sample calibration of soil water content sensor, the MAPE (mean absolute percentage error) was found to be 8.2 % which is an acceptable error in volumetric water content measurement. The accuracy of the conductivity measurement depends mainly on its correct calibration. In order to suppress the polarization effect and to maintain correct results in the long term, alternating excitation was used for the EC sensor and the electrodes were chosen as titanium. The designed module will be part of a larger autonomous system that can control irrigation and fertilization based on the data provided by this module.

## ACKNOWLEDGMENT

This work was supported by Brno University of Technology and was carried out with the support of the project FEKT-S-20-6205 Research in automation, cybernetics and artificial intelligence for Industry 4.0.

## REFERENCES

- [1] Bilskie, Jim. "Soil Water Status: Content and Potential." Campbell Scientific, 2001. <https://s.campbellsci.com/documents/es/technical-papers/soilh20c.pdf>
- [2] METER. Soil electrical conductivity: The complete guide to measurements. [online] Available at: <https://www.metergroup.com/en/meter-environment/education-guides/soil-electrical-conductivity-complete-guide-measurement> [Accessed 8 March 2022].

# Measurement System for Determining Piezoelectric Charge Coefficient

O. Kašpárek<sup>1</sup>, J. Fialka<sup>1</sup>

<sup>1</sup>Department of Control and Instrumentation, Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech Republic

E-mail: [203245@vut.cz](mailto:203245@vut.cz), [fialka@vut.cz](mailto:fialka@vut.cz)

**Abstract**— The paper focuses on a system of direct measurement of piezoelectric charge coefficient. The created system is based on quasistatic method when harmonic force signal is applied to the examined piezoelectric material. A program in the programming environment LabVIEW was created for control, measurement, and processing. Thanks to the universality of this system from the perspective of hardware, this system offers higher levels of modularity than  $d_{33}$  meters available on the market.

**Keywords**— piezoelectric effect, piezoelectric charge coefficient, quasistatic method, Berlincourt method, dynamic method

## 1. INTRODUCTION

Determining the parameters of piezoelectric materials is an integral part of the design of applications using these materials. It is necessary to know the exact properties of the behavior of materials in given applications, and in addition, each application may place different requirements on the piezoelectric material used. The piezoelectric charge coefficient is an important parameter for the selection of a suitable piezoelectric material as an actuator. Furthermore, it is one of the parameters evaluated during control measurements in the production of piezoelectric materials [1].

Not only one but several different experimental methods are given to determine the piezoelectric coefficients. When choosing a method, several factors must be taken into account, such as the size and shape of the sample, available laboratory equipment and more. Examples of possible methods are direct methods and frequency method. [2].

Within the direct methods of measuring the piezoelectric charge coefficient, it is possible either to use dedicated instruments for measuring charge coefficients called  $d_{33}$  meters or to create your measurement system based on appropriate standards. However, the standards contain basic instructions and do not standardize measurement methods.

This work aims to create a system for measuring piezoelectric charge coefficient using laboratory equipment that can replace  $d_{33}$  meters and which can offer higher levels of modularity than these instruments available on the market. In practice, the dedicated instruments are manufactured for measuring  $d_{33}$  and  $d_{31}$ , but in this case, a device was created only for measuring the parameter  $d_{33}$ . For  $d_{31}$  it would be necessary to measure on the electrodes perpendicular to the applied force.

## 2. QUASISTATIC METHOD

The measurement of the charge induced by mechanical stress or the so-called Berlincourt / quasi-static / dynamic method is the manufacturer's most widely used method for measuring the polarity and piezoelectric charge coefficient  $d_{33}$  [3]. Compared to the frequency method, the quasistatic method is not limited by the shape and dimensions of the specimen [2].

The piezoelectric charge constant is determined through measurement of the electrical response of the material under the action of an on-off or dynamic external force. Description of the quasi-static method leads into two methods based on the direct or inverse piezoelectric effect. That is, to measure charge or deflection. In this case, a method based on a direct phenomenon is used where the charge is measured depending on the excitation force. The method based on the direct phenomenon is described in [4] and the method based on the inverse phenomenon in [5].

European and American standards dealing with measuring methods for determining the coefficients of piezoelectric materials, namely EN50324-2 and IEEE Std 176-1987, describe this method in general. These standards do not specify exactly how the measurement takes place. The determination of the charge constant can therefore take place in two ways. Either through comparison of the generated charge with the force signal that the measured piezoelectric element stresses, or through comparison of the generated charge with the charge that is generated by a piezoelectric element with known parameters [3] [4].

Calculation of charge coefficient  $d_{33}$  when determining with force measurement [3]:

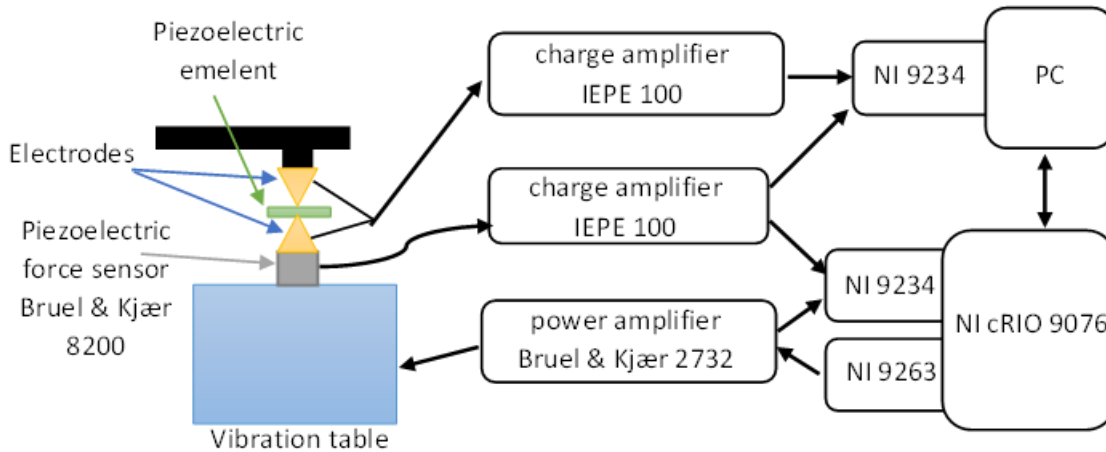
$$d_{33} = \frac{Q_3}{F_3} [\text{C} \cdot \text{N}^{-1}], [\text{m} \cdot \text{V}^{-1}] \quad (1)$$

where  $Q_3$  denotes generated charge of the measured element in the coordinate direction 3 and  $F_3$  is the magnitude of the loading force in the coordinate direction 3.

The charge coefficient  $d_{33}$  is mentioned in the equations and the description of the method, but this method is not limited to only this direction. The method can be used to measure charge coefficients in other directions, for example  $d_{13}$ ,  $d_{31}$ , and  $d_{15}$ . The only changes in such cases are a change in the settling of the sample and further addition of measuring electrodes to other corresponding sides of the material under examination.

### 3. MEASUREMENT SYSTEM

As was mentioned before, created system of measurement shown in Fig. 1 is based on the quasi-static method when the piezoelectric charge coefficient is determined through comparison of the generated charge with the dynamic force signal that the measured piezoelectric element stresses. This dynamic force has shape of a sinusoid with static frequency of 110 Hz and with user-configurable amplitude. This frequency was chosen for several reasons. Firstly, it fulfills the requirements mentioned in standard EN50324-2 which states that frequency should be in limits between 20 and 1000 Hz [3]. Secondly, this frequency is commonly used in dedicated  $d_{33}$  meters and thirdly, it is not a signal of 50 Hz or its higher harmonics, so the signal is separable from noise from the electrical network thanks to spectral analysis [6]. Furthermore, the dynamic force was chosen because it eliminates the influence of the pyroelectric effect on the measurement [2].



**Figure 1:** System devices diagram

From the devices seen in Fig. 1, only the voltage output module NI 9263, the acquisition modules NI 9234 and the CompactRIO controller NI cRIO 9076 have to be these exact types. Other devices can be replaced by other devices and their parameters alter in the control program. Thanks to this, the system of measurement offers high levels of modularity and flexibility from the perspective of available laboratory equipment.

The control software for this system is divided into two connected programs in the programming environment LabVIEW. The first program running on the CompactRIO controller NI cRIO 9076 is in

charge of generating the required force signal through altering the power supply to the vibration table and in charge of safety controls to protect the hardware from possible damage. Thanks to the integrated controller NI cRIO-9076, constant control and its independence from the system resources of the computer from which the measurement is started is ensured.

Although this separation of programs creates inefficiency in the usage of hardware, i.e. more modules are needed, this configuration avoids possible problems with accessing results of measurements which are stored in TDMS file. Both programs could be united and run-on integrated controller, but it would increase the required user's knowledge and experiences with integrated controllers.

The dynamic force signal within the measuring system is created by the action of the vibrating armature of the vibrating table on the structure ensuring the prestressing of the force. A measured piezoelectric element is connected between copper point electrodes which are placed on the vibrating armature and the structure. Furthermore, a force sensor is placed between the vibrating armature and the lower electrode to provide feedback for the program ensuring vibration control.

The second program which is running on the computer is ensuring data acquisition and data saving. The data acquisition is started when the program in CompactRIO controller transmits information that the RMS value of the force signal stabilizes at the value entered by the user with a maximum error of 0,5%.

From the measured waveforms of the force and charge from the tested piezoelectric element the amplitude spectrums are determined to obtain the amplitudes of the frequency component 110 Hz. Thanks to the use of the fast Fourier transform, and thus the spectral analysis, the calculation is not subject to error due to the presence of the 50 Hz harmonic component and its higher harmonics, which enter the measurement from the electrical network. To determine the amplitude of the required component as accurately as possible, the properties of the windows were used in the calculation of the fast Fourier transform, namely the Flat top window, which is optimal for the most accurate calculation of the amplitudes of the harmonic components of the signals. Thanks to the use of windows, the occurrence of spectral leakage is also prevented, i.e. the formation of non-existent frequency signal components in the spectrum.

#### 4. RESULTS AND DISCUSSION

The functionality of the system was tested on a set composed of multiple samples of soft piezoceramics NCE51 of different dimensions and shapes with piezoelectric charge coefficient  $d_{33} = 443 \text{ pC} \cdot \text{N}^{-1} \pm 5\%$ . This value of charge coefficient is stated by the producer for this material, but the actual value varies in different batches and is also dependent on the polarization of the specimen and his age. The measured results and reference values, i.e. values of piezoelectric charge coefficient measured by  $d_{33}$  meter ZJ-3C, are shown in Table 1 and Table 2.

**Table 1:** Measured  $d_{33}$  of soft piezoceramics NCE51 in the direction of polarization

dimensions [mm]	25x5x0,5	Ø3x5,7	Ø3,5x14	Ø20x2	Ø18x2	Ø16x2	Ø10x2
shape	thin plate	cylinder	cylinder	thin disc	thin disc	thin disc	thin disc
$d_{33}$ [pC·N <sup>-1</sup> ] (reference)	500	462	368	452	474	460	459
$d_{33}$ [pC·N <sup>-1</sup> ]	453	440	343	417	472	440	437
$\delta_r$ [%]	10,4	5,0	7,3	8,4	0,4	4,5	5,0

**Table 2:** Measured  $d_{33}$  of soft piezoceramics NCE51 in the opposite direction than polarization

dimensions [mm]	25x5x0,5	Ø3x5,7	Ø3,5x14	Ø20x2	Ø18x2	Ø16x2	Ø10x2
shape	thin plate	cylinder	cylinder	thin disc	thin disc	thin disc	thin disc
$d_{33}$ [pC·N <sup>-1</sup> ] (reference)	478	463	370	452	476	459	465
$d_{33}$ [pC·N <sup>-1</sup> ]	447	443	352	419	463	436	443
$\delta_r$ [%]	6,9	4,5	5,1	7,9	2,8	5,3	5,0



Note that the piezoelectric charge coefficient should not be affected by flipping the sample, i.e. affected by polarization direction. But the coefficient is not constant on the whole area of the samples so changes of the gripping point can influence the measured value. The differences between the charge coefficient of the same sample can be most probably attributed to this phenomenon.

## 5. CONCLUSION

For determining piezoelectric charge coefficient system of measurement based on the quasi-static method was created. This system thanks to adjustable parameters offers high degrees of modularity from the hardware perspective. The consequence is a system suitable for determining various effects on the measurement such as the shape of electrodes and which can be put together from common laboratory equipment instead of acquiring a special device called  $d_{33}$  meter. Furthermore, the system includes various precautions for ensuring safety of used devices, such as protection against current overexcitation of the vibrating table, to ensure that the equipment will not be damaged even when the system is used by less skilled operators.

This system was tested on a set of soft piezoceramics NCE51 with various shapes and dimensions, the results can be found in Table 1 and Table 2. Measured values are compared with reference piezoelectric charge coefficients which were determined with  $d_{33}$  meter ZJ-3C. The largest measurement error was determined for the sample in the shape of a thin plate with dimensions 25x5x0,5 mm. For the rest of the samples values of relative deviations are between 0,4 % and 8,4 %. It is worth noting that values of the piezoelectric charge coefficient determined by this measurement system are lower than reference values determined by  $d_{33}$  meter. Some error can be attributed to the fact that change of location of the gripping point influence the measured value. Furthermore, the generation of force contributes to the measurement error due to small changes of the effective value of the force signal during measurement, i.e. errors of the effective value of force signal are tolerated up to 0,5%.

Based on the results of testing measurements, it can be concluded that the created measurement system gives similar results as dedicated instruments for determining piezoelectric charge coefficients and offers an alternative to purchasing  $d_{33}$  meters for measurements that do not require high levels of accuracy.

## REFERENCES

- [1] *Piezoelectric ceramics: principles and applications*. Mackeyvill: APC International, 2002. ISBN 09-718-7440-9.
- [2] *IEEE Standard on Piezoelectricity*. IEE Standart 176-1987. New York, USA: The Institute of Electrical and Electronics Engineers, Inc, 1988.
- [3] EN 50324-2. *Piezoelectric properties of ceramic materials and components - Part 2: Methods of measurement*. Brusel: CENELEC, 2002.
- [4] WAANDERS, J.W. *Piezoelectric Ceramics: Properties and Applications*. 1. Philips Components, 1991, 91 s. OCLC: 45085184.
- [5] HUANG, Zhaorong, Qi ZHANG, Silvana CORKOVIC, Robert DOREY a Roger W. WHATMORE. Comparative Measurements of Piezoelectric Coefficient of PZT Films by Berlincourt, Interferometer, and Vibrometer Methods. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* [online]. 2006, 53(12), 2490-2523 [cit. 2022-03-10]. ISSN 0885-3010, doi:10.1109/TUFFC.2006.175
- [6] STEWART, Mark, Will BATTRICK a Markys CAIN. *Measurement Good Practice Guide No. 44: Measuring Piezoelectric  $d_{33}$  coefficients using the Direct Method* [online]. Teddington, Middlesex, United Kingdom: National Physical Laboratory, 2001 [cit. 2022-03-06]. ISSN 1368-6550, <https://eprintspublications.npl.co.uk/2768/1/mgpg44.pdf>

# A patch antenna array for IFF band

J. Mička<sup>1</sup>, J. Láčik<sup>2</sup>, M. Martinec<sup>3</sup>

<sup>1</sup>Author, Czech Republic, Brno University of Technology

<sup>2</sup>Supervisor, Czech Republic, Brno University of Technology

<sup>3</sup>Supervisor, Czech Republic, RAMET a.s.

E-mail: [203290@vut.cz](mailto:203290@vut.cz), [lacik@vut.cz](mailto:lacik@vut.cz), [mmartinec@ramet.as](mailto:mmartinec@ramet.as)

**Abstract**—This paper describes a design of an antenna array for the IFF band. The array consists of four antenna elements arranged as a  $2 \times 2$  matrix. A single antenna is designed as two stacked patches, fed by a slot. The feeding network is designed as an asymmetric stripline. Simulations of the proposed antenna array give good simulation results especially VSWR lower than 1.5.

**Keywords**— antenna array, stacked patches, aperture-coupled, stripline

## 1. INTRODUCTION

IFF (Identification Friend or Foe) technology was developed by the army for better recognition between own and enemy units, especially planes. Communicating objects are called an interrogator (e.g., control tower) and a transponder (e.g., plane). The interrogator sends a request to the transponder, which sends back a reply message. Two frequency bands are used for communication. Center frequencies of these bands are 1030 MHz and 1090 MHz. Each band has a 20 MHz bandwidth [1].

The antenna is a very important part of a transmitter, which achieves a good and effective transformation of an electromagnetic field between a transmission line and a free space [6]. The required antenna must reach the following goals at given frequency bands: VSWR lower than 1.5, gain at least 12 dBi, beam scanning angle  $10^\circ$  in the elevation. The antenna array must be smaller than  $1 \text{ m} \times 0.5 \text{ m}$ .

Good design is shown in [2], where a single patch fed by slot is used as a single antenna element. Thick foam with a low dielectric constant is used as a dielectric material between patch and ground plane. This single element achieves a large impedance bandwidth across both desired bands with  $S_{11}$  lower than  $-18 \text{ dB}$  (VSWR  $< 1.288$ ). But whole antenna array has the  $S_{11}$  lower only than  $-12 \text{ dB}$  (VSWR  $< 1.671$ ), it does not satisfy our requirement. The mentioned paper presents a feeding network that establishes sum and difference radiating beams. In this paper, only the sum radiating pattern is considered.

## 2. ANTENNA ARRAY

The antenna array is composed of four single elements as  $2 \times 2$  matrix (see Figure 1). Distance between centres of neighbouring elements is approximately  $0.6 \lambda_0$ . Where  $\lambda_0$  is the wavelength in free space. The feeding network is designed as asymmetric stripline. So the feeding lines are stacked between two substrate layers of different thicknesses (see Figure 2). The thicknesses of proposed substrates are  $h_{s1} = 1.524 \text{ mm}$  and  $h_{s2} = 0.508 \text{ mm}$ . The type of considered substrate is the Kappa<sup>®</sup> 438 with a relative permittivity of 4.38 [3].

The stripline is covered by two ground planes from the top and the bottom sides. This makes the stripline very immune to noise signals and eliminates the back radiation of the feedline. The asymmetric structure improves the coupling between the feeding line and the slot in the upper ground plane. The back radiation from the slot causes the appearance of stripline higher-order modes. These modes could be suppressed by placing the vias between the ground planes [4]. To achieve sufficient higher-order mode suppression the vias must be placed correctly. The good answer for this problem is given in [7]. Where the vias are used to create walls, which construct a resonant cavity with basic resonant mode  $TE_{10}$ . The length of the cavity  $L_w$  is approximately half of the wavelength in a dielectric substrate. The width  $W_w$  was estimated experimentally.

The feeding network uses T-junctions to split power from the input to the antennas. This has an advantage against the Wilkinson power splitter in absence of the discrete resistor. To achieve a good impedance match an impedance matching network must be considered. Because at the splitting point

the impedance is different from the characteristic impedance of the input line. The proposed impedance matching network is composed of three quarter-wavelength transformers. The characteristic impedance of each quarter-wave transformer was estimated by formula 1:

$$Z_n = \sqrt{Z_{n-1}Z_{n+1}}. \quad (1)$$

Where  $Z_n$  is the characteristic impedance of the transformer,  $Z_{n-1}$  is the characteristic impedance of previous line segment and  $Z_{n+1}$  is the characteristic impedance of following segment [8]. Characteristic impedance depends on the width  $W_f$  of the transmission line. The width for every segment was estimated using Ansys HFSS. These values are shown in the table I.

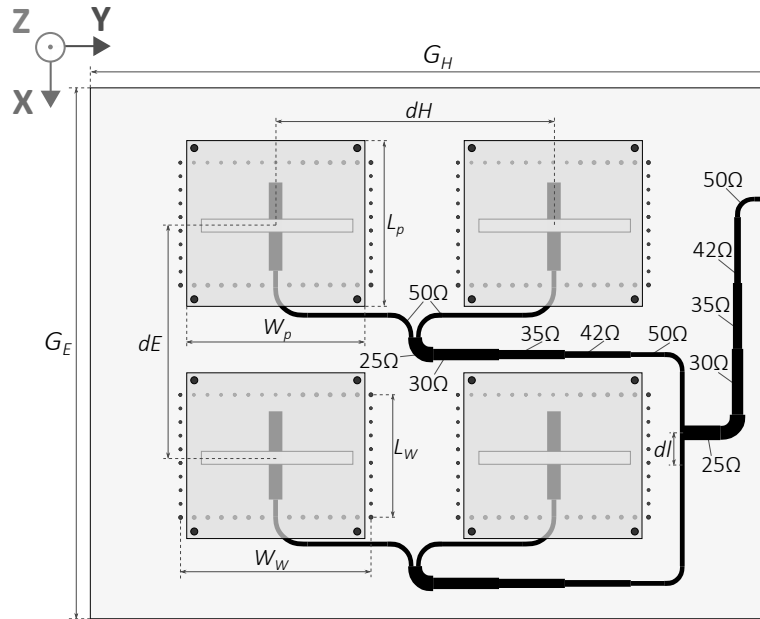
The phase shift between elements in the XZ-plane must be carried out to shift the main lobe maximum in the XZ-plane. This is done by the length difference  $dl$  of the feeding lines.

**Table I:** The widths of transmission lines [mm]

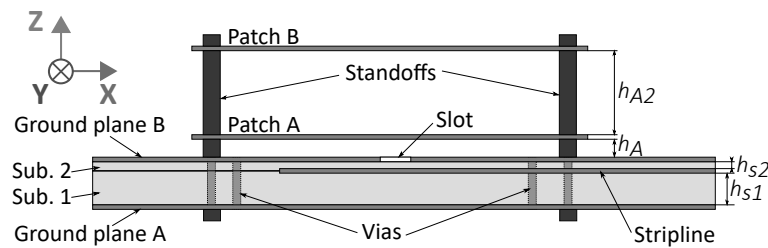
$Z_0[\Omega]$	50	42	35	30	25
$W_f[\text{mm}]$	0.7	0.98	1.35	1.67	2.15

**Table II:** Overall dimensions [mm]

$G_E$	$G_H$	$dE$	$dH$	$L_W$	$W_W$	$h_{s1}$	$h_{s2}$	$h_A$	$h_{A2}$	$L_p$	$W_p$	$dl$
380	430	174	174	70	120	1.524	0.508	4	11	113.5	113.5	8



**Figure 1:** Antenna array and feeding network sketch (top view)



**Figure 2:** Single antenna element sketch (side view)

### 3. SINGLE ANTENNA ELEMENT

The single element (see Figure 2) is designed as a multilayer with two vertically stacked patches. This type of antenna gives very good results mainly low VSWR and high gain at large frequency bandwidth [9] [10] [11].

Proposed patches are identical and square-shaped. Patches are held by non-metallic standoffs, which makes an air layer of height  $h_A$  between ground plane B and patch A, and height  $h_{A2}$  between patches. The lower patch is coupled to the feeding line by the slot. This coupling method gives a manufacturing advantage in comparison to probe feeding.

### 4. SIMULATION RESULTS

The antenna array was simulated by the Ansys HFSS software. Figure 3 shows VSWR at the input of the single antenna and designed antenna array. The antenna array has a lower VSWR than 1.4 at both frequency bands. Figure 4 shows a radiating pattern in the XZ-plane. It is visible that the main lobe maximum is well shifted to  $10^\circ$ . The radiating pattern in the YZ-plane is shown in figure 5. The cross-polarization is lower than  $-30$  dBi in the XZ-plane and lower than  $-50$  dBi. The radiation plots are represented as realized gain. Maximum gain is 12.52 dBi at 1030 MHz and 13 dBi at 1090 MHz.

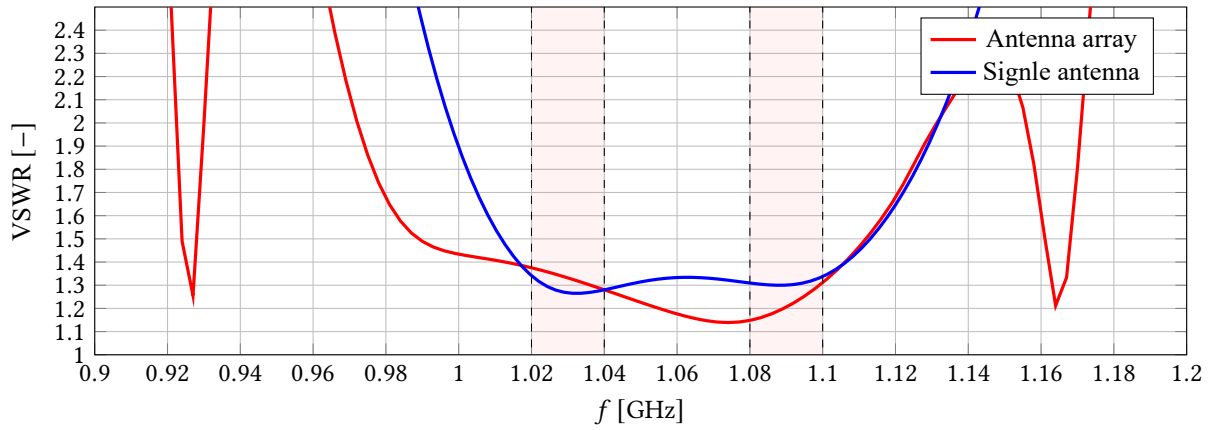


Figure 3: VSWR at array input

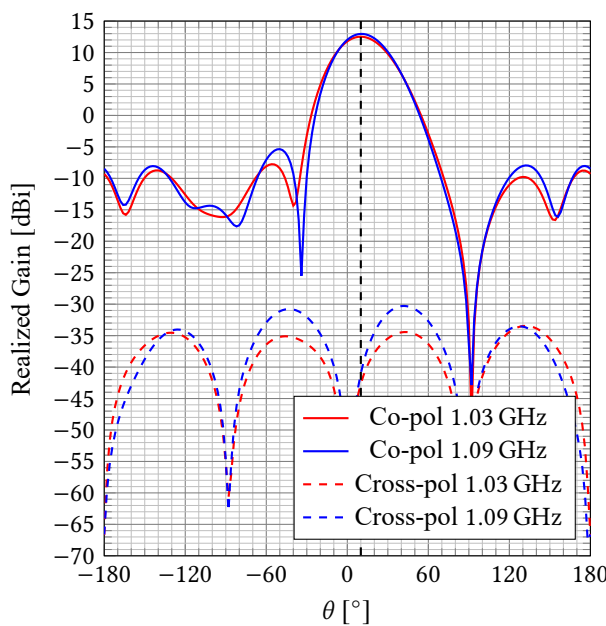


Figure 4: Radiating pattern (XZ-plane)

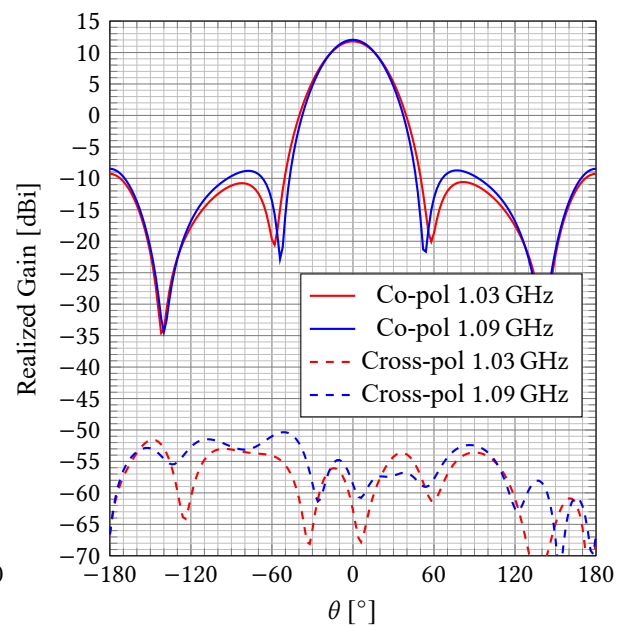


Figure 5: Radiating pattern (YZ-plane)

## 5. CONCLUSION

This paper presents the designed antenna array for the IFF band. The array consists of four antenna elements designed as multilayer stacked patches coupled with slots. Maximum dimensions of the whole array are 430 mm  $\times$  380 mm.

The array is well impedance matched, the VSWR is lower than 1.4 at both frequency bands. The scan angle was set to 10° by adjustment of feedline length between antenna elements. Maximum gain slightly varies with frequency but is higher than 12 dBi at both frequency bands.

The designed array satisfies all specified requirements. The next goal is the creation of the designed antenna and verification if the physical realization matches the proposed design.

## ACKNOWLEDGMENT

This work was supported by the Internal Grant Agency of Brno University of Technology, project no. FEKT-S-20-6526.

## REFERENCES

- [1] "IFF - IDENTIFICATION - FRIEND OR FOE", Accessed: January 2, 2022. [Online]. Available: <http://www.tscm.com/iff.pdf>
- [2] M. Abdolahi, Z. Pourgholamhossein, H. M. Sadeghi and M. Fadaei, "Design and analysis performance of a new patch array antenna for SSR," 2017 Progress In Electromagnetics Research Symposium - Spring (PIERS), 2017, pp. 2539-2545, doi: 10.1109/PIERS.2017.8262180.
- [3] "Kappa® 438 Laminates", Accessed: December 12, 2021. [Online]. Available: <https://rogerscorp.com/-/media/project/rogerscorp/documents/advanced-electronics-solutions/english/data-sheets/kappa-438-laminates-data-sheet.pdf>
- [4] A. Bhattacharyya, O. Fordham and Yaozhong Liu, "Analysis of stripline-fed slot-coupled patch antennas with vias for parallel-plate mode suppression," in IEEE Transactions on Antennas and Propagation, vol. 46, no. 4, pp. 538-545, April 1998, doi: 10.1109/8.664118.
- [5] P. Brachet and J. M. Baracco, "Dual-polarization slot-coupled printed antennas fed by stripline," in IEEE Transactions on Antennas and Propagation, vol. 43, no. 7, pp. 738-742, July 1995, doi: 10.1109/8.391150.
- [6] Thomas A. Milligan, "Properties of Antennas," in Modern Antenna Design , IEEE, 2005, pp.1-41, doi: 10.1002/0471720615.ch1.
- [7] Thomas A. Milligan, "Dipoles, Slots, and Loops," in *Modern Antenna Design*, IEEE, 2005, pp.217-284, doi: 10.1002/0471720615.ch5.
- [8] David M. Pozar, "Transmission line theory" in *Microwave engineering*, 4th. ed. Hoboken, New Jersey, USA: John Wiley & Sons, Inc., 2011. ch. 2, pp.48-94, ISBN 978-0-470-63155-3
- [9] S. D. Targonski, R. B. Waterhouse and D. M. Pozar, "Design of wide-band aperture-stacked patch microstrip antennas," in IEEE Transactions on Antennas and Propagation, vol. 46, no. 9, pp. 1245-1251, Sept. 1998, doi: 10.1109/8.719966.
- [10] F. Croq and D. M. Pozar, "Millimeter-wave design of wide-band aperture-coupled stacked microstrip antennas," in IEEE Transactions on Antennas and Propagation, vol. 39, no. 12, pp. 1770-1776, Dec. 1991, doi: 10.1109/8.121599.
- [11] R. B. Waterhouse, "Design of probe-fed stacked patches," in IEEE Transactions on Antennas and Propagation, vol. 47, no. 12, pp. 1780-1784, Dec. 1999, doi: 10.1109/8.817653.



**SAAB AB** serves the global market with world-leading products, services and solutions from military defence to civil security. With operations in over 100 countries, Saab continuously develops, adapts and improves new technology to meet customers' changing needs.



**SAAB**

## OUR VISION

It is a human right to feel safe. Since Saab was started, we have strived to keep society and people safe. Through systems and solutions that increase security, we can make this possible.

## OUR MISSION

By pushing mental and technological boundaries we make people safe. Our mission describes how we work to achieve our vision. By improving and updating technological systems and solutions, we increase security in society, for its citizens and for those whose job it is to maintain security. We also contribute insight into how threats to our security change and develop innovative new solutions that make society secure.

## OUR VALUES

### Trust

We are honest and reliable global citizens, and we keep our promises.

### Expertise

We combine a strong history of knowledge with constant learning.

### Drive

We have a passion for innovation, are open to change and committed to being fast and flexible.

[www.saab.com](http://www.saab.com)







## **DYNAMICS - WHO WE ARE**

- A leader in precision engagement systems
- Meeting the global market's requirements
- Worldwide customer base in over 60 countries



# **SAAB**

Business unit Dynamics is a workplace for more than 2200 employees all around the world. We offer a market-leading product portfolio comprising of ground combat weapons, missile systems, torpedoes, unmanned underwater vehicles, training systems and signature management systems for armed forces around the world as well as unmanned underwater vehicles for the offshore industry.

## **TRAINING AND SIMULATION**

Product portfolio covers the full range of training solutions including capabilities and products for live and virtual training.

We understand training and provide scalable state-of-the-art training solutions for military, security and emergency services all over the world.

### **Live training**

- Instrumented training systems

### **Live fire training**

- Range equipment

### **Virtual training and integration**

- Ground combat simulators

### **Training services**

- Logistical and operational support

[www.saab.com](http://www.saab.com)



# Detection of optic disc in retinal fundus image

M. Nohel<sup>1</sup>, R. Kolář<sup>1</sup>

<sup>1</sup>Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology, Brno, Czech Republic

E-mail: [xnohel04@vutbr.cz](mailto:xnohel04@vutbr.cz), [kolarr@vut.cz](mailto:kolarr@vut.cz)

**Abstract**— This paper presents a comparison of several image processing methods for optic disc detection in retinal images. The detection algorithms are tested on five publicly available retinal databases (total 2140 images). Thresholding, Hough transform, and matched filtering followed by postprocessing were tested for optic disc detection. It has been shown that matched filtering followed by peak detection achieves the best success rate (98.8 %).

**Keywords**—optic disc, detection, thresholding, Hough transform, matched filter

## 1. INTRODUCTION

Images of the retina (fundus images) are digital images of the eye taken by (retinal) fundus camera and are used to diagnose various eye diseases. These images are taken to check possible abnormalities occurring on the retina, which allows us to diagnose various diseases such as diabetic retinopathy, glaucoma, etc. In glaucoma progression, the segmentation of the optic disc (OD, outer border) and optic cup (OC, inner border) is an important part of diagnosis (see Figure 1). The cup-to-disc ratio is very often measured. If the ratio is 0.3 or less, the eye is considered to be healthy, but if the ratio is greater, it is an indicator of possible abnormality. Manual segmentation of the optic disc and cup is very time-consuming for the ophthalmologist. Therefore automatic image processing methods to segment these formations are sought. [1]

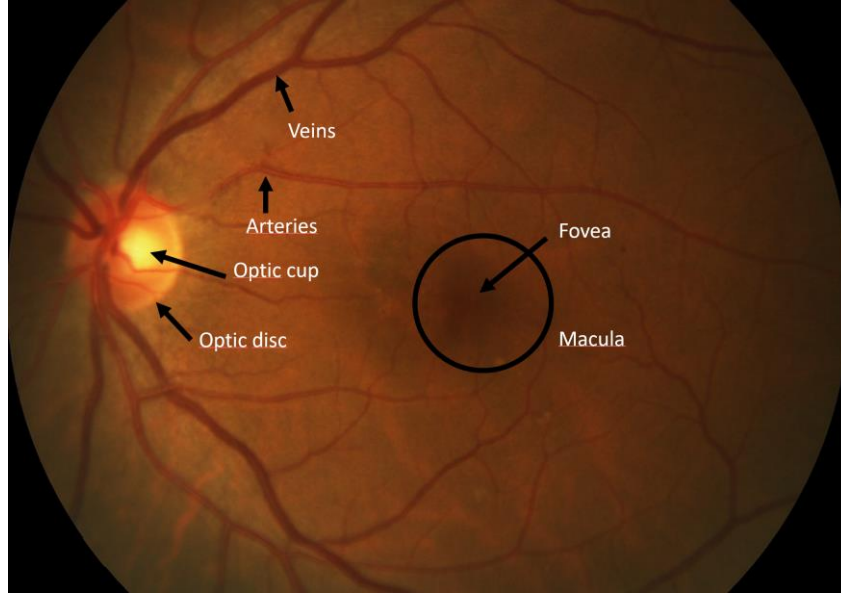
In recent years, methods based on deep learning have become very popular. These methods require a large number of learning data. There are many databases of retinal images with optical disc and cup annotations [2]-[6]. However, these individual databases have different parameters such as image size, FOV, illumination, and ophthalmoscope type. Therefore, the preprocessing step is needed. One of the preprocessing steps is the detection of the optical disc for its localization and subsequent segmentation using a specific segmentation network. This paper presents a comparison of several traditional optical disc detection methods on different databases.

## 2. MATERIALS AND DATA

This paper deals with the detection of the optic disc in retinal images (see Fig. 1). The publicly available databases Dristhi-GS [2], REFUGE [3], RIM-ONE [4], UoA\_DR [5], and RIGA [6] were used. Parameters of databases are shown in Table I. Due to different database parameters, these databases were preprocessed. All databases were resampled to unify spatial resolution – 55 pixels per one field-of-view (FOV). Thus, a total of 2140 retinal images with the same parameters are available, most of them are from healthy subjects, but a minor part of patients suffer from glaucoma or diabetic retinopathy.

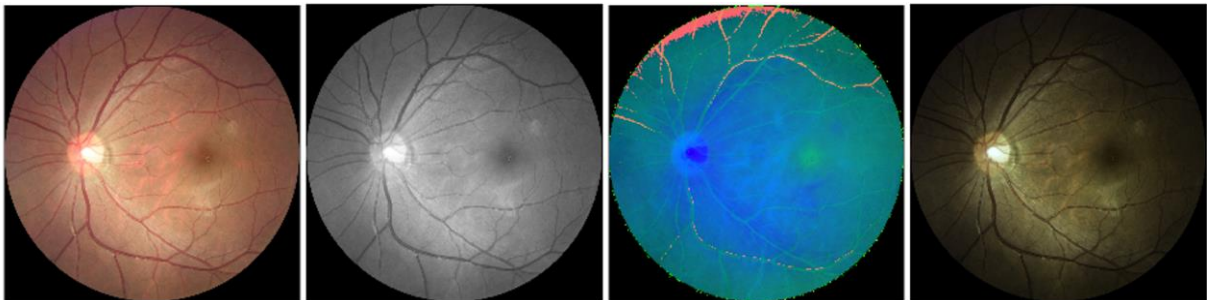
**Table I:** Parameters of individual databases

Databases	Num. of image	FOV	Size of FOV[px]	number of px per 1°
Dristi-GS	101	30°	1750×2048	68
REFUGE-train	400	45°	2056×2124	47
REFUGE-test, valid	400,400	45°	1565×1565	35
RIM-ONE r3	159	20°	1425×1055	53
UoA_DR	200	35°	2056×2124	61
RIGA	480	45°	1435×1435	32

**Figure 1:** An example of a retinal image from a fundus camera with basic retinal structures.

### 3. OPTIC DISC DETECTION METHODS

The above-mentioned databases also contain OD annotations, which enables finding OD centers as a centroid of the segmentation masks. The evaluation of the OD detection success rate consists in calculating the Euclidean distance of the detected disc positions from the centroid positions of the corresponding mask for a given image. The second evaluated criterion was detection accuracy. Detection was considered successful if the found position of the optical disc matched the annotation. Classical image processing methods such as thresholding, Hough transform, matched filtering and their combination were tested for detection. The images from each database were taken with different devices with different settings. For this reason, different data spaces were tested during data preprocessing to test if they could suppress the different luminance of the databases and increase the optic disc detection accuracy. Figure 2 shows an example of RGB, grayscale, HSV, and CIE XYZ color spaces for a selected image from the REFUGE test database.

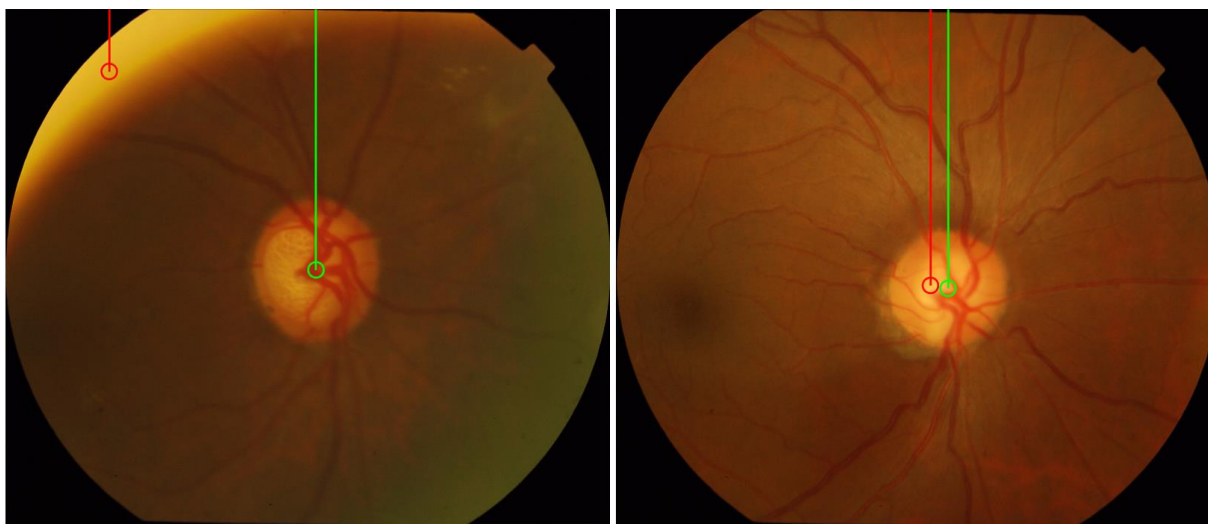
**Figure 2:** Example of different color spaces for a picture from the REFUGE test database. From left RGB, greyscale, HSV, CIE XYZ color space

As can be seen in Figure 2, the optical disc and especially the optical cup are relatively clearly visible in the individual color spaces. Therefore, in the first step, simple thresholding was tested in the indicated color spaces. The optical disc is the brightest part of the image and therefore the threshold was chosen as the 98% percentile of the histogram. The image was then thresholded and subsequently modified using a morphological closure operation. Next, the largest object in the resulting image was found and its centroid was calculated, which indicates the position of the detected optical disc. Another tested approach was based on the Hough transform using the OD circular model. The Hough transform was applied to the thresholded images with a lower threshold and the center of the circle that best approximates the optical disc was used as the optical disc coordinate. The last approach was based on matched filtering. Since it is possible to approximate the OD by 2D Gaussian function, the mask for the matched filtering was chosen as a 2D Gaussian function (with sigma equals to 40 estimated from the average size of the OD whose average radius is 150 px) that blurred the image and then a maximum in the image was detected that represented the OD position. During preprocessing, binary FOV masks were created and used for OD detection, since the OD center should not be at the edge of the FOV, the maxima in the image were searched for after applying matched filtering only on the reduced FOV (eroded by disk structuring element of the size of neighborhood of 40 px).

#### 4. IMPLEMENTATION AND DISCUSSION OF OD DETECTION

The implementation was performed in the MathWorks MATLAB 2021b programming environment. The Image processing toolbox for visualization of data was used and the Statistics toolbox was used for statistical analysis.

The best results are summarised in Table 2. Using thresholding, the best results were obtained over all databases for CIE XYZ color space with subsequent transformation of this space into a greyscale. Detection in the grayscale color space failed especially when an overexposure artifact was captured at the edge of the FOV or there were pathologies in the image that were very bright. The HSV color space proved to be unhelpful as the image in this space was different for each database and no suitable thresholds could be found in this space. Attempting to improve the results using the Hough transform did not lead to an improvement. This was mainly because false detections often occurred when a blood vessel was detected. The best detection results were achieved using matched filtering, where the pattern was a 2D Gaussian function that blurred the image and then detected the maximum in the blurred image, using the information that the FOV mask is known and the disc cannot be at the edge of the FOV. The detection success rate, in this case, was almost 99 %, with false detections occurring mainly in the image with large artifacts as shown in Figure 3.



**Figure 3:** Demonstration of false OD detection and correct detection. Estimated detection is depicted by the red color circle and correct OD position by green color

**Table II:** Summary of OD detection results

Method	Color space	Detection accuracy [%]	MSE [px]
Thresholding	gray	91.9	112.5
Thresholding	CIE XYZ	93.0	105.2
Thresholding+HT	CIE XYZ	87.0	85.9
Matched filtering + morfining	CIE XYZ	98.8	65.7

## 5. CONCLUSION

This contribution deals with OD detection on different databases that have different parameters. The first step was to create a modified database composed of publicly available databases with the same number of pixels per FOV degree. Various methods for image processing such as thresholding, Hough transform, and matched filtering were tested. The average OD radius size across all databases is approximately 150 px. The best OD detection results were obtained using matched filtering with postprocessing, where the detection success rate was nearly 99% and MSE approximately 66 px. To create a segmentation algorithm based on the use of deep learning that will work on different instruments with different image resolutions, proper detection of the optical disc region is essential. The above described and tested detection algorithm based on customized filtering provides a good basis for the subsequent region of interest selection and segmentation using deep learning.

## ACKNOWLEDGMENT

This work was supported by project no. 21-18578S from Czech Science Foundation.

## REFERENCES

- [1] N. Thakur and M. Juneja, "Survey on segmentation and classification approaches of optic cup and optic disc for diagnosis of glaucoma", *Biomedical Signal Processing and Control*, vol. 42, no. 1, pp. 162-189, 2018. doi 10.1016/j.bspc.2018.01.014
- [2] J. Sivaswamy, S. R. Krishnadas, A. Chakravarty, G. D. Joshi, Ujjwal, and T. A. Szed, "A Comprehensive Retinal Image Dataset for the Assessment of Glaucoma from the Optic Nerve Head Analysis", *JSM Biomed Imaging Data Papers*, vol. 2, no. 1, 2015
- [3] J. I. Orlando, H. Fu, J. Barbosa Breda, K. van Keer, D. R. Bathula, A. Diaz-Pinto, R. Fang, P. -A. Heng, J. Kim, J. H. Lee, J. Lee, X. Li, P. Liu, S. Lu, B. Murugesan, V. Naranjo, S. S. R. Phaye, S. M. Shankaranarayana, A. Sikka, J. Son, A. van den Hengel, S. Wang, J. Wu, Z. Wu, G. Xu, Y. Xu, P. Yin, F. Li, X. Zhang, Y. Xu, and H. Bogunović, "REFUGE Challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs", *Medical Image Analysis*, vol. 59, 2020
- [4] F. Fumero, S. Alayon, J. L. Sanchez, J. Sigut, and M. Gonzalez-Hernandez, "RIM-ONE: An open retinal image database for optic nerve evaluation", in *2011 24th International Symposium on Computer-Based Medical Systems (CBMS)*, 2011, pp. 1-6
- [5] R. J. Chalakkal, W. H. Abdulla, and S. Sinumol, "Comparative Analysis of University of Auckland Diabetic Retinopathy Database", *Proceedings of the 9th International Conference on Signal Processing Systems - ICSPS 2017*, pp. 235-239, 2017
- [6] A. A. Almazroa, S. Alodhayb, E. Osman, E. Ramadan, M. Hummadi, M. Dlaim, M. Alkatee, K. Raahemifar, V. Lakshminarayanan, J. Zhang, and P. -H. Chen, "Retinal fundus images for glaucoma analysis: the RIGA dataset", *Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*, p. 8-, Mar. 2018



# Assessing Movement of Articulatory Organs in Patients with Parkinson's Disease

K. Novotný<sup>1</sup>, J. Mekyska<sup>2</sup>

<sup>1</sup>Master Degree Programme (1), FEEC BUT

<sup>2</sup>Department of Telecommunications, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technická 12, 61600 Brno, Czech Republic

E-mail: [212566@vut.cz](mailto:212566@vut.cz), [mekyska@vut.cz](mailto:mekyska@vut.cz)

**Abstract**—Hypokinetic dysarthria is a motor speech disorder often present during Parkinson's disease. It affects the speech system, including articulatory abilities. There are several speech parameters describing this domain, so it is suggested to deal with their mutual comparison. This work aims to design and describe an algorithm for calculating the parameters of articulation, adapted for the Czech language, and then compare their discriminative power. The acoustic analysis of speech included in it is done via the Praat program and basic machine learning algorithms such as Expectation-Maximization, K-means and linear regression are used for the subsequent data processing. The Mann-Whitney U test, descriptive statistics and Random Forest machine learning model using cross-validation and balanced accuracy is used for evaluation. The results are scripts for automatic assessment of vowel space area, for calculating articulation parameters and for their evaluation. The outputs of the analysis of speech recording database prove that differences in articulation can indeed be observed between normal and dysarthric speech. Based on the mutual comparison of results, it is therefore proposed in the work which parameters are being appropriate for further dealing with this issue.

**Keywords**—Hypokinetic dysarthria, acoustic analysis, speech signal processing, speech parametrization, formant frequencies, articulation, machine learning

## 1. INTRODUCTION

Motor speech disorders (MSD) are manifested in areas such as respiration, phonation, articulation or prosody. According to [1] 9.3 % of all MSD is a disorder called hypokinetic dysarthria (HD), which occurs in up to 90 % of patients with Parkinson's disease (PD).

HD is often associated with an articulation disorder (speech is wiped, faded, etc.), which can be quantified using various methods based on formant frequencies [2, 3, 4, 5, 6, 7]. However, these methods are often language dependent and their generalization has not been tested yet.

The aim of this work is to adapt the existing methods for the Czech language and design new algorithms for quantification of articulation based on formant frequencies, and subsequent testing of the ability of these parameters to differentiate healthy controls (HC) and patients with PD.

## 2. MATERIALS AND METHODS

**Database** – The input data for this work were audio recordings from the PARCZ database, which contains 100 patients with PD (40 women/60 men; mean age  $67.36 \pm 8.35$ ) and 52 HC (26 women/26 men; mean age  $63.69 \pm 9.09$ ). It was recorded at the 1st Department of Neurology, St. Anne's University Hospital Brno within the project of the Ministry of Health of the Czech Republic no. NT13499.

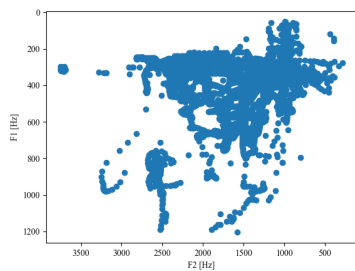
**Formants** – The main regions of energy concentration in the sound spectrum of an acoustic signal are called formants. Thanks to the formants, it is possible to distinguish individual vowels – in the Czech language: [a], [ɛ], [i], [o], [u]. The order of speech formants is determined from those at lower frequencies to higher ones. The most mobile is the second formant – F2. This formant corresponds to the oral cavity and is strongly influenced mainly by the position of the tongue and the size of the jaw



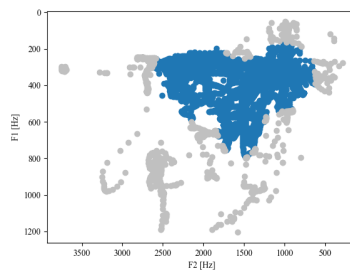
opening. The second most mobile is the formant of the throat cavity – F1. It is also influenced by tongue, although not to such an extent. This time it is the position of the root of the tongue. The movement of the tongue therefore changes the volume ratio of these two cavities. By plotting the pairs F1 – F2 in the graph, where the frequency F2 is on the  $x$ -axis and the frequency F1 on the  $y$ -axis, it is possible to obtain a space with their higher concentration called the Hellwag triangle. An important hypothesis of this work is that when the mobility of articulatory organs deteriorates, there is a centralization in the distribution of formant pairs.

**Proposed algorithm** – To get a better description of articulatory disorders, it is more appropriate to determine the formants from continuous speech. Therefore, in 2013, an algorithm was developed to automatically calculate the Hellwag triangle [8]. Its advantage is automation, as well as the ability to process a speech signal of any length (assuming that it contains all the necessary vowels) and also considering all the vowels contained in the signal (not only those at the vertices of the triangle according to the classical method).

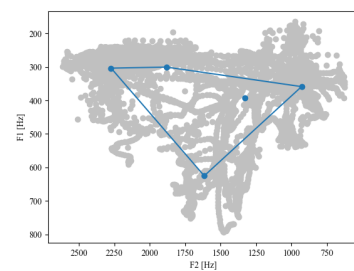
The steps of its equivalent after optimization for Czech speakers proceed as follows. First, F1 and F2 are calculated frame by frame from the speech signal in places with voiced speech (see Figure 1). Subsequently, outliers are removed based on the Gaussian Mixture Model (GMM). Parameters for GMM calculation are obtained using Expectation-Maximization algorithm (EM). Then, those pairs that do not reach the set limit derived from the average likelihood of all observations are discarded (see Figure 2). From the remaining pairs of formants, the cluster centers belonging to the individual vowels are determined. To do this, the k-means algorithm is used, based on a predetermined number of resulting clusters ( $k$ ), which corresponds to the total number of different vowels. After finding the centers, you can draw their convex hull and get its vertices (see Figure 3).



**Figure 1:** Extracted F1 – F2



**Figure 2:** Filtered F1 – F2



**Figure 3:** Convex Hull

Subsequently, from these centers (F1 – F2 pairs), those representing the vowels [i], [a] and [u] are selected. Thanks to these three vertices of the imaginary triangle, it is possible to calculate the basic parameters characterizing the mobility of the articulatory organs – Vowel Space Area (VSA), Logarithmic Vowel Space Area (LnVSA) and Vowel Articulation Index (VAI). The Vowel Space Hull Area (VSHA) parameter is calculated from all vertices of the convex hull. For the Articulatory-Acoustic Vowel Space (AAVS) parameter, a matrix of all values after filtration is used, and for our own Density Percentage (DP) parameter (representing the graphic parameter – Vowel Space Density – VSD) and the relative standard deviations of the first and second formant it is the matrix of all values, including those filtered out. A detailed explanation of the properties and calculation procedure for each parameter is described in work [9].

The calculated parameters for all speakers are processed collectively in the next step. First, the effect of age and gender is removed from all results using linear regression. Subsequently, linear regression is used again, but only for speakers with PD. Thanks to it, the effect of Levodopa Equivalent Dose (LED) medication is removed.

**Statistical analysis** – Descriptive statistics of individual parameters are determined for each group separately. Then all values of one parameter from both groups of speakers undergo a non-parametric Mann-Whitney U test (M-W test). It assesses the statistical significance of the PD effect on a given parameter. The value  $\alpha = 0.05$  was chosen as the level of significance.

**Machine Learning** – To determine the common ability of these parameters to differentiate HC and patients with PD, a Random Forest (RF) machine learning model with fine-tuned (thanks to the grid search method) optimal hyperparameters was used. The comparison of the results was performed after three times repeated 10-fold cross-validation based on the comparison metric – balanced accuracy (BA).

### 3. RESULTS

Eight parameters were calculated for each speaker in the database and these parameters were consequently evaluated. According to the M-W test, three parameters were identified as statistically significant – VSHA ( $p_{\text{VSHA}} = 0.0378$ , HC > PN), AAVS ( $p_{\text{AAVS}} = 0.0038$ , HC > PN) a DP ( $p_{\text{DP}} = 0.0138$ , HC > PN).

In the best trained RF model, an average BA of 70.2 % was achieved with a sensitivity (SEN) of 89.0 % and a specificity (SPE) of 51.3 %. Complete results for both the descriptive statistics and machine learning with tables and graphs can be found in work [9].

### 4. DISCUSSION

According to the results of statistical analysis, VSHA and AAVS appear to be the parameters with the highest discrimination power. The fact that the VSA and LnVSA parameters did not appear as good as the VSHA may be evidence of the significance of the size of the entire vowel space area, including the “cropped” areas when plotting the Hellwag vocal triangle. Thus, it is not only the position of the formants in its vertices that is decisive, but the overall distribution of all F1 – F2 pairs. Similar results are reported by the study from 2019 [5]. It achieves the highest resolution for the second mentioned AAVS parameter too. This study also evaluates positively the distinction between PD and HC by observing differences in the size of areas in VSD. It can be said that in this work it represents our DP parameter whose discriminative power was also strong enough. As in studies [6, 7], VSA and LnVSA do not show promising results. The VAI parameter is on the bottom limit of statistical significance. The results confirm the theory that its values are indeed close to one and lower for speakers with PD than for HC. However, this difference is not very significant. A lower VAI in the vowel space area indicates a higher centralization of the F1 – F2 formant pairs. In speech it can be caused by the stiffness of the tongue or the articulatory organs in general. The tested relative deviations of the formants result in a better resolution for the relative standard deviation of the second formant. This one is related to the front-backness of the vowel and is therefore mainly influenced by the mobility of the tongue. Thus, this outcome could support the hypothesis of its stiffness in patients with PD.

The large range between the minimum and maximum of the VSA, LnVSA and VSHA parameters is most likely due to inappropriate filtering of outliers. In general, this block appears as one of the biggest weaknesses in the data processing algorithm. The filtration parameters strongly influence the calculated values of the parameters for assessing the movement of articulatory organs, which consequently turn out to be very sensitive. We can propose a new way of filtering, based on VSD parameter, i.e. DP, which would allow only pairs of F1 – F2 formants from areas with a normalized density of internal distribution higher than a set limit. Imaginary smooth borders could provide a fairer sieve than GMM.

For trained machine learning models, the SEN was in the vast majority of cases higher than the SPE. This may be due to the fact that the database contains more speakers with PD than HC and thus offers better conditions for training the model to recognize them better. The BA result itself may not seem very delightful. However, according to [10] from 2021, only a fraction of people with PD suffer from articulation problems, so in light of this information, the results are not so bad after all.

Apart from the way of filtering the outliers, a small sample size of our dataset is also a limitation of this work. A higher number of speakers would provide both a more objective evaluation of individual parameters and better generalisation of the machine learning models.

## 5. CONCLUSION

This work had four main goals. The first goal was to design and implement new algorithm for determination of Hellwag vocal triangle from a speech recording. Despite its successful completion, a future work is planned – the limitations of the algorithm and the design of the solution were described above.

The second goal was to use the proposed algorithm to calculate the parameters assessing the movement of articulatory organs. Eight parameters were obtained here. A possible limitation of the simpler ones is the incorrect choice of formant pairs for the vertices of the Hellwag triangle.

The third goal, i.e. the mutual comparison and evaluation of the discrimination power of each individual parameter, was also achieved. Based on descriptive statistics and the M-W test, we can report that VSHA and AAVS have the highest discrimination power.

The last goal was to test the discrimination power of the combination of these parameters. Employing the RF algorithm, we were able to classify PD with BA = 70.2 %.

## ACKNOWLEDGMENT

This study was supported by a grant from the Czech Ministry of Health no. NU20-04-00294.

## REFERENCES

- [1] J. R. Duffy, *Motor Speech Disorders E-Book: Substrates, Differential Diagnosis, and Management*. Mayo Clinic College of Medicine, Rochester, Minnesota: Elsevier Health Sciences, 3 ed., 2019.
- [2] S. Sapir, L. O. Ramig, J. L. Spielman, and C. Fox, “Formant centralization ratio: A proposal for a new acoustic measure of dysarthric speech,” *Journal of Speech, Language, and Hearing Research*, vol. 53, no. 01, pp. 114–125, 2010.
- [3] S. Sapir, L. O. Ramig, J. L. Spielman, and C. Fox, “Acoustic metrics of vowel articulation in parkinson’s disease: Vowel space area (vsa) vs. vowel articulation index (vai),” *Models and Analysis of Vocal Emissions for Biomedical Applications - 7th International Workshop, MAVEBA 2011*, vol. 7, no. 01, pp. 173–175, 2011.
- [4] B. Story and K. Bunton, “Vowel space density as an indicator of speech performance,” *The Journal of the Acoustical Society of America*, vol. 141, no. 05, pp. EL458–EL464, 2017.
- [5] J. Whitfield, “Exploration of metrics for quantifying formant space: Implications for clinical assessment of parkinson disease,” *Perspectives of the ASHA Special Interest Groups*, vol. 4, no. 04, pp. 1–9, 2019.
- [6] J. Whitfield and A. Goberman, “Articulatory-acoustic vowel space: Application to clear speech in individuals with parkinson disease,” *Journal of Communication Disorders*, vol. 51, 09 2014.
- [7] J. Whitfield and A. Goberman, “Articulatory-acoustic vowel space: Associations between acoustic and perceptual measures of clear speech,” *International journal of speech-language pathology*, vol. 19, no. 04, pp. 184–194, 2017.
- [8] S. Sandoval, V. Berisha, R. Utianski, J. Liss, and A. Spanias, “Automatic assessment of vowel space area,” *The Journal of the Acoustical Society of America*, vol. 134, no. 11, pp. EL477–83, 2013.
- [9] K. Novotný, “Hodnocení hybnosti mluvidel na základě akustické analýzy řeči,” Master’s thesis, Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, Brno, 2021.
- [10] J. Rusz, T. Tykalova, M. Novotny, D. Zogala, K. Sonka, E. Ruzicka, and P. Dusek, “Defining speech subtypes in de novo parkinson disease,” *Neurology*, vol. 97, no. 21, pp. e2124–e2135, 2021.

# Comparison of full-size and patches-based learning approaches for aneurysm segmentation in TOF-MRI data

J. Vývoda<sup>1</sup>, R. Jakubíček<sup>1</sup>

<sup>1</sup>Department of Biomedical Engineering, FEEC, Brno University of Technology, Brno, Czechia

E-mail: [xvyvod01@vut.cz](mailto:xvyvod01@vut.cz), [jakubicek@vut.cz](mailto:jakubicek@vut.cz)

**Abstract**—The paper is interested in segmentation of intracranial aneurysms. Intracranial aneurysms are life-threatening issue. In this paper there are proposed two methods for this segmentation problem. First one is segmentation with use of full size images, the other one uses patches of the image, which could help decrease the ration between pixels representing background and pixels representing aneurysms. Data from ADAM challenge 2020 are used to train and evaluate these approaches. Using full images show better results in dice coefficient, which is 0.16 greater, then patched image approach.

**Keywords**—Intracranial aneurysm, aneurysm, machine learning, detection, magnetic resonance, U-net, segmentation

## 1. INTRODUCTION

Detection of intracranial aneurysms (IA) is a very difficult, but very important task, since it affects 5-10 % of the whole population [1]. In a lot of cases, the presence of aneurysm is only discovered during subarachnoid haemorrhage (SAH) as a consequence of a rupture of this IA. The mortality of SAH caused by ruptured aneurysm is between 23 % and 51 % and thereafter, fairly large percentage of people will remain disabled [2].

There are many methods for a diagnosis of IAs such as digital subtraction angiography (DSA), computed tomography angiography (CTA) or magnetic resonance angiography (MRA). But even so, locating an aneurysm in these images is difficult, time consuming task and the doctors must be very experienced in this field. Therefore, it is important to develop an aneurysm detection algorithm, that would support and facilitate the work of doctors.

These days, the deep learning methods are the state-of-the-art approaches. In [3] a DeepMedic CNN architecture was used for segmentation of TOF-MRI data. As a result, authors achieved an average sensitivity of 0.89 by this convolutional network, and when combined with human readers, they achieved on average sensitivity of 0.98 [3]. Another approach of detecting aneurysms is described in [5]. Here they used a surface approach instead of a voxel approach. First they extracted the surface of the brain arteries and used PointNet++ for their classification. After that, they did a segmentation using SO-Net. As a result, authors have achieved dice coefficient = 0.72 % which is much more, then they have achieved using voxel-based methods [5]. In article [4] authors use patches of images to train CNN for segmentation of brain tumors. Their complex segmentation algorithm achieved dice coefficient 0.90 [4].

As can be seen from the articles above, segmentation usually might be part of an IA or tumor detection. This paper compares two approaches of segmentation using U-net. First one is a normal approach, where a whole image is fed to the network. The other one uses patches of the images, as were used in [4]. Main idea is to divide the image into number of patches and after segmentation, reassemble the image back to it's original size. In this paper only random cutouts are used instead of using image divided into patches, as patches might be worse option for training the network. The main advantage of this approach is decreasing the ration between number of background and aneurysm pixels. Both techniques are compared on TOF-MRI data from ADAM challenge from 2020.

## 2. DATA PREPROCESSING

The data used in this paper were provided by the Aneurysm Detection and Segmentation (ADAM) challenge. The dataset contains data of 113 cases. There are 23 unique patients with aneurysms and SAH, 35 baseline and 35 follow-up cases with aneurysms and SAH. Then there are also 20 patients who do not have any aneurysms or SAH. The data contain TOF-MRA image and manually created references

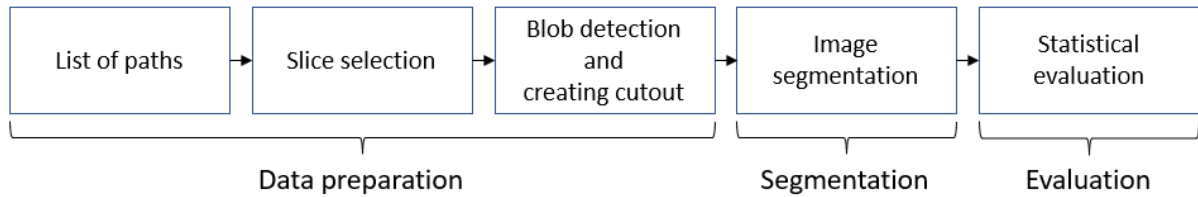
(masks) of the images and locations of aneurysms in text files. All images were already preprocessed using correction to adjust the bias field. For further processing and to reduce computational complexity, same voxel size for all images with a 0.5 mm edge length, was used. Data were also normalized to a range from 0 to 1.

### 3. PROPOSED PIPELINES

Before the segmentation, data had to be prepared based on the approach. First of all, a list of paths to each patient was created. Then according to the text file with aneurysm, position slices were selected and were added to the list of paths. Based on this list, slices that are used for training of the U-net, were selected. As mentioned above, some patients did not have aneurysms at all. In this cases, random slices were selected. As these were NIfTI files, the SimpleITK library was used to create a function for browsing through the individual slices in the NIfTI file.



**Figure 1:** Block diagram of proposed workflow of full image segmentation.



**Figure 2:** Block diagram of proposed workflow of segmentation using patches, where blob detection and creating cutout is a necessary step compared to full image segmentation.

#### 3.1. Data preparation

For segmentation of the full images, data had to be prepared as a list of paths, as mentioned above and then resized to size  $200 \times 200$  pixels, since they all had to be the same size to enter the network. Proposed workflow is shown in Figure 1.

Due to the big difference between number of positive and negative pixels, the cutout segmentation could be beneficial. Therefore a function based on a blob detector (based on difference of gaussian) was used to create cutouts of size  $50 \times 50$  from groundtruth masks. These cutouts are created randomly around the detected blobs. Corresponding cutouts were also extracted from the original TOF-MRI image and fed to the network. Proposed workflow is shown in Figure 2. Dividing image into patches is not used for training the network, because there would be a lot of patches that only contain background. Same slices were selected for training both models, as well as for testing, so that the results are comparable.

The network architecture is fairly basic and taken from [6]. The key part of the network are two consecutive convolutions. Kernel size  $3 \times 3$  and stride equal to 1 are used for the convolutions. Batch normalization is used to improve training of the network and ReLU activation function is used here. Dimensional reduction is achieved by maxpooling layer with kernel size  $2 \times 2$  and stride equal to 2. At the end of this U-net, there is final  $1 \times 1$  convolution, that creates a feature map. To visualize the segmented images and evaluate the network, this feature map is converted to a range from 0 to 1 using a sigmoidal function followed by thresholding.

#### 3.2. Image segmentation

U-net used in this paper is trained from scratch. Binary cross entropy with sigmoidal layer was used as a loss function. Due to the large disproportion between the number of positive and zero pixels, positive pixels were weighted, for full images by 150 and for cutouts by 40. Batch size was equal to 8 images.

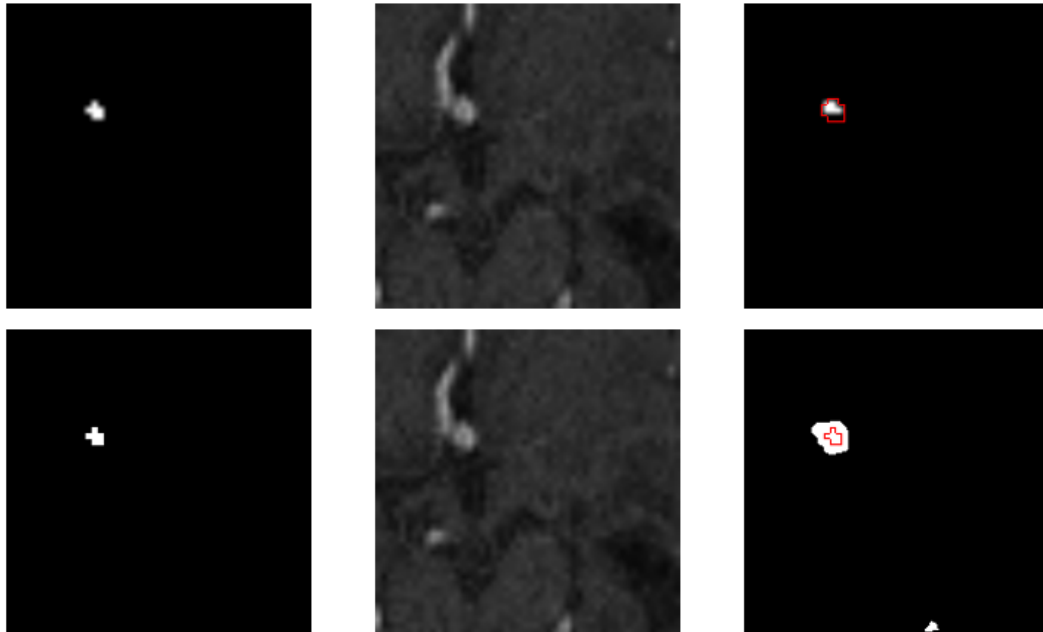
Adam was used as the optimization algorithm for both approaches. The learning rate for full image segmentation was set to  $lr = 1 \times 10^{-4}$  for first forty epochs and then reduces to value  $lr = 1 \times 10^{-5}$  until the end of 135th epoch. Learning rate for cutout image segmentation was set to  $lr = 1 \times 10^{-4}$  for first sixty epochs and then reduces to value  $lr = 1 \times 10^{-5}$  until the end of 135th epoch.

#### 4. EVALUATION OF BOTH APPROACHES

Data of 25 % of cases were selected for evaluation of both approaches. These patients were not used for training the model. Dice coefficient was chosen as main statistical metric for evaluation of the segmentation, which evaluates the similarity of segmented image and groundtruth. Also precision and recall were used as statistical metrics for segmentation. For evaluation of the results, positive predictive value (PPV) and Manhattan distance was computed, as well. Positive predictive value was computed as a ration between number of true positive segments (aneurysms in groundtruth masks) and number of all positive segments (from segmented images). Manhattan distance was computed between centroids of segments aquired by segmentation and centroids of ground truth mask segments (real aneurysms). The lowest distance values were always selected for the images. It was assumed that the segment with the lowest value was the most accurate estimate. The results are summarized in Table I.

**Table I:** Comparison of the segmentation results of both approaches. The results are calculated as mean values over all slices and there are also standard deviations (Std).

Used approach	Dice/Std	Precision/Std	Recall/Std	PPV/Std	Man. distance/Std
Full image segmentation	0.78/0.12	0.75/0.10	0.83/0.12	0.21/0.23	36.71/43.86
Cutout image segmentation	0.62/0.08	0.50/0.07	0.82/0.04	0.33/0.29	26.33/41.97



**Figure 3:** The first row represents zoomed-in result of full image based network. The second row shows result of patch based approach. There is the groundtruth mask on the left, original TOF-MRI image in the middle and segmented image on the right. In the segmented images, there are shown the contours (red) of the true aneurysms.

According to the achieved results, even though dividing image into patches reduce the ratio between negative and positive pixels, full image segmentation reaches better results of dice coefficient and also precision metric. The recall values are almost equal. As can be seen from the Table I, positive predictive value is better with use of cutouts, which means, that there are less false positive segments, then in the



other approach. Manhattan distance is a bit higher in full image approach, so the segments are not as precisely segmented as in the cutout approach. In Figure 3, there is an example of image segmentation by both approaches. Images in the first row represent results of full image approach and as can be seen, only part of the aneurysm is segmented, whereas in the second row (cutout approach), the segmentation of the aneurysm is not accurate and often the surrounding structures are segmented, as well.

## 5. CONCLUSION

Both of the approaches are able to learn to segment aneurysms or subarachnoidal haemorrhages with large area. According to the results shown in Table I, full image based network achieves better results, but predicts more false positive segments. This issue can be solved by further processing, for example by another convolutional network, that would decide, which segments are real aneurysms and which are not.

## ACKNOWLEDGMENT

I would like to thank the authors of [7] for providing the ADAM challenge data and allowing its use for this paper.

## REFERENCES

- [1] CARANCI, F., F. BRIGANTI, L. CIRILLO, M. LEONARDI a M. MUTO. Epidemiology and genetics of intracranial aneurysms. *European Journal of Radiology* [online]. 2013, 82(10), 1598-1605 [cit. 2022-03-12]. ISSN 0720048X. Available from: [doi:10.1016/j.ejrad.2012.12.026](https://doi.org/10.1016/j.ejrad.2012.12.026).
- [2] UEDA, Daiju, Akira YAMAMOTO, Masataka NISHIMORI, et al. Deep Learning for MR Angiography: Automated Detection of Cerebral Aneurysms. *Radiology* [online]. 2019, 290(1), 187-194 [cit. 2022-03-12]. ISSN 0033-8419. Available from: [doi:10.1148/radiol.2018180901](https://doi.org/10.1148/radiol.2018180901).
- [3] FARON, Anton, Thorsten SICHTERMANN, Nikolas TEICHERT, et al. Performance of a Deep-Learning Neural Network to Detect Intracranial Aneurysms from 3D TOF-MRA Compared to Human Readers. *Clinical Neuroradiology* [online]. 2020, 30(3), 591-598 [cit. 2022-03-12]. ISSN 1869-1439. Available from: [doi:10.1007/s00062-019-00809-w](https://doi.org/10.1007/s00062-019-00809-w).
- [4] BEN NACEUR, Mostefa, Mohamed AKIL, Rachida SAOULI a Rostom KACHOURI. Fully automatic brain tumor segmentation with deep learning-based selective attention using overlapping patches and multi-class weighted cross-entropy. *Medical Image Analysis* [online]. 2020, 63 [cit. 2022-03-13]. ISSN 13618415. Available from: [doi:10.1016/j.media.2020.101692](https://doi.org/10.1016/j.media.2020.101692).
- [5] YANG, Xi, Ding XIA, Taichi KIN a Takeo IGARASHI. A Two-step Surface-based 3D Deep Learning Pipeline for Segmentation of Intracranial Aneurysms [online]. 2020, , 1-9 [cit. 2022-03-12]. Available from: <https://arxiv.org/abs/2006.16161>.
- [6] DU, Getao, Xu CAO, Jimin LIANG, Xueli CHEN a Yonghua ZHAN. Medical Image Segmentation based on U-Net: A Review. *Journal of Imaging Science and Technology* [online]. 2020, 64(2), 20508-1-20508-12 [cit. 2022-03-12]. ISSN 1062-3701. Available from: [doi:10.2352/J.ImagingSci.Technol.2020.64.2.020508](https://doi.org/10.2352/J.ImagingSci.Technol.2020.64.2.020508).
- [7] TIMMINS, Kimberley M., Irene C. VAN DER SCHAAF, Edwin BENNINK, et al. Comparing methods of detecting and segmenting unruptured intracranial aneurysms on TOF-MRAS: The ADAM challenge. *NeuroImage* [online]. 2021, 238 [cit. 2022-03-25]. ISSN 10538119. Available from: [doi:10.1016/j.neuroimage.2021.118216](https://doi.org/10.1016/j.neuroimage.2021.118216)

# Dynamic metabolomic prediction based on genetic variation for *Hordeum vulgare*

P. Nemčková<sup>1</sup> and J. Schwarzerová<sup>1,2</sup>

<sup>1</sup>Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology, Brno, Czech republic

<sup>2</sup>Molecular Systems Biology (MOSYS), University of Vienna, Vienna, Austria

E-mail: [xnemce05@vutbr.cz](mailto:xnemce05@vutbr.cz), [Jana.Schwarzerova@vut.cz](mailto:Jana.Schwarzerova@vut.cz)

**Abstract**—*Hordeum vulgare*, like many other crops, suffers from the reduction of genetic diversity caused by climate changes. Therefore, it is necessary to improve the performance of its breeding. Nowadays, the area of interest in current research focuses on indirect selection methods based on computational prediction modeling. This study deals with dynamic metabolomic prediction based on genomic data consisting of 33,005 single nucleotide polymorphisms. Metabolomic data include 128 metabolites belonging to 25 Halle exotic barley families. The main goal of this study is creating dynamic metabolomic predictions using different approaches chosen upon various publications. Our created models will be helpful for the prediction of phenotype or for revealing important traits of *Hordeum vulgare*.

**Keywords**—Machine learning, Single nucleotide polymorphism, genomic prediction, *Hordeum vulgare*

## 1. INTRODUCTION

Genomic prediction is one of the best computational approaches for acquisition of information that is important to breed crops [1]. Genomic predictions are also one of the best methods for selection of the most perspective plants for breeding or for optimization of growing healthy plants [2]. Thanks to that, plants can grow faster with reduced usage of water. Thus, this methodology has a huge potential in ecology and biotechnology. In addition, the same methodology can be useful for early disease detection in biomedicine.

Currently, research focuses on corss-linking genotypes and phenotypes. The main direction of interest focuses on metabolomics with metabolite concentration in the spotlight. This is due to the fact, that metabolites represent the response of biological processes in organisms. The genotype-metabolotype association is sensitive to environmental changes [3]. The first step in current research leads to creating and improving techniques that will have a potential to describe the relationship between genome and metabolome.

Halle Exotic Barley (HEB) population [4] was developed using nested association mapping (NAM), where exotic alleles were placed into a genetic background of German spring barley. Genotyping with the Infinium iSelect 9k SNP chip was used for obtaining the genomic information [5], that was used in this study. The metabolomic data taken from study by Gemmer et al. [6], were processed using MassHunter Qualitative Analysis software. This resulted in data for 1,307 lines where 158 metabolites (METs) were defined. In this study, we focused on the prediction of metabolite concentration related to the HEB family based on genomic variance represented as single nucleotide polymorphism (SNP). This study uses genomic and metabolomic dataset of HEB population mentioned above and brings three models for metabolomic prediction based on genetic variation.

## 2. DATASET

*Hordeum vulgare* [7] has wide usage in the food industry as it is the 4th most important crop in the world. According to climate changes, the genetic diversity of barley has been reduced. The genomic dataset was taken from the study by Maurer et al. [8]. The study by Gemmer et al. [6] presented metabolomic information which is related to the above mentioned genomic dataset. Thus, these metabolite data were

used in our study, too.

Genotyping was performed by 50k Illumina Infinium iSelect SNP Array [8]. Genomic data include 1,363 measures of *Hordeum vulgare* nested association mapping population HEB-25. In metabolomic data, 1,361 concentrations of 1,419 were connected to one of the HEB families, which were analyzed in our study.

### 3. METHODS

#### 3.1. Pre-processing

Firstly, data were reduced by a number of HEB lines and united by accession number for genomic and metabolomic values. The final SNP table includes 3,877 SNP markers, 13 metabolites and 1,307 HEB lines. The genomic dataset was transformed from scaling  $\langle 0, 2 \rangle$  to  $\langle -1, 1 \rangle$  for optimization of the input for predictions.

#### 3.2. Dynamic metabolomic prediction

In general, the metabolomic prediction based on genetic variation used static genomic information such as SNP table. This study focused on the dynamic metabolomic prediction using three different prediction approaches, namely: support vector regressor (SVR), least absolute shrinkage and selection operator (LASSO) and sparse partial least squares regression (sPLS), chosen based on publications [9, 10].

Each of the prediction method was implemented for a different setting of parameters. The different rate of training and testing data in ratio from 9:1 to 6:4 was applied. In this study, parameter alpha in LASSO model is set to 0.3.

sPLS is a method based on covariance that combines spectral decomposition technique and multiple regression analysis. The main principle of this method relies on searching for linear combinations of variables. In this study, the number of components of the sPLS model is set to 10. Due to that, the dimensionality of the data is reduced.

SVR uses a non-linear radial basis function and a polynomial kernel function. The principle of SVR consists of transforming data into a hyperplane represented by support vectors. In this study, the regularization parameter C is set to 1 and the  $\epsilon$ -insensitive region is set to 0.1. All created and used scripts are available on Github: 'PetraNemcekova/Genomic\_Prediction'.

### 4. RESULTS AND DISCUSSION

The study brings three prediction models based on SNP using three different approaches. Evaluation parameters represent mean absolute error (MAE) and mean squared error (MSE) [11].

Summary of results of all approaches using SVR, LASSO and sPLS are shown in Table I. The best predicting SVR model is trained on 70% of the training dataset. The MAE of the SVR model is 0.188 and MSE is 0.058.

**Table I:** Summary of results of trained prediction models using SVR, LASSO and sPLS with different percentage of data used for training.

	SVR		LASSO		sPLS	
	MAE	MSE	MAE	MSE	MAE	MSE
10% test	0.198	0.058	0.206	0.058	0.413	5.743
20% test	0.203	0.107	0.202	0.069	0.546	0.851
30% test	0.188	0.058	0.272	1.996	0.511	0.576
40% test	0.198	0.074	0.256	1.513	0.279	0.576

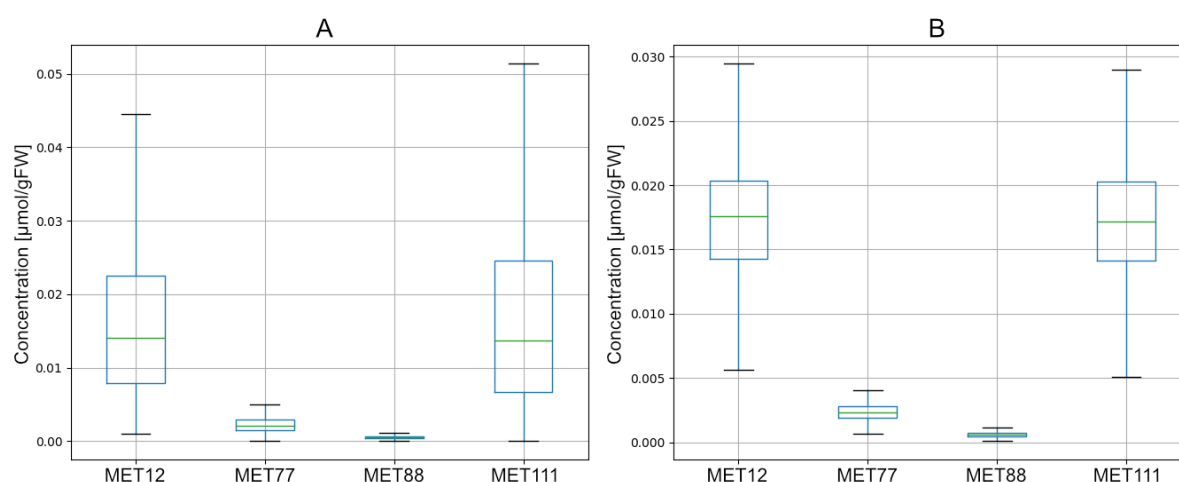
The best evaluation parameters of the SVR model is connected to metabolite MET25, MAE is 0.067 and MSE is 0.005. The worst predicted results, MAE is 0.198 and MSE is 0.088, were found for disaccharide MET122. SVR is an appropriate regression model for metabolomic prediction based on genomic variation.

The best evaluation parameters of LASSO models are 0.206 MAE and 0.058 MSE. The standard deviations

(SD) are 0.089 and 0.025. The minimal MAE is associated with the metabolite MET39, the minimal MSE is connected to metabolite MET88. LASSO was proved to be a suitable regression model for metabolomic prediction based on SNP.

In sPLS models, the best division of training and testing data is ratio 7:3 according to obtained values of MAE and MSE. The best sPLS model has 0.511 MAE with 0.933 SD and MSE is 0.576 with 17.767 SD.

The minimal error in the sPLS model is connected to the metabolite MET88, with the MAE being 0.017. According to MSE, the best-predicted evaluation is associated with MET39 with the value of MSE being 0.0007.



**Figure 1:** Box plots of original and predicted values by sPLS model. A section shows boxplots of original values and B section represents boxplots of predicted values using sPLS. The boxplots show the statistics of ascorbic acid MET12, ribonic acid MET77, hydroxylase inhibitor MET88 and transthyretin MET111.

Figure 1 shows boxplots of original and predicted data using the sPLS method of 4 randomly chosen metabolites. The values in these boxplots are visualized in the same numerical order. Because of that, evaluation parameters such as MAE and MSE are sufficient for the description of models.

The model with the best evaluation parameters was created using SVR. Thus, this model was selected and modeled using cross-validation with 5 repeats and the number of splits 5. Thanks to that, we obtained the final model with better accuracy. The evaluation parameters of 5-cross-validation have the value of MAE 0.150 and the MSE 7.608.

## 5. CONCLUSION

Nowadays, ecology research of plants has huge potential due to the combination of laboratory analysis and new computational algorithms. Thanks to that, we can open new way of possibilities based on prediction modeling. This can help with prediction of disease phenotypes or with selection of the most perspective plants in an early development of the crops. One of these major ways is a dynamic metabolomic prediction based on genetic variation. The reason for this is that especially linking of genotype and metabolite is sensitive to environmental changes such as climate changes. This study focused on dynamic metabolomic prediction based on genetic variation in *Hordeum vulgare*. It is a representative of crops which have a main role in the food industry and it is necessary to improve the performance of its breeding. The main goals of our study are creating and comparing prediction models that can be used for describing the relationship between genome and metabolome. These models can be used for the prediction of phenotype or revelation of the important traits of *Hordeum vulgare* using the genotype-metabolotype association. In total, we modeled three prediction models based on three different modeling approaches such as LASSO, sPLS and SVR. All models were evaluated using MSE and MAE. The model that connected to the best evaluation parameters was created using the SVR method. A cross-validation was used to improve this model. This study brings three models that can be used for prediction of metabolomic information based on genetic variation. Thanks to that, this study partly fills the blank space in solving this problematics and opens new opportunities for predicting the genotype-phenotype association. The

whole methodology was implemented in Python and is available on Github: 'PetraNemcekova/Genomic\_Prediction'.

## ACKNOWLEDGMENT

This work has been supported by grant FEKT-K-21-6878 realised within the project Quality Internal Grants of BUT (KInG BUT), Reg. No. CZ.02.2.69 / 0.0 / 0.0 / 19\_073 / 0016948, which is financed from the OP RDE.

## REFERENCES

- [1] L. M. Zingaretti et al., "Exploring Deep Learning for Complex Trait Genomic Prediction in Polyploid Outcrossing Species", *Front. Plant Sci.*, 2020, doi [10.3389/fpls.2020.00025](https://doi.org/10.3389/fpls.2020.00025).
- [2] M. E. Goddard, B. J. Hayes, "Genomic selection", *J Anim Breed Genet*, 2007, doi [10.1111/j.1439-0388.2007.00702.x](https://doi.org/10.1111/j.1439-0388.2007.00702.x).
- [3] A.-M. Waldvogel et al., "Climate Change Genomics Calls for Standardized Data Reporting", *Front. Ecol. Evol.*, 2020, doi [10.3389/fevo.2020.00242](https://doi.org/10.3389/fevo.2020.00242).
- [4] P. Herzig et al., "Contrasting genetic regulation of plant development in wild barley grown in two European environments revealed by nested association mapping", *Journal of Experimental Botany*, 2018, doi [10.1093/jxb/ery002](https://doi.org/10.1093/jxb/ery002).
- [5] A. Maurer et al., "Modelling the genetic architecture of flowering time control in barley through nested association mapping", *BMC Genomics*, 2015, doi [10.1186/s12864-015-1459-7](https://doi.org/10.1186/s12864-015-1459-7).
- [6] M. R. Gemmer et al., "Can metabolic prediction be an alternative to genomic prediction in barley?", *PLoS One*, 2020, doi [10.1371/journal.pone.0234052](https://doi.org/10.1371/journal.pone.0234052).
- [7] S. Abbo et al., "Agricultural Origins: Centers and Noncenters; A Near Eastern Reappraisal", *Critical Reviews in Plant Sciences*, vol. 29, no. 5, p. 317-328, 2010, doi [10.1080/07352689.2010.502823](https://doi.org/10.1080/07352689.2010.502823).
- [8] A. Maurer et al., "50k Illumina Infinium iSelect SNP Array data for the wild barley NAM population HEB-25", *e!DAL*, 2019, doi [10.5447/ipk/2019/20](https://doi.org/10.5447/ipk/2019/20).
- [9] S. Wang et al., "Identification of optimal prediction models using multi-omic data for selecting hybrid rice", *Heredity*, 2019, doi [10.1038/s41437-019-0210-6](https://doi.org/10.1038/s41437-019-0210-6).
- [10] C. Du et al., "Genomic selection using principal component regression", *Heredity*, 2018, doi [10.1038/s41437-018-0078-x](https://doi.org/10.1038/s41437-018-0078-x).
- [11] R. Sun et al., "Prediction of Liver Weight Recovery by an Integrated Metabolomics and Machine Learning Approach After 2/3 Partial Hepatectomy", *Front Pharmacol*, 2021, doi [10.3389/fphar.2021.760474](https://doi.org/10.3389/fphar.2021.760474).

# Analysis of brain tumors based on line of interest

K. Širůčková<sup>1</sup>, P. Solár<sup>2</sup>, P. Marcoň<sup>1</sup>

<sup>1</sup>Faculty of electrical engineering and communication Brno University of Technology, Czech Republic

<sup>2</sup>Department of Neurosurgery, St. Anne's University Hospital Brno, Czech Republic

E-mail: [xsiruc03@vutbr.cz](mailto:xsiruc03@vutbr.cz), [peter.solar@fnusa.cz](mailto:peter.solar@fnusa.cz), [marcon@vut.cz](mailto:marcon@vut.cz)

**Abstract**—Due to the high resolution of soft tissue, magnetic resonance imaging plays a crucial role in the diagnosis and therapy planning in the neurosurgery field whereas it is necessary to determine which pathology in the brain tissue is involved. Glioblastoma multiforme, metastatic tumors and abscesses are examined in detail from magnetic resonance images. In clinical practice, all mentioned pathologies are diagnosed through invasive methods in the form of biopsy followed by histology of the affected tissue. This work is focused on an alternative non-invasive method of tumor diagnosis. The method is based on the data analysis from defined curves (drawn into the apparent diffusion images) that lead from the tumor area. The analysis of curves could lead to non-invasive diagnostics of the pathological tissue.

**Keywords**—Magnetic resonance imaging, apparent diffusion coefficient, glioblastoma multiforme, metastasis, abscess

## 1. INTRODUCTION

The human brain can be considered one of the most complex objects. The structure of the brain is vertically divided into two hemispheres (right and left) and their surface is gyrified (gyri – folds). In both hemispheres, several lobes can be distinguished, where various sensory and motor information coming from the external environment are processed. The hemisphere filling is made up of gray and white brain matter. The gray matter is the site of the accumulated bodies of neurons and the white matter is rich in nerve fibers, which form the functional bundles of the so-called nerve pathway. All areas of the brain are interconnected by so-called synapses, which ensure that the brain works as a whole and is able to ensure our actions in continuity.

Brain and spinal cord tumors are among the relatively common cancer with high mortality. Therefore, in addition to prevention, early and accurate diagnosis of the pathology is important. Brain tumor surgery is a compromise between minimal damage to healthy brain tissue and effective removal of all tumor tissue. Morphological magnetic resonance imaging of the brain is now a commonly used imaging method in the diagnosis of various pathologies. Magnetic resonance imaging can distinguish anatomical structures in many imaging modalities with great accuracy. This work deals with the analysis of three types of brain pathologies - glioblastoma multiforme (GBM), metastatic tumors and abscesses [1].

Currently, the comparison of ADC values is devoted to e.g. [2], where there is an effort to distinguish between GBM and anaplastic astrocytoma, or [3], which deals with the issue of curve placement and differentiation between GBM and metastatic lesions. No article was found to examine the differences between the brain pathologies we studied.

## 2. METHODOLOGY

The following subchapters describes the patient population, image acquisition and used methods.

### 2.1 POPULATION

Patients were subjected to magnetic resonance imaging. For this experiment there were used two types of magnetic resonance tomograph. Mostly was used tomograph from General Electric with a magnetic

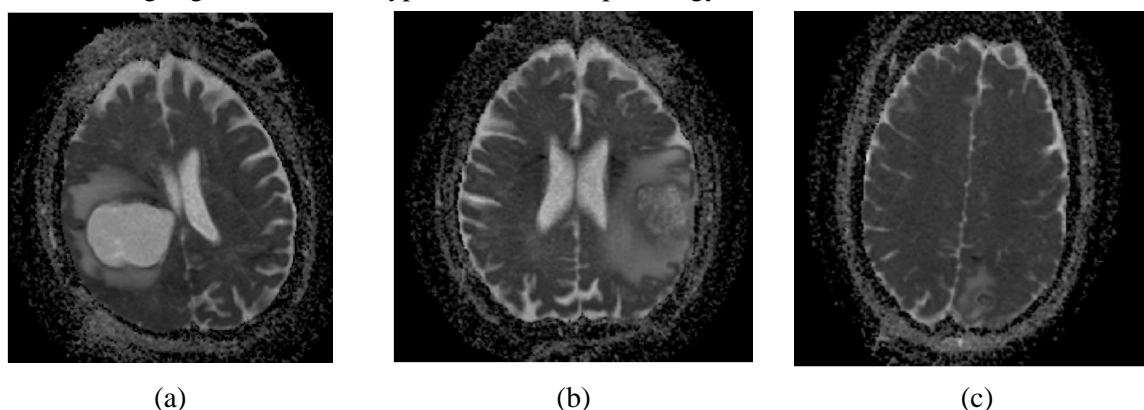


field induction size of 3 T. In few cases was used Philips magnetic resonance with size of 1.5 T. There were 30 patients examined with three types of brain pathology. A total of 10 patients with a diagnosis of metastatic brain disease aged 29 to 72 years (mean age 61.4 years) were included in the statistical testing. Of these, 5 were men (age range 64 to 72 years, mean  $\pm$  standard deviation: 66.8 years  $\pm$  2.7) and 5 women (29 to 72, mean 56  $\pm$  16.2). Patients with diagnosis of brain abscess were aged from 37 to 82 years (mean age 60.7 years). There were 6 men (age range 37 to 82 years, 57.7 years  $\pm$  15) and 4 women (47 to 87, 65.3  $\pm$  15). Patients with diagnosis of GBM were aged from 22 to 80 years (mean age 60.9 years). There were 7 men (age range 22 to 80 years, 59 years  $\pm$  18) and 3 women (61 to 70, 65.3  $\pm$  3.7).

## 2.2 IMAGE ACQUISITION

The images were taken by doctors at the Neurosurgical Clinic of St. Anne's Hospital. The data were provided in DICOM format. Diffusion weighted images were used, also known as ADC maps, described in more detail [4] [5]. From the sequence of images, a section with the best visible pathology was selected, in which the lines of interest were placed. The size of the metastasis and GBM images was 256x256. In the case of the abscess, the image sizes were the same in some patients, but also smaller, e.g., 160x160 pixels.

In the following Fig.1 there are all types of examined pathology.



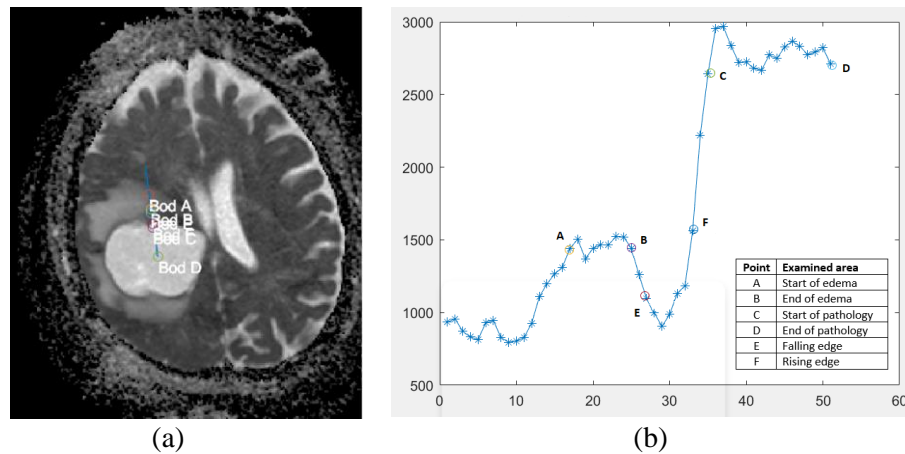
**Figure 1:** Examined pathology – (a) glioblastoma multiforme, (b) metastasis, (c) abscess

GBM is the most aggressive, fastest growing brain tumor. On a neuroradiological examination, it appears as a well-defined tumor, while it is star-shaped, and its protrusions are difficult to observe. Metastases come in many forms, from dotted bearing to large. The typical size of a metastatic tumor is about 2 cm. As metastatic brain tumors grow, they build up pressure and change the function of the surrounding brain tissue. The brain abscess manifests itself as a purulent deposit usually caused by a bacterial infection (anaerobic streptococcus, Escherichia coli, Golden Staphylococcus). During the growth of the abscess, the deposit is inaccurately demarcated, but after fourteen days of growth, the deposit necrotizes, and the boundaries become clear [6].

## 2.3 METHODS

The subject of the research are images weighted by apparent diffusion coefficient, obtained from magnetic resonance. A working program was created in Matlab R2020b, which allows to draw a curve (line of interest) in the tumour area and then statistically evaluate the descriptive quantitative quantities as standard deviation, median and mean. First, the program loads one pre-selected (e.g., using Weasis, Fiji, etc.) slice from the given patient in DICOM. After loading, the user places a line leading from the edematous to the pathological area. The curves were placed under the supervision of doctors. Subsequently, a 2D graph of values is drawn, in which the user is allowed to select the examined area of the tumour and edema thanks to the possibility of marking the examined points.

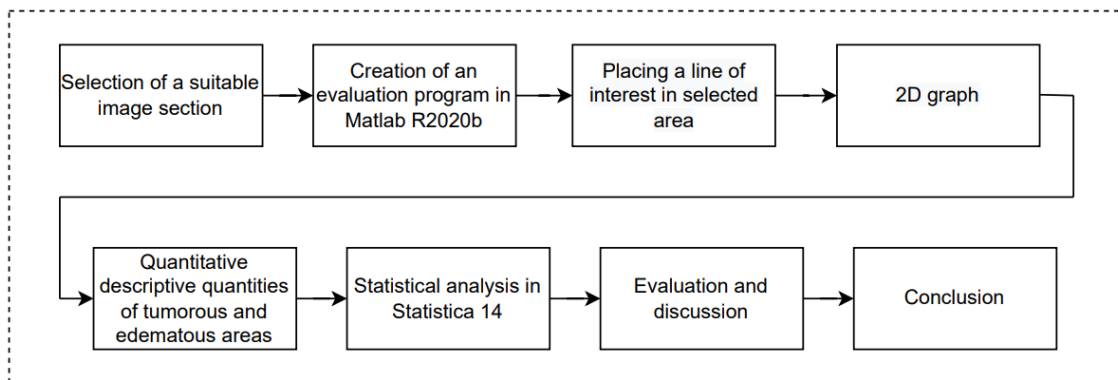
It is shown on Fig. 2.



**Figure 2:** Placing line of interest and plotting the appropriate graph

The obtained values are subjected to statistical evaluation in the program Statistica 14 and the results are further discussed in next chapter.

The following Fig. 3 shows the described work procedure.



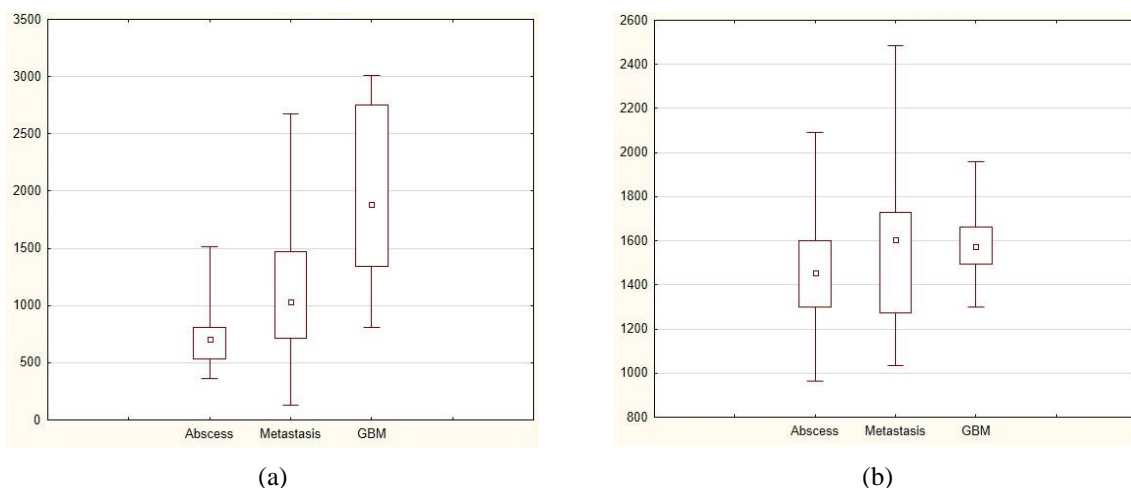
**Figure 3:** Processing procedure

### 3. RESULTS

The subject of the statistical analysis was to determine whether there was a statistically significant difference between the median and the average of the values measured from the above-mentioned pathologies. Statistical analysis was performed in software program Statistica 14. Based on standard procedures for performing statistical analysis, it was necessary to first verify the normality of the data distribution using Shapiro-Wilk test and histograms. The *significance level*  $\alpha$  was set to  $\alpha = 0.05$  in all tested cases, as is known for biological tests [7]. In only one case, the value of  $\alpha$  was exceeded, and the data, therefore, showed a normal distribution only for the measured values of the average at the abscess values.

Subsequently, the mean values were compared between the investigated pathologies. The Wilcoxon paired test revealed that a statistically significant difference occurred between the compared pair of abscess pathologies – GBM ( $p = 0.000898$ ). The compared pairs abscess – metastasis ( $p = 0.273234$ ) and metastasis – GBM ( $p = 0.366749$ ) exceeded the value of  $\alpha$ , and therefore, it is not possible to determine the difference between the investigated pathologies, which can be visually verified by the box graph in Fig. 4 (b). There a small square indicates the median value, the range 25 – 75% is indicated by a rectangle.

Furthermore, the median values were evaluated. A statistically significant difference was found in all three cases of the compared pairs, e.g., abscess – metastasis ( $p = 0.000074$ ), GBM – metastasis ( $p = 0.000002$ ) and abscess – GBM ( $p = 0$ ).



**Figure 4:** Box plot for (a) median and (b) mean value.

#### 4. CONCLUSION

From the available data, it was found that there is a statistically significant difference between the measured values for tumorous objects in the brain, in all three cases of the compared pairs using the median values. The mean values obtained from the lines of interest do not show a statistically significant difference in most cases (abscess – metastasis and metastasis – GBM).

The main advantage of this approach over the previous ones is that the line is drawn in the area of interest instead of placing a (circular) region of interest (ROI) similar as [3], thanks to which it is possible to study the dynamics and rate of change of values of two different tissues (edema - tumor). Also, no article comparing these three pathologies was found. This article will then be followed by other articles and projects that will focus on this problematic in more detail.

#### ACKNOWLEDGMENT

The research was carried out under support of internal grant of St. Anne's University Hospital Brno (FNDN\_IP21IPV15AXA\_PN) and the assistance provided by the general student development project being executed at Brno University of Technology.

#### REFERENCES

- [1] MERKUNOVÁ, Alena and Miroslav OREL. Human Anatomy and Physiology: 4 ort he Humanities. Prague: Grada Publishing, 2008, 304 pp. ISBN 978-80-247-1521-6.
- [2] GIHR G, HORVAT-Rizea D, HEKELER E, Ganslandt O, HENKES H, Hoffmann KT, SCHERLACH C, SCHOB S. Diffusion weighted imaging in high-grade gliomas: A histogram-based analysis of apparent diffusion coefficient profile. PLoS One. 2021 doi: 10.1371/journal.pone.0249878.
- [3] LEE, EJ, K TERBRUGGE, D MIKULIS, DS CHOI, SY MOON a SK LEE. Diagnostic value of peritumoral minimum apparent diffusion coefficient for differentiation of glioblastoma multiforme from solitary metastatic lesions. American Journal of Roentgenology. 2011, 196(1), 71-76: doi:10.2214/AJR.10.4752
- [4] MORI, Susumo. Introduction to diffusion tensor imaging. China: Elsevier, 2009, 176 s. ISBN 978-0-444-52828-5.
- [5] BERG, Heidi Johansen a Timothy E. J. BEHRENS. Diffusion MRI. Elsevier, 2009. ISBN 978-0-12-374709-9.
- [6] SEIDL, Zdeněk a Manuela VANĚČKOVÁ. Diagnostic radiology: Neuroradiology. Prague: Grada Publishing, 2014. ISBN 978-80-247-4546-6.
- [7] PROCHÁZKA, Bohumír. Brief biostatistics for physicians. Charles University in Prague: Karolinum, 2015, 125 s. ISBN 978-80-246-2783-0.

# NXP SEMICONDUCTORS CZECH REPUBLIC



**NXP Semiconductors** se zabývá vývojem a výrobou mikroprocesorů a analogových součástek pro automobilový a spotřební průmysl, počítačové sítě a bezdrátové komunikace. V České republice je společnost NXP Semiconductors zastoupena vývojovými centry v Rožnově pod Radhoštěm a Brně. V moderních laboratořích pracuje mezinárodní tým čítající více než 300 odborníků z oblastí **Vision, SW Tools, Multicore, Motor Control, Security and Connectivity, Embedded Software, Operating Systems, Analog Products, Customer Support, Software Validation and Verification, Industrial Connectivity, Pricing a Marketing** a v neposlední řadě zákaznické centrum spolu s technicko-obchodním oddělením.

**Kariéra ve firmě NXP** znamená krok do budoucnosti – jak profesně, tak osobnostně. Společnost NXP je globálním hráčem v oblasti technologie a vývoje aplikací, měnících náš každodenní život. Náš tým tvoří ambiciózní profesionálové, pracujících v mezinárodním prostředí. **Hledáme nové, motivované a talentované kolegy, kteří se spolu s námi postaví výzvám budoucího vývoje v otevřeném prostředí, které firma NXP nabízí.**





# NXP SEMICONDUCTORS CZECH REPUBLIC

Hledáme talentované STUDENTY



## Právě nabíráme na tyto studentské pozice:

**Embedded HW/SW Developer** – Rožnov, Brno

**SW Technical specialist** – Brno

**Embedded SW Developer** – Rožnov

**Systems Engineering Intern** – Rožnov, Brno

**SW R&D MICR Intern** – Rožnov/Brno

**ML/AI SW Intern** – Rožnov/Brno

**Audio Linux Framework Intern** - Brno

## Studentům nabízíme:

Práci na reálných projektech, technický mentoring

Výborné mzdové ohodnocení a časovou flexibilitu

Pomoc s bakalářskou, diplomovou či disertační prací

Spolupráci s kapacitami svého oboru

[michaela.latiokova@nxp.com](mailto:michaela.latiokova@nxp.com)

[www.nxp.com/czechrepublic](http://www.nxp.com/czechrepublic)

## Zajímáte se o:

*Vývoj Emb. Aplikací*

**Strojové učení/umělá  
inteligence**

*Internet of Things*

**Vývoj validačních /  
testovacích technologií**

*SW Konfigurační nástroje*

**AUTOSAR a multicore  
řešení**

*Embedded operační  
systémy*

**Security v embedded  
systémech**



# An app to demonstrate the risks of granting permissions in Android OS

Jakub Michálek<sup>1</sup>, and Václav Oujezský<sup>2</sup>

<sup>1</sup>Brno University of Technology, The Czech Republic

<sup>2</sup>Brno University of Technology, The Czech Republic

E-mail: [186140@vut.cz](mailto:186140@vut.cz), [oujezsky@vut.cz](mailto:oujezsky@vut.cz)

**Abstract**—This article presents an application for demonstrating the risks of granting unsafe authorizations on Android devices. These permissions allow the application to access sensitive user data. The application uses practical examples to illustrate what user data can be abused by enabling permission. Nowadays, the security of user data is a growing topic. Therefore, mobile users themselves must be aware of the risks they expose themselves to by granting permissions.

**Keywords**—Android, Kotlin, mobile application, permissions

## 1. INTRODUCTION

Almost everyone has a smartphone these days. Over 70 % of them are based on the Android operating system. As these devices provide users with many different functions, they are increasingly used for security-critical purposes such as internet banking, payment confirmation, or logging into online accounts. They also constantly collect sensitive user data such as text messages, location information, call history, photos, and videos.

Permissions allow support user privacy by preventing applications from accessing confidential user data and confidential actions. Each android application runs in its own isolated space and by default has no permission to perform any operations that could potentially adversely affect other applications, the operating system, or the user. Therefore, if it needs to use resources or information outside of this isolated space, it needs to declare permissions to access those resources and set up requests for those permissions. The three most important permissions groups are defined, namely permissions granted at installation, application runtime, and special permissions. Each of these groups has a limited range of data that can be accessed and actions that the application can perform once granted [1].

Permissions granted while the application is running are also known to be dangerous. Permissions give access to sensitive user data to an application and allow it to perform actions that significantly affect the system and other applications. In Android 6.0 and above, permissions are required when the application runs. In earlier versions, the operating system required all permissions to be enabled when an application is installed on a device.

This article describes a developed application whose purpose is to demonstrate to users one of the security risks of Android, namely the granting of insecure permissions. Enabling these permissions on a device allows sensitive data to be extracted from the user's device. The risks of each permission are illustrated with practical examples directly in the application.

## 2. THE APPLICATION OF THE RISKS OF SECURITY AUTHORISATIONS

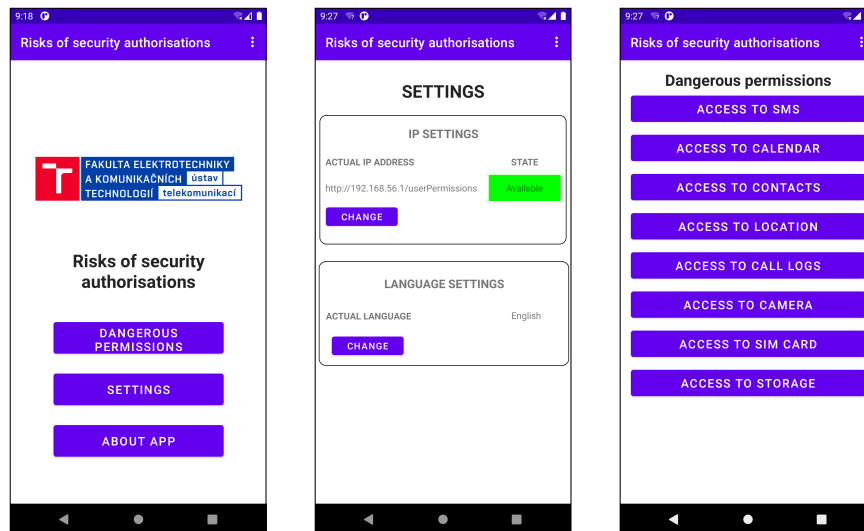
### 2.1. Proposal

The application's primary purpose is to present the risks of granting dangerous permissions to users. It consists of two parts – mobile and server. The mobile part is programmed using Kotlin language, the one of the official programming language used today to create Android applications. The server part uses the PHP (Hypertext Preprocessor) language to communicate with the MySQL database and the HTML (Hypertext Markup) language to display data from the database and interact with the user.

### 2.2. Implementation

A user is presented with a menu that allows them to navigate to examples with risk permissions, settings and information about the application, shown in Figure 1. In the settings, it is possible to change the





**Fig. 1:** The main menu screen, settings and list with the examples of dangerous permissions.

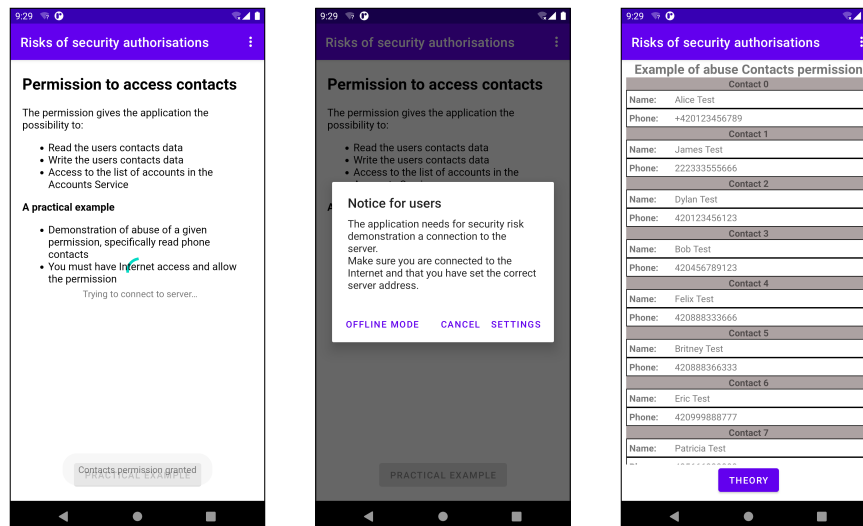
application language and the address of the server with which the application communicates. The server availability is required to use all functions. If the user navigates to the risk permissions examples, a menu with eight different permissions is displayed, which are listed with their function in Table I.

After selecting one of the examples, the user is first introduced to what is involved in enabling this permission and what is required to view the practical example. The user's account is also created in the database on the server. If a practical example is selected, the server is available, and the permission has already been enabled, sensitive data and photos are uploaded. If the server is unavailable, a dialog with the option to display an offline example is displayed, as is shown in Figure 2. If this option is selected, the offline example is displayed with the data intended to be sent to the server. If the permission has not been granted, a dialog is displayed depending on the user's previous decisions. If required for the first application run, the dialog contains an option to allow or deny permissions. Suppose the dialog has already been rejected and the user still wants to go to a practical example. In that case, the permission must be approved directly in the phone settings in the application manager.

The practical example then shows what sensitive data can be obtained with the selected permission. The web application loaded in the mobile is used for this purpose. The user's login credentials are generated and after logging in, it is shown that the sensitive data is outside the device itself and could be used by a potential attacker to his advantage. When the application exits, all sensitive user data stored on the server are deleted. The server hosts a web application and a MySQL database. The web application is used to display data from the database that has been stolen from a device using the mobile application to demonstrate the dangers of granting permissions. It is a simple application that consists of a user

	Permissions used	Allows
1	READ_CALENDAR	Read calendar events
2	READ_CALL_LOG	Read call logs
3	CAMERA	Access camera
4	READ_CONTACTS	Get contacts
5	ACCESS_FINE_LOCATION	Access phone location
6	READ_PHONE_STATE	Get SIM card information
7	READ_SMS	Read SMS messages
8	READ_EXTERNAL_STORAGE	Access files in storage

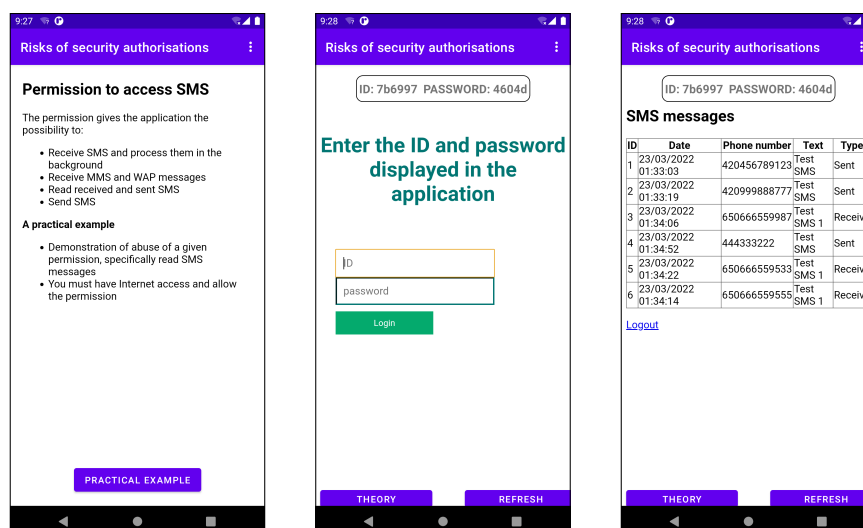
**Tab. I:** The permissions used in the application.



**Fig. 2:** A screen with the theory (left), unavailable server dialog (center) and an offline practical example of Contact access permission abuse.

login/logout page and a database data display page. The login page contains a form that requests the user's ID and password. This data is generated by the mobile application itself and stored in a database located on the server. If the correct login credentials are entered, the user is redirected to a page displaying the data from the database. If the user is not in the database an error message is displayed stating that the login details are incorrect. After a successful login, the retrieved data from the user's device stored in the database is displayed in the form of tables based on the login data. Figure 3 shows the example of SMS access permission abuse. If the stored data also includes photos from the user's devices, these are also displayed. The MySQL database is used to store the device data. It contains a static table for storing the users of the application. Tables storing data directly from devices using the mobile application are dynamically added depending on the currently displayed practical example.

Communication between the mobile application and the webserver is only one-way and takes place from the application to the webserver. Functions from the Volley library are used to send the data. This is the official HTTP (Hypertext Transfer Protocol) and HTTPS (Secure) library developed by Google for fast and easy communication over the Internet on Android [2]. In the presented version of the application, HTTP protocol communication is used. PHP scripts are created on the webserver to receive data from the application.



**Fig. 3:** A screen with the theory (left) and a practical example of SMS access permission abuse.

### 2.3. Verification

The application is now in the development stage and testing phase. The latest version of the application is available on the developer's website [3] and the source code on GitHub [4]. A video demonstrating all the functionalities of the application is uploaded on the YouTube platform [5].

The full functionality of the latest version of the application has been tested by using three tests. The first test involves testing the application in the mobile device emulator in the Android Studio IDE using a local server. This test is used to detect critical errors and complex functions. The second test is a user test with a server located on the hosting. This test involves testing the application on several real devices. It is used to detect errors that may not appear on the emulator. It also tests the functions for sending data from the device. The third test is performed in the Firebase CodeLab environment. This test is used to test whether the application is stable.

The application has passed all the tests mentioned above. All critical errors and shortcomings have been identified and corrected. As the testing phase is still ongoing, the application may contain some other bugs that have not yet been detected but should not affect the main functionality.

### 3. CONCLUSION

The aim of the paper was to present the developed application "Security of risk permissions". The purpose of this application is to present to users the risks that can occur when granting unsafe permissions. Enabling these permissions then gives the application the ability to access sensitive data stored on the mobile device and misuse it by an attacker without any knowledge of the application user. The risks posed by each permission are first described in a short theory and then demonstrated with a practical example. With the help of this application, the user can realize the risks that can arise by unthinkingly approving any permissions and thus protecting his sensitive data from applications that he does not trust completely. In the next version of the application, HTTPS protocol support will be implemented to communicate with the webserver and more examples will be added.

### REFERENCES

- [1] Developers, "Permissions on android," Google LLC, 2022. [Online]. Available: <https://developer.android.com/guide/topics/permissions/overview>
- [2] V. Tyagi, "Volley library in android," GeeksforGeeks, 2022. [Online]. Available: <https://www.geeksforgeeks.org/volley-library-in-android/>
- [3] J. Michalek, "Security application," 2022. [Online]. Available: <http://unsecureapp.tode.cz/>
- [4] vaklur, "Userpermissions," GitHub, Inc., 2022. [Online]. Available: <https://github.com/vaklur/UserPermissions>
- [5] J. Michalek, "Application – risks of security authorisations," Youtube, 2022. [Online]. Available: <https://www.youtube.com/watch?v=MntV5ga39Ls>

# Traffic Analysis Using Machine Learning Approach

O. Zelený<sup>1</sup> and T. Frýza<sup>1</sup>

<sup>1</sup>Department of Radio Electronics, Faculty of Electrical Engineering and Communication  
Brno University of Technology, Technická 12, 616 00 Brno, Czech Republic

E-mail: [xzelen23@vut.cz](mailto:xzelen23@vut.cz)

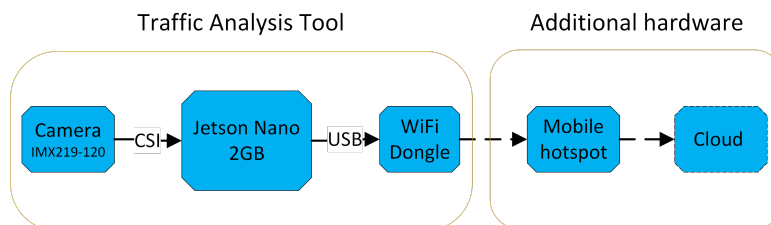
**Abstract**—This paper provides insight to the YOLOv5 deep learning architecture and its use for vehicle detection and classification in order to improve traffic management in larger cities and busy roads. The paper presents simple system with one fixed camera and Jetson Nano, a computer for embedded and AI application, to detect and classify vehicles.

**Keywords**—Deep learning, Computer vision, Traffic analysis, Convolutional Neural Networks, You Only Look Once, COCO dataset

## 1. INTRODUCTION

For the past few decades the traffic in large cities and highways has become source of many problems such as efficiency and safety. These reasons lead cities to started deploying means of traffic control to allow smooth and safe flow of traffic. The main instrument of traffic control are traffic lights, which however, have the disadvantage of having "fixed" time when the green light is on. In order to improve the flow of traffic, system analyzing the traffic in real time is needed [1][2][3]

This paper proposes system based around Jetson Nano, a platform for artificial intelligence, which is used to handle the task of object detection and classification from live video feed captured by a camera. The block diagram of described system is shows in Figure 1.



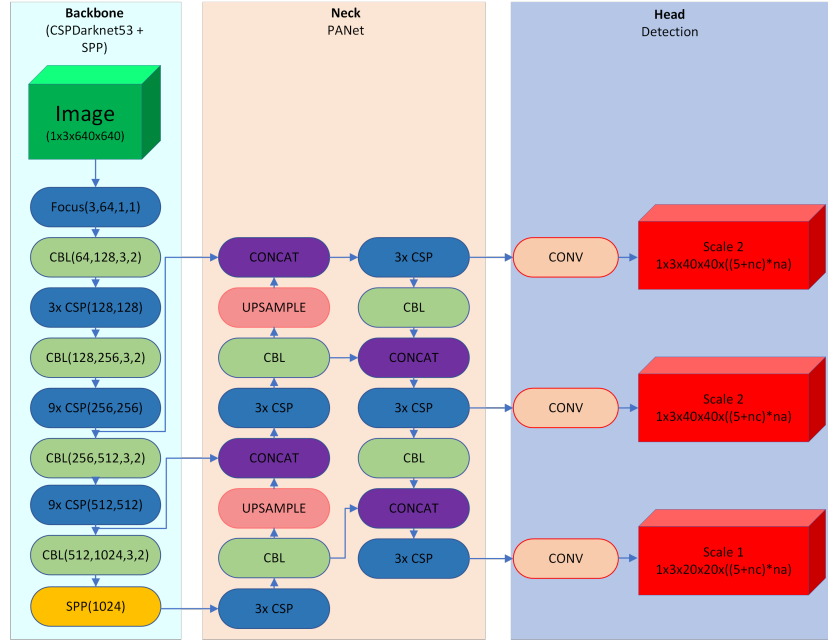
**Figure 1:** Proposed system for traffic flow analysis

In machine learning, a machine is trained by data and algorithms to learn how to perform a given task. Deep learning is considered a subset of machine learning that uses a programmable neural network to allow machines to make decisions without the help of humans. Algorithm running on Jetson Nano utilizes Neural Networks, which consist out of layers of neurons, a basic building block of neural networks inspired by human brain [4].

## 2. MODEL

YOLO (You Only Look Once) architecture is one of the fastest architectures for object detection and classification and it has become very popular in recent years. The architecture has gone through 5 iterations of development and the newest version is YOLOv5 [5] by Glenn Jocher. This architecture, as well as its previous version, is fully convolutional and it uses three difference scales in order to extract wide variety of features for detection of small and large objects. Each of these scales divides image or feature map into  $n$  by  $n$  grid, where each cell predicts bounding box for object which center is located in each particular cell. The architecture of YOLOv5 consists of slightly modified CSPDarknet53 in combination with SPP (Spatial Pyramid Pooling), which are used to extract important features and PAN (Path Aggregation Network), used to further improve the extracted features. The backbone consists of

CBL (Convolutional Base Layer) and CSP (Cross Stage Partial) network which are repeated three or nine times. Detailed view of the architecture of YOLOv5 can be seen in Figure 2.



**Figure 2:** Architecture of YOLO version 5

### 3. COCO DATASET

COCO (Common Objects In Context) [6] is a large dataset containing 123 thousand annotated images with more than 886 thousand instances divided into 80 classes. The dataset is split into 118 thousand training images and only 5 thousand validation images. Since the dataset contains a large amount of classes which our system does not need, only some classes are used for the learning process. In particular, the classes that will be used are Car, Motorcycle, Bicycle, Truck, and Bus. The exact number of instance per each class can be seen in Table I.

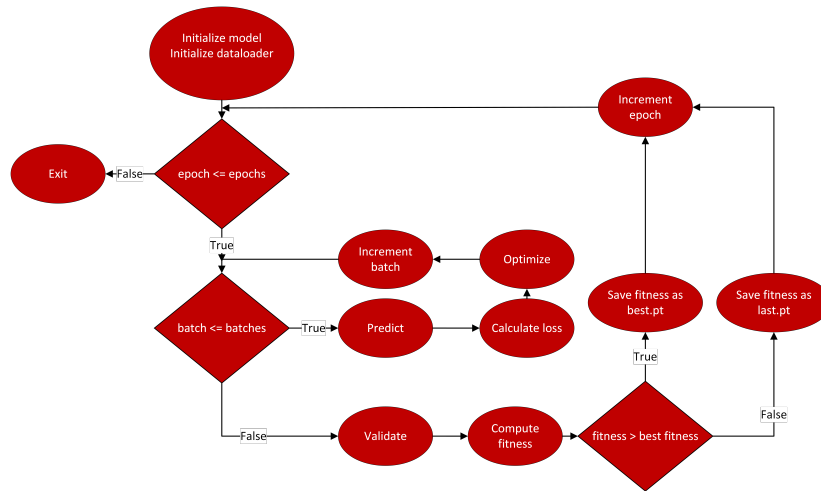
**Table I:** Selected instances in COCO dataset.

Class	Number of instances
Car	12,786
Motorcycle	3,661
Bicycle	3,401
Bus	4,141
Truck	6,377

### 4. IMPLEMENTATION

All of the code needed for the model to work was implemented in Python using PyTorch framework. Figure 3 shows flowchart of the training loop which incorporates numerous techniques and option to improve the optimization. The training algorithm uses numerous techniques of data augmentation including mosaic, rotation or translation. The network was optimized using SDG (Stochastic Gradient Descent) optimizer which has shown better results than Adam optimizer. In order to further improve the learning process, the algorithm computes four different metrics based on which it evaluates the fitness of the network. The algorithm save model parameters both for last and best epoch, which also allows the algorithm to resume optimization in case of interruption. Due to complexity of learning process and high hardware requirements for faster training, Google Colab, a cloud based environment was used, since it

allows for utilization GPU (Graphics Processing Unit) like Tesla P100 with 16GB of memory, which was used to train the network.

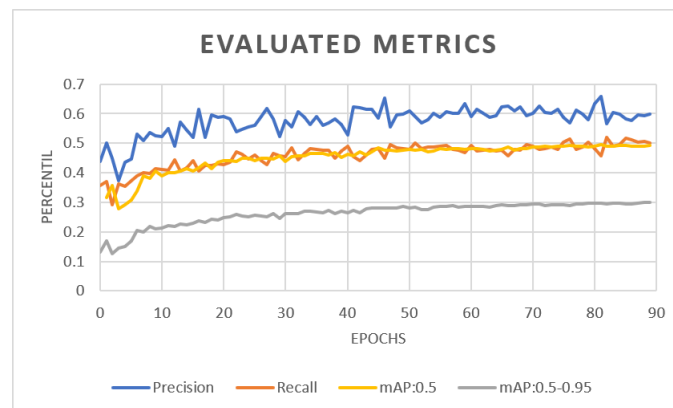


**Figure 3:** Training loop.

The detection or inference handled by separate Python file allows the use of images, videos or cameras as a source of data for detection. The model utilizes saved parameters of the model from learning process to make predictions to each individual cell. Using thresholds and non-maximum suppression only valid predictions are sorted out and their bounding boxes are added to the image.

## 5. OPTIMALIZATION

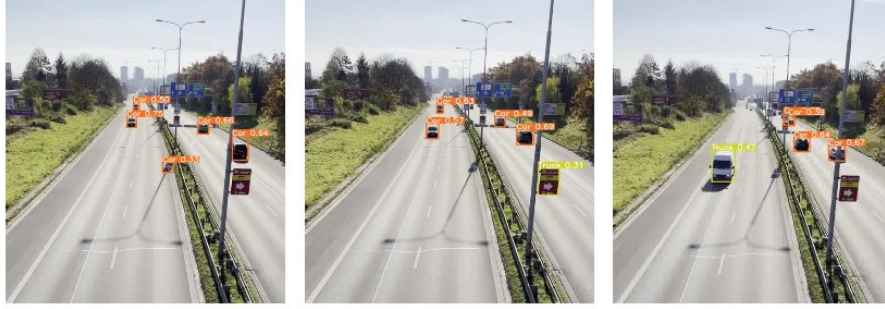
The algorithm uses pretrained weights on COCO dataset and fine-tunes them to the subset of COCO dataset containing only vehicles. This is necessary due to the small amount of data containing vehicles due to which the algorithm was not able to reach reasonable precision when learning from scratch. The precision was mainly evaluated based on mAP (Mean Average Precision) with the threshold of 50% and mAP with thresholds from 50 to 95% with step of 5%. The evolution of evaluated metrics is show in Figure 4.



**Figure 4:** Metrics evolvement over time.

The fitness of the networks was evaluated based on mAP:0.5 and mAP:0.5-0.95 with the contribution of 10 and 90% respectively. Starting from pre-trained weights, the training took approximately 18 hours on Tesla P100 GPU with the batch size of 16 and image size 640 by 640. The YOLOv5 model has numerous subversion which have been tested and all of them can be handled by Jetson Nano in terms of computational capacity, however, due to the importance of inference speed, the small version was selected for its performance in speed and accuracy. The real functionality of the system was tested on scenario with low traffic from Hradecka street in Brno, which is shown in Figure 5.





**Figure 5:** Detection on real scenario

The precision of the system in real scenario is highly dependant on the camera location and the angle at which it captures the road. The system was tested both with the camera facing the incoming traffic and also from the side of the road, where detections made from the side of the road were more confident as well as it had less false detections.

## 6. CONCLUSION

The presented system reaches average precision about 65% and is able to sufficiently detect vehicles with small amount of false detections. The implementation is currently able to run at approximately 30 frames per second on GTX960 and 9 frames per seconds on Jetson Nano. The model shows good performance in classifying vehicle classes, whose precision can be seen in Table II.

**Table II:** Precision on specific class.

Class	Precision
Car	75.5%
Motorcycle	61.7%
Bicycle	47.3%
Bus	83.6%
Truck	55.0%

The project will continue its development focusing on improvements in accuracy and also object tracking, which is necessary in order to track vehicles through time and to compute the speed of the vehicles. Further, the algorithm will be sending data to cloud, which will process the data for important statistical information. The complete system will be then evaluated in long-term test on more frequent roads with the focus on performance in generating statistical data.

## REFERENCES

- [1] Chauhan, M., Singh, A., Khemka, M., Prateek, A. & Sen, R. Embedded CNN based vehicle classification and counting in non-laned road traffic. *CoRR*. **abs/1901.06358** (2019), <http://arxiv.org/abs/1901.06358>
- [2] Najeeb, S., Raza, R., Yusuf, A. & Sultan, Z. Fine-Grained Vehicle Classification in Urban Traffic Scenes using Deep Learning. *CoRR*. **abs/2111.09403** (2021), <https://arxiv.org/abs/2111.09403>
- [3] Satar, B. & Dirik, A. Deep Learning Based Vehicle Make-Model Classification. *CoRR*. **abs/1809.00953** (2018), <http://arxiv.org/abs/1809.00953>
- [4] Chollet Francois, "Deep learning with Python, Second edition," Shelter Island, NY: Manning Publications, 2021. ISBN 1617296864
- [5] Jocher, G. YOLOv5. *GitHub Repository*. (2020), <https://github.com/ultralytics/yolov5>
- [6] Lin, T., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Dollár, P. & Zitnick, C. Microsoft COCO: Common Objects in Context. *CoRR*. **abs/1405.0312** (2014), <http://arxiv.org/abs/1405.0312>

# Planning of Single Frequency Networks for DVB-T2

Vojtech Barta<sup>1</sup>, Ladislav Polak<sup>1</sup>, Pavel Cidl<sup>2</sup>

<sup>1</sup>Brno University of Technology, Faculty of Electrical Engineering and Communication,  
Dept. of Radio Electronics, Technicka 12, 616 00 Brno, Czech Republic

<sup>2</sup>Czech Telecommunication Office, Jurkovicova 244/1, 638 00 Brno-Lesna, Czech Republic

E-mail: [xbarta45@vut.cz](mailto:xbarta45@vut.cz), [polakl@vut.cz](mailto:polakl@vut.cz), [cidlp@ctu.cz](mailto:cidlp@ctu.cz)

**Abstract:** Nowadays, especially in Europe, the second generation of Digital Video Broadcasting Terrestrial system, abbreviated as DVB-T2, gradually replaces its predecessor DVB-T. Compared to DVB-T, DVB-T2 is a more versatile system with several configuration settings. Thanks to this, the planning of Single Frequency Networks (SFNs) used to effectively cover a certain area by DVB-T2 signal can be more flexible and straightforward. This paper deals with the design of SFNs for DVB-T2, especially with focus on the study of DVB-T2 system parameters (e.g. guard interval) and transmission conditions (e.g. multipath propagation) directly affecting performance of SFNs. For this purpose, RadioLab, RadioBase and Networks SFN utilities from CRC data software are used. Real DVB-T2 system configuration for Multiplex 21 (MUX 21) used in Brno (Czech Republic) is employed in this study. The simulation-based results can help to identify places with lower signal coverage or interfering sources as well as to find appropriate approaches to solve these issues.

**Keywords:** DVB-T2, SFN, guard interval, multipath propagation, RadioLab

## 1. INTRODUCTION

The second generation of Digital Video Broadcasting Terrestrial (DVB-T2) system [1], compared to first generation DVB-T [1], excels with very flexible configuration of the signal processing chain. Thereby, the broadcasting of TV signal can be adapted for different transmission scenarios (mobile, portable, fixed). Thanks to this, the so-called Single Frequency Networks (SFNs) [2] used to cover a land by a DVB-T2 signal can be planned more effectively. The performance of SFNs is influenced not only by the transmission conditions of a certain area, planned to be cover, but also by the used system configuration of DVB-T2.

This short paper deals with the design and performance analysis of SFNs for DVB-T2, especially from the viewpoint of different transmission (e.g. multipath propagation) and DVB-T2 system parameters (e.g. the length of guard interval – GI). For this purpose, software utilities from CRC Data<sup>1</sup>, developed for planning of SFNs, are employed. The simulation-based results can help to identify problematic places at signal coverage of a certain area as well as to help find and appropriate solution to fix issues.

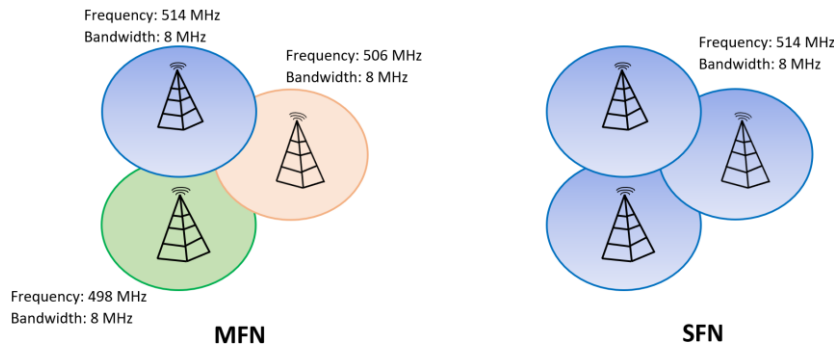
The paper is organized as follows. The basic concept of SFNs and their features are briefly introduced in Section 2. The design of SFN and evaluation of simulation-based results are discussed in Section 3. Finally, Section 4 concludes this paper.

## 2. SFN IN DVB-T/T2

SFN is an effective way to cover a certain area of land by a TV signal. In such a network, compared to mono frequency network (MFN), several transmitters (TXs) simultaneously broadcast the same signal over the same frequency (TV) channel. Illustration of an SFN network is shown in **Figure 1**. Efficient utilization of the available radiofrequency (RF) band and increasing the coverage area are the main aims at the employing of SFNs. To achieve these ones, strict requirements must be fulfilled at the planning of SFNs [2]. SFN schemes can only be built in a certain area, not in whole country.

---

<sup>1</sup><https://www.crcdata.cz/index.php>



**Figure 1:** Comparison of MFN versus SFN.

The terrestrial transmission channel is rich in multipath propagation [1] caused by reflections of the signal from different obstacles (e.g., buildings, trees). Thanks to this, several signals with different power, delay and phase shift are received at the antenna of receiver (RX), which can result in inter-symbol interferences (ISI). DVB-T2 is a Coded Orthogonal Frequency Division Multiplex (COFDM) based system employing to choose between several lengths of so-called guard interval (GI) [1]. At the appropriate set of GI noises due to ISI can be suppressed.

As it was mentioned, the planning of SFNs is conditioned by several factors. The distance between TXs must set carefully to minimize interfering of TXs by each other. This distance depends on the length of GI and the selected so called OFDM mode [1]. From this point of view, SFN is a special case of multipath propagation. In some cases, partial contributions of TXs to the signal at the RX not only do not interfere but, in certain circumstances, can improve the conditions of signal reception [1]-[6].

The employing of SFNs has several advantages. For instance, the available RF band can be managed effectively (transmission of higher number of TV and radio programs). Next, the transmission power is used more efficiently and in the case of signal leakage from one TX, signal from other TX can be used. On the other hand, SFNs have also some disadvantages. The TXs and broadcasted data must be synchronized in time and frequency. If one TX violates the rules of SFN then it becomes as an interferer for the other TXs. This requires continuous monitoring the performance of SFNs [1]-[3].

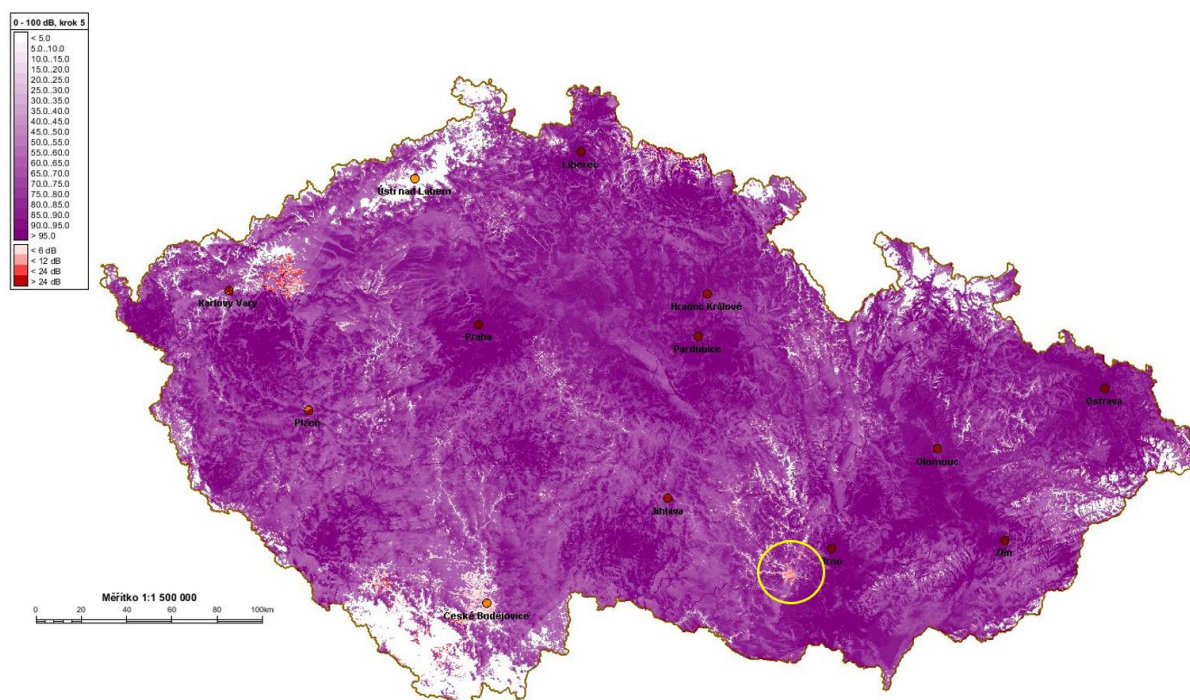
### 3. PLANNING AND EVALUATION OF SFN

In this work, RadioLab, RadioBase and Networks SFN utilities from CRC data software were used to planning SFNs and study their performances. RadioLab is a computational system for the analysis and visualization of the propagation of radio signals over the Earth's surface (in free space). Among others, it can be used to analyze terrain profile, calculate signal propagation attenuation and calculate area distribution of electromagnetic field strength. RadioBase is a data subsystem for managing data of area coverage transmitters. Networks SFN is a computational system for coverage and interference analysis of different services (FM, DVB, DAB). It is also used for the design of SFNs for DVB-T/T2 and T-DAB. Among others, the system is capable for evaluating interferences from other broadcast systems.

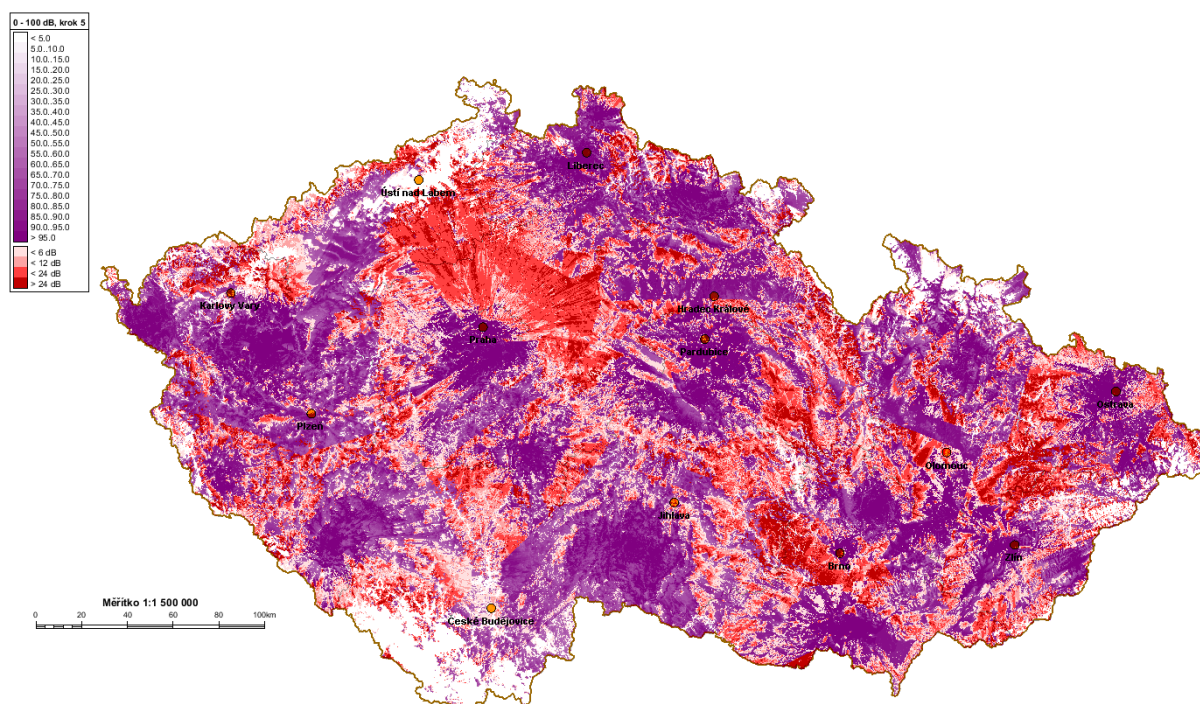
To study the DVB-T2 signal coverage of Czech Republic, broadcast parameters used for Multiplex 21 (MUX 21, Channel 26) were used (see the caption of Figs. 2 and 3). The bandwidth of TV channel is 8 MHz. The individual transmitters together with their real parameters from the Czech Telecommunications Office (ČTÚ) database were uploaded to the program. The ITU-R P.1812-4 propagation model was used to calculate the signal diagrams of the useful and interfering signals from TXs. The way in which the receiver evaluates the individual signals against GI was the following. It was chosen that the RX "hangs on" to the first suprathreshold signal, i.e. a signal with intensity greater than the minimum level for a given SFN.

The coverage for the selected DVB-T2 implementation is shown in **Figure 2**. It is visible that practically the whole land of Czech Republic is covered by sufficiently strong signal (highlighted in magenta). There are only a few places where interfering signals can occur (see red colors near to city Karlovy Vary). Such a phenomenon is probably caused by either an interfering TX or multipath signals with a delay higher than the duration of GI (rising of ISI). The GI has a length of 448  $\mu$ s. Areas uncovered by the signal of MUX21 (see white color near to cities České Budějovice and Ústí nad Labem) are dominantly occurred at places near to borders of Czech Republic.





**Figure 2:** Coverage of the Czech Republic by DVB-T2 signal: Multiplex 21 (Channel 26), guard interval 1/8, OFDM mode 32K Extended, 256-QAM modulation, code ratio 2/3, Pilot Pattern PP8, rotated constellation, working frequency 514 MHz.



**Figure 3:** Coverage of the Czech Republic by DVB-T2 signal: Multiplex 21 (Channel 26), guard interval 1/128, OFDM mode 32K Extended, 256-QAM modulation, code ratio 2/3, Pilot Pattern PP8, rotated constellation, working frequency 514 MHz.

An inappropriate configuration of SFNs is captured in **Figure 3**. For this scenario, the duration of GI was changed from 448  $\mu$ s to 28  $\mu$ s. It represents the lowest value for duration of GI in the system DVB-T2. As a result, strong and destructive interference between TXs of SFN can be observed in the whole country of Czech Republic (see red colors mainly near to cities Brno and Prague). Due to

deleterious self-interference effects, only a few locations, highlighted in magenta, are covered with a strong signal.

#### 4. CONCLUSION

In this brief work, planning of SFN for DVB-T2 used Czech Republic was explored and provided. For this purpose, professional software tools from CRC Data were employed. Simulation-based results confirmed theoretical assumptions, i.e. the right level of signal of TXs and the duration of GI play key roles in prevention against deleterious interference effects in SFNs. All the results, presented in this paper, are a part of preliminary diploma thesis.

Yellow circle in **Figure 2** marks the area near to city Brno, where simulation-based analysis revealed problems with the reception of TV signal from MUX 21. First experimental measurements in this terrain confirmed lower quality of TV signal. In the future, attention will be devoted to this certain area. In cooperation with ČTÚ, extensive measurements in this field will be conducted. After the analysis of these results, appropriate approaches will be proposed to improve the signal reception in this problematic area. For this purpose, the above mentioned SW tools will be used.

#### ACKNOWLEDGMENT

This paper was supported by an internal grant of the Brno University of Technology FEKT-S-20-6325.

#### REFERENCES

- [1] FISCHER, Walter a Horst von RENOUARD, 2008. Digital video and audio broadcasting technology: a practical engineering guide ; 2<sup>nd</sup> ed. Berlin: Springer. Springer series on signals and communication technology. ISBN 978-3-540- 76357-4.
- [2] RICNY, V. Single Frequency Networks (SFN) in Digital Terrestrial Broadcasting. Radioengineering, vol. 16, no. 4, Dec. 2007, pp. 2-6. ISSN: 1210-2512.
- [3] ITU-R, Frequency and network planning aspects of DVB-T2, Report ITU-R BT.2254-4 2020
- [4] NEPAL, Subodh, et al. A Comparative Study on Radio Coverage Due to DVB-T2 Multi-Frequency Network (MFN) and Single Frequency Network (SFN). In: 2020 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB). IEEE, 2020. p. 1-6, [doi 10.1109/BMSB49480.2020.9379904](https://doi.org/10.1109/BMSB49480.2020.9379904)
- [5] RUCKVERATHAM, Bundit; PROMWONG, Sathaporn. Evaluation of SFN gain for DVB-T2. In: 2017 International Conference on Digital Arts, Media and Technology (ICDAMT). IEEE, 2017. p. 85-88. [doi 10.1109/ICDAMT.2017.7904940](https://doi.org/10.1109/ICDAMT.2017.7904940)
- [6] RUCKVERATHAM, Bundit; PHOMWONG, Sathaporn. A Study of Single Frequency Network for DVB-T2 base on Measurement Data. In: 2018 Global Wireless Summit (GWS). IEEE, 2018. p. 379-382. [doi 10.1109/GWS.2018.8686516](https://doi.org/10.1109/GWS.2018.8686516)

# An Anti-spam Software Plugin For E-Mail Classification Based On Geographical Location Of The Sender

M. Kovařík<sup>1</sup> and T. Caha<sup>1</sup>

<sup>1</sup>Department of Telecommunications, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technická 12, 616 00 Brno, Czech Republic

E-mail: [190001@vut.cz](mailto:190001@vut.cz), [tomas.cahal@vut.cz](mailto:tomas.cahal@vut.cz)

**Abstract**—Every e-mail sent must be transferred by one or more mail servers. Some of these servers are connected to the Internet using a public IP address. IP addresses can be geographically located, e.g., using geolocation databases. This means e-mails can be classified by the location of the sender. This paper introduces an anti-spam plugin for such classification. The plugin is implemented for SpamAssassin anti-spam software. The basic functionality was implemented. By using this plugin, one can define a penalty score for specific countries as e-mail origins.

**Keywords**—anti-spam plugin, e-mail, geolocation, IP2Location, SpamAssassin

## 1. INTRODUCTION

The problem of spam has been ongoing for several decades now and it is not going anywhere anytime soon. Spam has many forms, but it is mostly spread through e-mail messages. Since the cost of sending spam is nearly non-existent and anyone can do it, there needs to be a protection from it[1]. This job is done by e-mail filters. These filters decide if the message is a spam or not. Filtering techniques range from simple whitelists or blacklists to more complex Bayes filtering[2].

Geolocation of an IP address has many uses, including an automatic selection of language and currency on a website, more relevant advertisements, news, weather forecast, and cybersecurity. IP geolocation can be divided into 2 categories – active and passive methods. Active methods usually rely on measuring RTT (Round-Trip Time) from known locations to the sought end station. Passive methods use established databases containing updated information on a range of IP addresses[3].

This paper presents a tool as one of the possible ways to be used in spam filtering. The developed plugin called *Geolock* grants the user the ability to block all e-mails sent from specific countries.

## 2. ORIGIN OF E-MAIL

An e-mail is a text file with the structure defined by RFC 5322. The message is divided into two parts, namely a header and an optional body[4]. Generally speaking, the message body contains the actual text of the message. Nowadays, research focuses on the analysis of the text of the message body, but this is not the subject of this paper. Listing 1 shows an excerpt of header of the e-mail. Header fields consist of header name followed by a colon and by value.

When trying to determine the geographical origin of the e-mail, *Received* fields in the header can be used. When a mail server (in particular Simple Mail Transfer Protocol – SMTP server) receives a message for delivery or further processing, it must alter the message by prepending *Received* information. Simply put, the *Received* header closest to the sender is the lowest among other headers and the *Received* header closest to the recipient is the highest. RFC 5321 also states that the SMTP server must not change the order of existing lines or insert *Received* lines in any other location[5].

If all the above-mentioned conditions are met, the geographical origin of the e-mail is the last (the lowest) *Received* in the header of the e-mail.



### Listing 1: Excerpt of a header of an e-mail

```
Received: from mandark.labs.netnoteinc.com ([213.105.180.140]) by
dogma.slashnull.org (8.11.6/8.11.6) with ESMTP id g4SK1YO24988 for
<jm@jmason.org>; Tue, 28 May 2002 21:01:34 +0100
Received: from hotmail.com (kbl-mdb6237.zeelandnet.nl [62.238.24.141]) by
mandark.labs.netnoteinc.com (8.11.2/8.11.2) with SMTP id g4SK1R731575 for
<jm@netnoteinc.com>; Tue, 28 May 2002 21:01:28 +0100
Received: from sparc.zubilam.net ([146.172.86.49]) by
a231242.upc-a.zhello.nl with esmtp; Tue, 28 May 0102 11:05:33 -0300
Received: from unknown (118.186.232.151) by sparc.zubilam.net with NNFP;
28 May 0102 08:00:33 +0700
Received: from unknown (75.72.44.176) by rly-xw05.oxyeli.com with asmt;
28 May 0102 14:55:33 +0200
Received: from unknown (HELO rly-xr02.nikavo.net) (75.249.246.124) by
rly-xw05.oxyeli.com with asmt; 28 May 0102 16:50:33 +0300
Reply-To: <Sheila7316x53@hotmail.com>
Message-Id: <003b45b76c1e$6862a6d3$3da65ae7@dnvdoi>
From: <Sheila7316x53@hotmail.com>
To: Preferredfriends@aol.com
Subject: Learn How To Make $8,000 within 7-14 days! 3341iYj15-964K-13
Date: Tue, 28 May 0102 20:37:24 -0100
```

### 3. DESCRIPTION OF DEVELOPED PLUGIN *GEOLOCK*

The custom plugin, called Geolock, was developed for SpamAssassin, an open-source anti-spam platform, in the latest version 3.4.6[6]. SpamAssassin was chosen because it is one of the most popular and most used anti-spam software solutions available. SpamAssassin is written in Perl and all its components, including custom plugins, have to be written in the same programming language. This means that the plugin Geolock is written in Perl.

The plugin uses the IP2Location geolocation database. IP2Location offers free and commercial databases. As of now, the plugin is using the free database IP2Location LITE DB5 with limited accuracy. The DB5 database contains the country, city, and coordinates of an IP address[7]. IP2Location was used because of its high accuracy and low error rate compared to its competitors.

The plugin consists of 2 files, *Geolock.pm* and *Geolock.cf*, where the *Geolock.cf* file is a configuration file for the whole plugin, where among other things, the user chooses what countries one wishes to block. The *Geolock.pm* file contains the plugin itself.

The file *Geolock.cf* contains the following lines:

#### Listing 2: Configuration file *Geolock.cf*

```
header    BLOCKED_COUNTRY    eval:get_country()
score     BLOCKED_COUNTRY    50.0
describe  BLOCKED_COUNTRY    The country of origin is blocked

rule UK US SK

add_header all Country The country of origin is _MYTAG_
```

The first line describes what subroutine from the Perl module should be evaluated. The subroutine returns 1 as true or 0 as false. The second line sets how many points is the e-mail awarded in case of the rule being true. The official recommended value for a message to be marked as spam is 5. Since some users may have set the value higher, the awarded score for this plugin is set at 50 points to be certain that the plugin blocks the selected country. The third line describes what message is added to the tests triggered section of the SpamAssassin header added to the e-mail header.

The line starting with "rule" sets the list of blocked countries. The countries have to be separated by space and the user can block from 0 countries up to how many countries he wishes. In this example, the countries the United Kingdom, United States of America, and Slovakia are being blocked. While further research should indicate if a finer definition of blocked origin is needed, the plugin is easy to modify to block based on cities or regions.

The last line tells SpamAssassin always to add a message to the SpamAssassin header. Variables in this line are set in the plugin itself depending on the country of origin and if the rule is triggered.

The *Geolock.pm* file has 2 main parts. The first one can be seen in listing 3. The subroutine is responsible for reading the selected countries to block from the configuration file *Geolock.cf* and saving them for later usage for comparison against the country of origin of the e-mail.

**Listing 3:** Subroutine `parse_config` in `Geolock.pm`

```

sub parse_config {
    my ($self, $opts) = @_;
    my $key = $opts->{key};
    my $rule = "";

    if ($key eq "rule") {
        $rule = $opts->{value};
        $opts->{conf}->{"rule"}{rule} = $rule;
        $self->inhibit_further_callbacks();
    }
    return 0;
}

```

The subroutine seen in listing 4 is the one being evaluated. Firstly the subroutine gets the IP addresses of e-mail servers the e-mail went through and saves them into an array. For it to be in chronological order, the array is reversed and the first element is looked up in the locally saved geolocation database. Lastly, the returned value is compared to the list of blocked countries, and in case there is a match, the subroutine returns 1 as true and sets the variables for the correct message to be displayed in the e-mail header.

**Listing 4:** Main subroutine `get_country` in `Geolock.pm`

```

sub get_country {
    my ($self, $pms) = @_;
    my $msg = $pms->{msg};

    my @countries;
    foreach my $relay (@{$msg->{metadata}->{relays_untrusted}}) {
        push(@countries, $relay->{ip});
    }
    my @reverseC = reverse(@countries);
    my $countryshort;

    require Geo::IP2Location;
    my $obj = Geo::IP2Location->open("/etc/mail/spamassassin/DB/IP2LOCATION-LITE-DB5.IPV6.BIN");

    if (!defined($obj)) {
        print STDERR "Geo::IP2Location::get_last_error_message()";
    }
    $countryshort = $obj->get_country_short($reverseC[0]);
    $obj->close();

    my $rule = $pms->{conf}->{"rule"}{rule};
    if (!length $rule){
        dbg("Geolock: No blocked country found in the configuration file.");
    }
    my @array;
    @array = split ' ', $rule;

    if ( grep( /$countryshort/, @array ) ) {
        $pms->set_tag("MYTAG", "BLOCKED : $countryshort");
        return 1;
    }
    else {
        $pms->set_tag("MYTAG", "NOT BLOCKED : $countryshort");
        return 0;
    }
}

```

For the plugin to correctly work, the e-mail has to have come from an external network; in other words, there needs to be a public IP address that can be looked up in the geolocation database. In order for the plugin to run, a line of code must be included in SpamAssassin's file `init.pre`. This line tells SpamAssassin to load the plugin for use.

**Listing 5:** Loading plugin in `init.pre`

```
loadplugin Geolock Geolock.pm
```

## 4. USE CASE EXAMPLE

The user can test an e-mail manually by using the following command:

**Listing 6:** Testing an e-mail

```
#spamassassin -t < emailToTest.txt
```

In this example, the country of origin is US, which is being blocked. In the listing 7 you can see the resulting SpamAssassin header with the rule from the Geolock plugin being triggered. This results in the mail being awarded extra 50 points to the other triggered rules and extra header being added to the message with more detailed information.

**Listing 7:** Excerpt of added SpamAssassin header

X-Spam-Checker-Version: SpamAssassin 3.4.6 (2021-04-09) on mail.kovar.com		
X-Spam-Flag: YES		
X-Spam-Level: *****		
X-Spam-Status: Yes, score=61.9 required=5.0 tests=BLOCKED_COUNTRY, DATE_IN_FUTURE_06_12,FSL_HELO_FAKE,INVALID_DATE,MISSING_MIMEOLE,MSGID_OUTLOOK_INVALID,MSOE_MID_WRONG_CASE,RATWARE_MS_HASH,RDNS_NONE,TO_NO_BRKTS_MSFT,TRACKER_ID autolearn=no autolearn_force=no version=3.4.6		
X-Spam-Country: The country of origin is BLOCKED : US		
Content analysis details: (61.9 points, 5.0 required)		
pts	rule name	description
0.0	FSL_HELO_FAKE	No description available.
0.4	INVALID_DATE	Invalid Date: header (not RFC 2822)
0.0	DATE_IN_FUTURE_06_12	Date: is 6 to 12 hours after Received: date
3.9	MSGID_OUTLOOK_INVALID	Message-Id is fake
50	BLOCKED_COUNTRY	The country of origin is blocked
0.1	TRACKER_ID	BODY: Incorporates a tracking ID number
1.3	RDNS_NONE	Delivered to internal network by a host with no rDNS
3.4	MSOE_MID_WRONG_CASE	No description available.
1.8	MISSING_MIMEOLE	Message has X-MSMail-Priority, but no X-MimeOLE
1.0	RATWARE_MS_HASH	Bulk email fingerprint (msgid ms hash) found
0.0	TO_NO_BRKTS_MSFT	To: lacks brackets and supposed Microsoft tool

## 5. CONCLUSION

The SpamAssassin plugin *Geolock* for classifying e-mails based on the geographical location of the e-mail sender was developed. The process of finding the origin of an e-mail was mentioned. The implementation was described in detail, including presenting the overall functionality on a test e-mail. Further work should indicate if a finer definition of the blocked origin (city or area instead of a country) is needed. Another area of focus will be on thorough testing of the reliability of classification, the handling of specific cases, such as the transmission of e-mail within an internal network (no public IP address available), and re-thinking determining e-mail origin with regard to the possibility of spoofed *Received* headers, which can be done by the sender or malicious mail server along the path.

## REFERENCES

- [1] RAO, Justin M a David H REILEY. *The Economics of Spam*. Journal of Economic Perspectives [online]. 2012, 29. 8. 2020, 2012(Vol. 26 3), 87-110. ISSN 0895-3309. doi:10.1257/jep.26.3.87
- [2] WAKCHAURE, Sushma L, Shailaja D PAWAR, Ganesh D GHUGE a Bipin B SHINDE. *Overview of Anti-spam filtering Techniques*. International Research Journal of Engineering and Technology (IRJET) [online]. Jan. 2017(Volume: 04 Issue: 01), 429-434. ISSN 2395-0056. Available at: <https://www.irjet.net/archives/V4/i1/IRJET-V4I169.pdf>
- [3] CAHA, Tomáš a Oldřich KLÍMA. *Odhad pozice IP adresy s použitím dat z více geolokačních databází*. Elektrorevue [online]. Brno, 29. 2. 2020, 2020(Volume: 22 Issue: 1), ISSN 1213-1539. Available at: <http://elektrorevue.cz/cz/download/odhad-pozice-ip-adresy-s-pouzitim-dat-z-vice-geolokacnich-databazi/>
- [4] RESNICK, Peter W., 2008. *Internet Message Format*. 1. San Diego: RFC Internet Requests for Comments, 57 s. RFC, 5322. ISSN: 2070-1721. Available at: <https://tools.ietf.org/pdf/rfc5322.pdf>
- [5] KLENSIN, John C., 2008. *Simple Mail Transfer Protocol*. 1. Cambridge: RFC Internet Requests for Comments, 95 s. RFC, 5321. ISSN: 2070-1721. Available at: <https://tools.ietf.org/pdf/rfc5321.pdf>
- [6] *Apache SpamAssassin* [online]. Available at: <https://spamassassin.apache.org/>
- [7] *IP2Location IP Geolocation and IP2Proxy Databases*. IP2Location: IP Address to IP Location and Proxy Information [online]. Available at: <https://www.ip2location.com/database>

# Evaluation of broadband couplers for noise injection into powerlines

Ladislav Malček<sup>1</sup>, Petr Musil<sup>1</sup> and Petr Mlýnek<sup>1</sup>

<sup>1</sup>Department of Telecommunications, Brno University of Technology, Brno, Czech Republic

E-mail: [xmalce00@vutbr.cz](mailto:xmalce00@vutbr.cz), [xmutil56@vutbr.cz](mailto:xmutil56@vutbr.cz), [mlynek@vut.cz](mailto:mlynek@vut.cz)

**Abstract**—This paper deals with the problem of coupling noise into the transmission power line. The main goal is development of device which can effectively inject noise into transmission lines. This device will be used for testing industrial or commercial-used BPL (Broadband Powerline) modems and their ability to communicate under certain conditions. In the first part, there is an overview of coupler types. Next, the appropriate type was chosen according to required qualities. PCB (Printed Circuit Board) design of final device was developed on the principle of capacitive coupler. In the end, final PCB design was tested with several combinations of magnetic coils and capacitors.

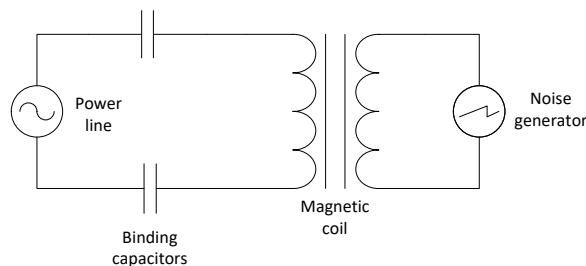
**Keywords**— BPL modems, capacitors, coils, coupler, noise, PCB design

## 1. INTRODUCTION

BPL (Broadband Powerline) modems provide solid and reliable communication without need of installing new cables between the modems. For evaluation of such communication, noise generator together with broadband coupler can inject signals into the powerline and therefore produce conditions which lead to decrease of SNR (signal to noise ratio) within the BPL channel. In this article, development and testing of coupler for injecting noise into transmission line for BPL modem communication is presented. In the first reference testing of bitrate and network throughput, the tests showed inconsistent conditions for the test during measurements with noise. A noise generator was used for simulating electromagnetic interference, which can occur in real networks. In original tests, the simple coil coupler was used for injection noise into transmission line, but this simple principle was very sensitive, and noise was injected only in certain frequency bands. Therefore, the main goal of this work is development of the coupler which can inject noise more effectively and evenly into the whole communication line. The proposed coupler could be used for evaluation of BPL modems in real conditions including tests focused on exploring limits of the technology.

## 2. PRINCIPLE

According to [1] several types of couplers exist based on the type of physical connection. The capacitive coupling is the serial insertion of one or more capacitors in the direct contact with the power cable. The main benefits of capacitive coupler are low insertion loss and high power transfer. Figure 1 shows a scheme of simple capacitive coupler. A coupling capacitor has a low impedance for high frequency signal which comes from the noise generator and at the same time it blocks the main signal with low frequency from power line. The inductive coupler is very easy to implement into the circuits, it offers degrees of freedom for system design by varying geometric shape of coils to tune the various parameters. The resistive coupler is used mainly for receiving signal from power line, so it is not suitable for construction of noise coupler. After considering the main features of each type of coupler, the capacitive coupler was finally chosen for noise coupler construction [2].



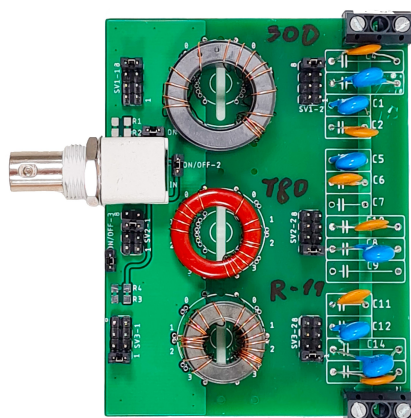
**Figure 1:** Scheme of simple coupler

The main part of this device is simple signal transformer made from magnetic core. In contrast with the initial solution for injection noise, coil and wiring is permanently fixed onto the PCB board. This design prevents parts from moving and changing position. Because the noise generator generates noise in band from 0 to 50 MHz, it is important to inject the whole frequency spectrum into communication line. Transformer with a certain combination of magnetic core and specific winding could create a filter for some frequency band. All these cases are implemented into coupler design. As it was previously mentioned, in the case of capacitive coupler the capacitor is the main connecting article, because it blocks low frequency from the power lines. This effect prevents that all the unnecessary signals from power line will not negatively affect noise generation [3].

### 3. PRINTED CIRCUIT BOARD

A development board was designed for testing basic coupler. It is basically a PCB board, on which three separable couplers can be attached. The PCB was designed by electronic design automation application Eagle. Figure 2 shows the finished development board. The number of turns on primary and secondary winding can be adjusted by jacks on each side of the coil. These adjustments could be made during testing; therefore, this system is flexible to examine the impact of different inductance of the coil. On a secondary winding of transformer are space for binding capacitors. Signal from the secondary winding is coupled into phase and neutral wire via the capacitors.

The final design of the coupler is based on design of development board, but it has a place for one coil instead of three. The combination of coil and capacitor was implemented into the final PCB.



**Figure 2:** Development board with three different coils

### 4. TESTING

A final testing of BPL noise coupler is done with industrial BPL modems which can communicate by phase, neutral and ground wire, however the profile of the modems was set to use only SISO (Single Input, Single Output) mode. This mode supports only the communication by phase and neutral wire (L-N). The design of the coupler allows injection of noise only to these wires.

The testing network must be isolated from the conventional power network as much as possible. To minimize interference from power network the 50 meters cable is used as power connection. It works as frequency filter, and it also helps to spread noise from network along its length. Also, the isolation transformer is used to further reinforce the signal insulation from the power grid.

The measurement focused on the noise level through all the spectrum on the lines. A specialized software which is supplied from BPL modems manufacturer, is used for the measurement. This program communicates with the modems and evaluates all the required parameters, such as noise, signal to noise ratio and others. This method is probably the best way to properly evaluate the noise level in the communication lines, because measurement is done with the units themselves.

Another measurement, which was used to test the efficiency of injection noise into communication line was measurement of bitrate and the amount of transferred data. For this purpose, two laptops were used. One was configured as a server, another one as a client, and they were connected by the modems to each other. Each measurement ran for 60 seconds, and it was done for all tested scenarios.

## 5. RESULTS

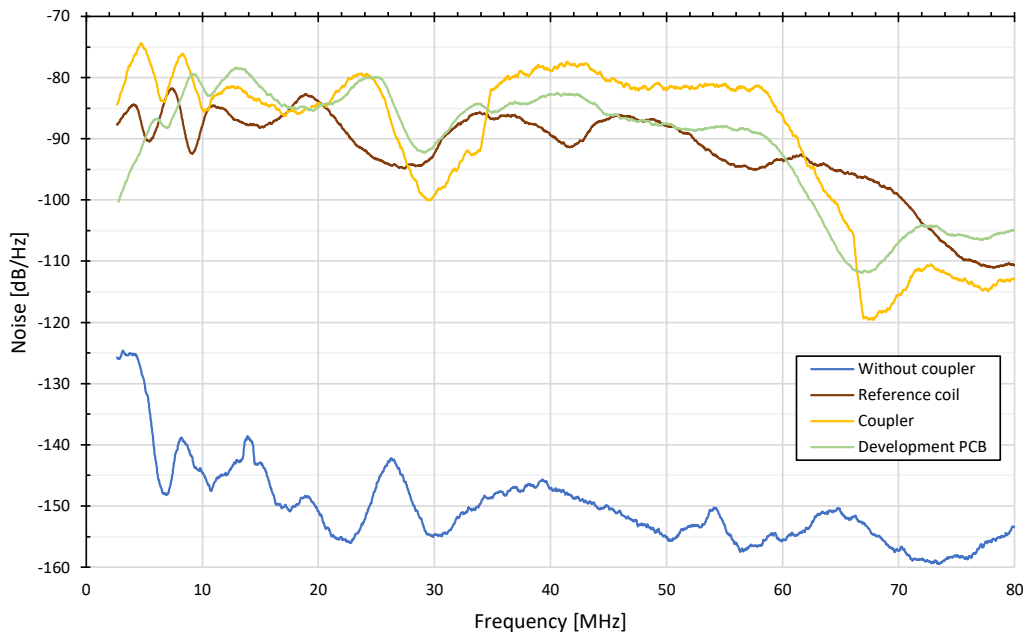
The final device as well as the development board were compared to reference measurement without any noise and with a simple coil coupler which was used earlier in the laboratory for BPL communication tests. Table I shows the results of data throughput of each measurement. For all the measurements was used only SISO mode of industrial BPL modems. This mode provides communication only by phase and neutral wire.

**Table I:** Results of the measurements of bitrate and transferred data.

Number	Coupler	Bandwidth	Transfer
		[Mbits/s]	[MBytes]
1	None	375	2630
2	Reference coil	105	752
3	Development board	66	495
5	Final coupler	132	946

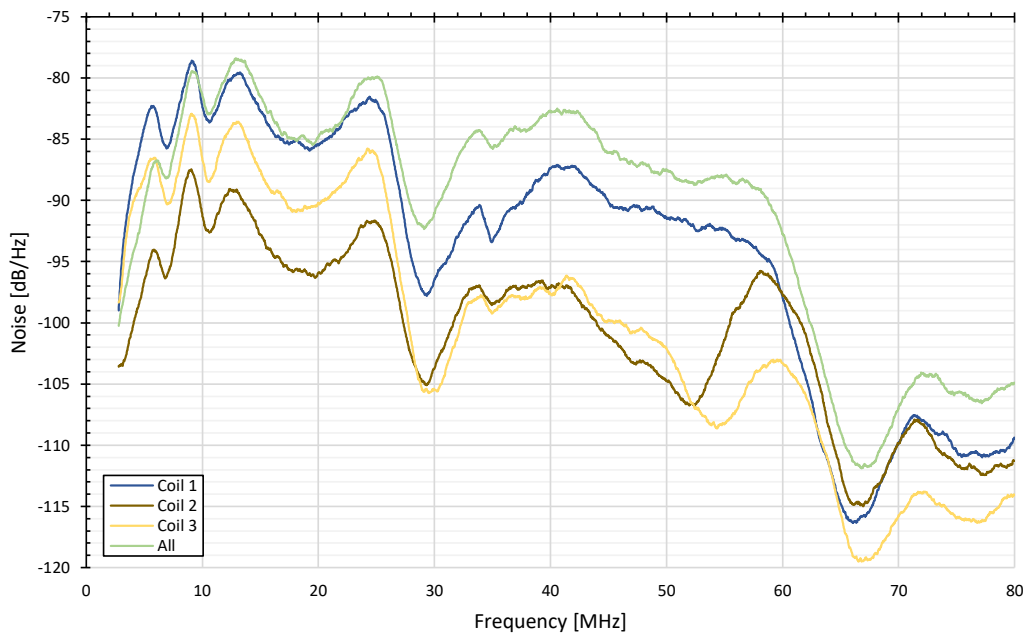
Figure 3 shows evaluation of power of noise in the measured spectrum. The graph of reference coil shows wavy course in whole measured spectrum. In contrast, the development board has straighter characteristic, but with the main drop at 30 MHz. The final coupler has the beginning of the course very similar to the reference coil and development board, but noise is much powerful in the area to roughly 10 MHz. The significant drop of signal appeared in the band from 25 MHz to 35 MHz, this drop is much more significant than in the previous case. The best part of the characteristic occurred in the band from 35 MHz to 60 MHz. The course of the graph is almost constant and noise has a level of -80 dB.

Figure 4 shows evaluation of each coil attached to the development board. Each of the coils has different parameters, such as inductance, frequency limit and core material. Graph of each characteristic has a very similar shape in band to 30 MHz but this different attenuation of noise. The coil number 2 (T80-2) with datasheet from [4] has highest attenuation of noise in whole frequency spectrum. The datasheet [5] of coil number 3 (ESD-R-19) claimed that frequency limit is 10 MHz, but the coil can inject frequency up to 25 MHz. The best characteristic has the coil number 1 (ESD-R-25D) with datasheet from [5], with the frequency limit of 300 MHz.



**Figure 3:** Comparison of noise characteristics of couplers





**Figure 4:** Comparison of coils on development PCB

## 6. CONCLUSION

This paper presented possibilities of usage of different types of couplers for injecting noise into power lines. These couplers will be used for testing purposes of BPL modems. The evaluation of the effect of three different coils is also shown in this paper. The coils are compared between each other and together as well. The best results in terms of lowest bitrate had development board with three different coils.

The results of testing different capacitance couplers showed very complex technical design of the coupler because the behavior depended on many parameters. The coils have the non-linear characteristic in terms of transforming the signals. Mainly in the lower frequency spectrum to approximately 25 MHz, coils reported similar courses of noise injection, but on higher frequencies the differences were more significant.

Very important element of coupler is binding capacitor which is directly connected to powerlines. The aim was to use capacitors with different values of capacitance for the maximum possible injection of noise. The addition of 22 nF binding capacitor caused a change in frequency characteristic of the coupler by enlarging the drop on 30 MHz, but improved binding on higher frequencies.

The further research that I am planning to conduct is going to be part of my master thesis.

## REFERENCES

- [1] L. Costa, C. Queiroz, B. Adebisi, V. Costa, and M. Ribeiro, "Coupling for Power Line Communication: A Survey", *Journal of Communication and Information Systems*, p. 16, 2017, [doi:10.14209/jcis.2017.2](https://doi.org/10.14209/jcis.2017.2).
- [2] A. Krutina and J. Bartovský, "PLC communication in Smart Grid", *Proceedings of the 2014 15th International Scientific Conference on Electric Power Engineering (EPE)*, pp. 165-168, 2014, [doi:10.1109/EPE.2014.6839518](https://doi.org/10.1109/EPE.2014.6839518).
- [3] M. Sibanda, P. van Rensburg, and H. Ferreira, "A Compact Economical PLC Band-Pass Coupler with Impedance Matching", *2013 IEEE 17th International Symposium on Power Line Communications and Its Applications*, pp. 339-344, 2013, [doi:10.1109/ISPLC.2013.6525874](https://doi.org/10.1109/ISPLC.2013.6525874).
- [4] Micrometals, "Power Conversion Line Filter Applications", 2007, [online datasheet](#).
- [5] KEMET Electronics Corporation, "ESD-R Toroidal Cores for Round Cables for Low High Frequency (Bare, coated with case)", 2018, [online datasheet](#).

# Hledáme na senior / junior pozice



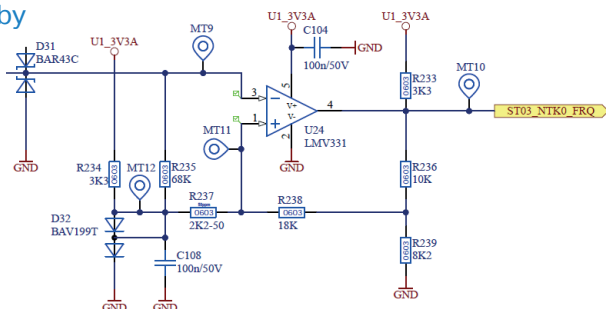
Division of Aerospace  
& Advanced Control

- System Engineer Aerospace
- Model Based Design Engineer (Matlab, dSPACE)
- Test Engineer
- Embedded HW Engineer
- Embedded SW Engineer



## Na čem budeš pracovat

- Vývoj řídicích systémů (ECU, FADEC) pro letecké motory
- Vývoj řídicích systémů pro palivové a hydraulická aktuátory
- Vývoj výkonové elektroniky (AC/DC, DC/DC měniče)
- Verifikace řídicích systémů v provozu
- Certifikace v souladu s leteckou legislativou
- Zavedení do sériové výroby

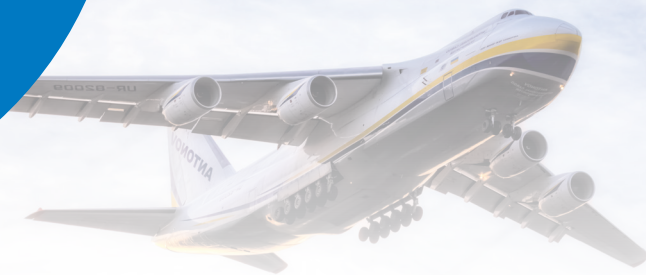


## Kontakt:

✉ [kariera@unis.cz](mailto:kariera@unis.cz)



[www.unis-daac.cz/cs/kariera](http://www.unis-daac.cz/cs/kariera)



## UNIS, a. s.

Hlavní činností společnosti **UNIS, a.s.** je realizace projektů v investiční výstavbě zaměřené do oborů zpracování ropy a zemního plynu, petrochemie, chemie a energetiky v dodavatelské formě tzv. „na klíč“. K ostatním aktivitám společnosti **UNIS** patří – Zákaznický i vlastní vývoj a výroba SW i HW v oblastech leteckého, mikroprocesorového a automobilového průmyslu. „Výrobní informační systém MES“ (Manufacturing Execution System), vyvíjený a distribuovaný pod značkou **PHARIS®**, je nasazován pro komplexní řízení výroby v diskrétních výrobcích. Zkušební laboratoře **VTP UNIS** nabízí zkoušky elektromagnetické kompatibility **EMC**, zkoušky mechanické a klimatické odolnosti.

### Nabízíme

- Zajímavou a dlouhodobou práci na vývojovém projektu
- **Kontinuální profesní a mzdový růst**
- Pružnou pracovní dobu a systém odměňování
- **1 týden dovolené navíc**
- Stravování v místě pracoviště

Místo výkonu práce:  
**Brno**

### Hledáme

- Analytické a komunikační schopnosti
- **Schopnost formulovat technické požadavky na zařízení**
- Základní inženýrské znalosti z oboru elektro a programování
- **Dobrou znalost oborové technické angličtiny**
- Aktivní a inovativní přístup

Typ nabídky:  
**HPP, brigáda, stáž, trainee program, diplomová práce, ostatní**

## Kontakt:

✉ [kariera@unis.cz](mailto:kariera@unis.cz)



[www.unis-daac.cz/cs/kariera](http://www.unis-daac.cz/cs/kariera)

# An Implementation of Lattice-based Proof-of-Work on Blockchain

A. Krivulčík<sup>1</sup> and S. Ricci<sup>1</sup>

<sup>1</sup>Department of Telecommunications, Brno University of Technology, Technická 12, Brno, 616 00, Czech Republic

E-mail: [xkrivu00@vut.cz](mailto:xkrivu00@vut.cz), [ricci@vut.cz](mailto:ricci@vut.cz)

**Abstract**—Cryptocurrencies and blockchain are skyrocketing in recent years. They rely on Proof-of-Work (PoW) mechanisms for generating a new transaction and turn this process into "work" (i.e., puzzles) where miners are paid for. With the advent of quantum computers, also PoW starts to migrate to post-quantum cryptographic alternatives. To the best of our knowledge, we present the first implementation of a lattice-based PoW based on the Shortest Vector Problem (SVP). By implementing in Python 3 and with the use of the NumPy library, we wrote a software that uses this concept on an artificial blockchain and demonstrates its real-world use. Even if this proposal has drawbacks on GPU optimisation and storage requirements, it shows its potential in use. The experimental results show that by balancing the size and generated range of a given matrix and vector, we can easily manipulate the time required to solve SVP challenge.

**Keywords**—Proof-of-Work, Lattice-based Cryptography, Blockchain, SVP problem, Post-quantum Cryptography, Python

## 1. INTRODUCTION

Cryptocurrency usage in the global market has grown significantly over the past decade. MarketWatch [1] estimates that by 2030, cryptocurrency usage will have an annual growth rate of 30%. One of cryptocurrency security primitives is the PoW and it is used to provide a certain amount of computational effort that has been spent. A PoW consists of solving a computational puzzle. The puzzle is very hard to compute but its correctness is easy to verify. PoW that is used in Bitcoin, Ethereum, and many other cryptocurrencies, consists of computing the hash function of a given block with the addition of a random nonce to achieve a hash with a required structure. Anyone can read and verify the PoW, but the data cannot be changed or tampered with after it is written to the blockchain. Hash functions such as SHA-256 for Bitcoin and Keccak-256 for Ethereum are proven to be secure and reliable.

However, a quantum era has been unfolding and Grover's search algorithm simplifies the collision and symmetric key brute force search to sub-linear complexity with an increase of keys and parameters sizes [2]. In this context, cryptocurrencies also start to migrate to post-quantum algorithms and PoW can be based on post-quantum problems such as the SVP problem on a lattice. Lattice can be imaged as a finite number of vectors in a given dimension. The problem called SVP is about finding the shortest vector in a given lattice, or matrix, so that its Euclidean norm (in other words length) is the shortest.

The lattice vectors are randomly generated with very big numbers, thus making computations harder. Therefore, the PoW hash function challenge can be replaced with an SVP challenge on a randomly generated lattice. To the best of our knowledge, we present the first implementation of the algorithm proposed in [3]. Moreover, experimental results on the lattice-based PoW efficiency and a comparison with currently deployed Bitcoin PoW are shown.

## 2. IMPLEMENTATION

The implementation is made in the Python 3 programming language and is publicly available in the repository<sup>1</sup>. A small change in the NumPy library is made to make computations over such high numbers as  $1.5 \cdot 10^{540}$  possible. In particular, the Python object needed to be replaced with an object from library Decimal.

Algorithm 1 depicts the steps needed to run the PoW. In particular, the first step is to generate a random prime of a given size. By utilising native Python library Crypto.Util and its function `getPrime(n)`, we

---

<sup>1</sup><https://github.com/Mysfrit/Lattice-PoW>

---

**Algorithm 1** PoW challenge structure

---

```
-  $n \leftarrow 150$ 
-  $\alpha \leftarrow$  Target magnitude from already solved vectors (SVP challenge).
-  $Gen(p, B, \alpha)$ , where  $p$  is prime of length  $10n$ , create Base matrix  $B$  with uniform distribution  $U(0 \in [p - 1])$ .
Return  $c = (\alpha, n, B, p)$ 
-  $Solve(c)$ 
while  $||V|| \leq \alpha * p^{1/n}$  do
    - Generate a non-zero vector  $v$  of size  $n$ .
    -  $V = B * v$ 
    - Calculate norm  $||V||$ 
end while
- Return  $(c, v, V, \alpha, B)$ , Write values to blockchain
-  $Verify(c, v, V, \alpha, B)$ 
```

---

receive a random prime number of bit size  $n$ . The second step is to create a matrix basis  $B$ . For this step, we are using the library NumPy. First with function `numpy.zeros(n)`, we create matrix of size  $n \times n$  with zeros. After that, we add ones on diagonals. Again by utilising library NumPy, we generate uniformly distributed random values from range  $(0, p - 1)$  and add them to the first row of the matrix. Then we add the generated prime on the first row and first column.

The third and last step is to start generating random vectors. For this, we are again using NumPy function `numpy.random.randint(low, high, n)`, where `low` and `high` are thresholds for values in the vector and  $n$  is the size of a given vector. Matrix multiplication with the generated vector and matrix basis is applied and the magnitude of the output vector is calculated. The final step is to compare the magnitude with a required value and loop this process until we find a satisfactory vector<sup>2</sup>.

### 2.1. Speed and memory usage

In this section, we introduce the speed optimisation achieved and compare the memory usage of the proposed PoW with Bitcoins PoW. In fact, our code is optimised to do so with parallel computing, utilising all possible CPU resources. For the time being, the biggest bottleneck is generating random numbers for the vector. With the benchmark CPU, we could achieve around 2500 guessed vectors per second (i.e., 25000 generated numbers per second with  $n = 100$ ).

Bitcoin block size is around 1 megabyte including transactions and other variables. At first, it appears to be manageable, but if this network is scaled into five years, every single kilobyte can make a difference. Referring to the paper [3], there are several variables:

- $n$  - recommended around 150
- prime  $p$  - Recommended size is  $10n$  bits
- Base matrix  $B$  - consisting of prime  $p$  and uniformly generated numbers  $x_i \in U(0 - [p - 1])$ , with size  $n \times n$

By taking these variables into account, the required storage space for a verification pass from a few bytes (in Bitcoins) to around tens or hundreds of kilobytes. Therefore, just a verification part of one block with the matrix of size  $150 \times 150$  would take 99.3 KBytes. For example, a Bitcoin blockchain of size 379 GB with 725000 blocks would be around 48% bigger with the use of a lattice PoW. Of course, if this PoW would be used on a real network, the parameters would need to be even bigger to compensate for larger computational power. That could lead to a significant increase in size.

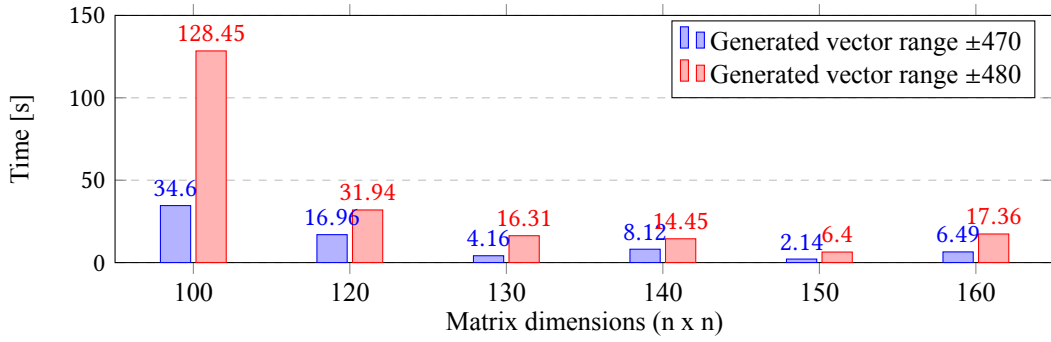
## 3. EXPERIMENTAL RESULTS

The complexity of computation, i.e., finding a vector with required properties, is linked to the range of a generated vector. For example, taking into account a matrix basis of size  $120 \times 120$  with range of generated vector  $\langle -470, 470 \rangle$ , the challenge is solved in around 16.96 seconds<sup>3</sup>, with around 2500 guessed vectors per second. But a vector with a range of  $\langle -480, 480 \rangle$  needs on average 2.2 times longer. These values are plotted in Figure 1. As with Bitcoin, we can artificially increase the difficulty of finding the solution

---

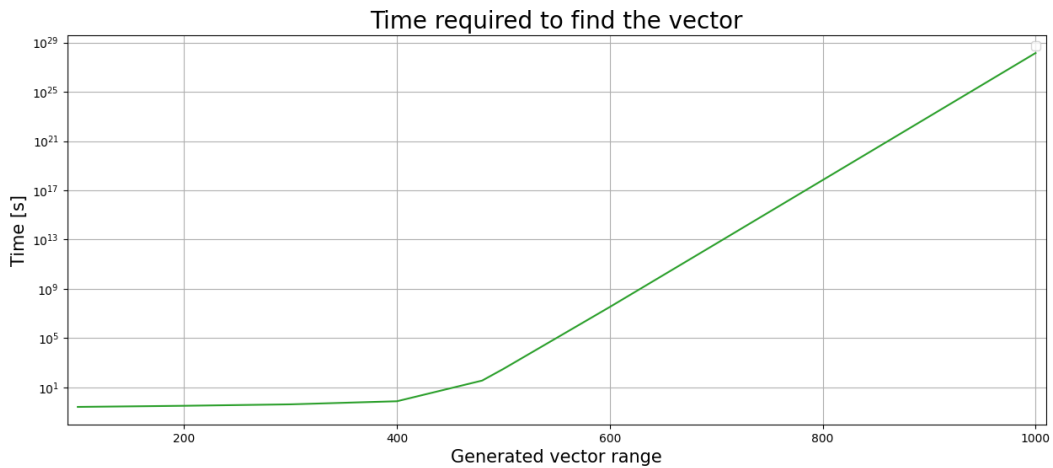
<sup>2</sup><https://www.latticechallenge.org/svp-challenge/halloffhp>

<sup>3</sup>Calculated on a CPU Intel(R) Core(TM) i5-8400 CPU @ 2.80GHz



**Figure 1:** Comparison of time taken to find a vector with sufficient norm and different vector ranges.

to slow down the network computations. With this knowledge, we can estimate the behaviour of the network. Note that increasing the size  $n$  of a given basis and vector does not necessarily increase the computational time, so it should be stable at 150. For instance, if the network uses 10-minute intervals as in Bitcoin, the network would need to adjust according to the time required to solve the latest block and adjust the vector generation range accordingly. However, it has be noted that finding a vector is purely a game of chance. There could be a faster or slower solution for the same vector at two different times.



**Figure 2:** Time required to find the vector.

Figure 2 shows our challenge solving times for vectors ranges smaller than  $\langle -500, 500 \rangle$  and our estimation from this threshold onward. In particular, comparing our results to the SVP challenge website<sup>4</sup>, we estimated that finding a solution of a targeted generated vector of range  $\langle -1000, 1000 \rangle$  would take on average  $4 * 10^{20}$  years. To put this time into perspective, the universe is old  $1.6 * 10^{10}$  years. At first glance, this number should be non-achievable for sole individuals, but we need to remember that the network may consist of a hundred thousand users trying to solve a given problem. Furthermore, these vectors are easier to solve on faster computers. For example, a vector of size 180 with generated range can be solved on 2 Xeon processors with 4 Nvidia GPU cards in 28 days by a single user.

#### 4. CONCLUSION

Cryptocurrencies are being used more frequently every year. With the advent of quantum computers, their security and cryptographic primitives will start to migrate to post-quantum secure algorithms. To the best of our knowledge, we present the first implementation of the lattice-based PoW proposed in [3]. Moreover, experimental results on the lattice-based PoW efficiency and a comparison with currently deployed Bitcoin PoW are carried out.

<sup>4</sup><https://www.latticechallenge.org/svp-challenge/halloffame.php>



As a result of our analysis, we conclude that solving the SVP problem depends more on the range of generated guessed vector rather than on the dimension of the lattice matrix. Moreover, we estimate that the computational time to get a resulting vector will need to be more than double by increasing the vector size by a factor of 10. Even if a real-life implementation of this proposal can lead to blockchain resistance to quantum computer attacks, we predict that the higher memory consumption of the whole blockchain due to storing such large numbers will become a drawback that needs to be resolved.

Furthermore, there is a possibility of deadlock when a generated lattice does not have a vector that meets the required magnitude. In this case, the network should regenerate a new matrix and restart the process after some arbitrary amount of time.

As future work, we plan to speed up the matrix generation running the computation on a graphic card from Nvidia, where guesses could achieve values over 200 million guessed numbers per second (in our case 2 million vectors per second)<sup>5</sup>.

## REFERENCES

- [1] Cryptocurrency Market Size, Growth 2022 Global Opportunities, Trends, Regional Overview, Global Leading Company Analysis, And Key Country Forecast to 2030. Market Watch [online]. 2021, December 22 2021 [cit. 2022-03-06]. Available from URL: <https://on.mktw.net/3hU5nz0>
- [2] BERNSTEIN, Daniel J. Introduction to post-quantum cryptography. Post-quantum cryptography. Springer, Berlin, Heidelberg, 2009. 1-14. Available from URL: [doi:10.1007/978-3-540-88702-7\\_1](https://doi.org/10.1007/978-3-540-88702-7_1)
- [3] BEHNIA, Rouzbeh, Eamonn W. POSTLETHWAITE, Muslum Ozgur OZMEN and Attila Altay YAVUZ. Lattice-Based Proof-of-Work for Post-Quantum Blockchains. Data Privacy Management, Cryptocurrencies and Blockchain Technology [online]. Cham: Springer International Publishing, 2022, 2022-01-01, 310-318 [cit. 2022-03-01]. Lecture Notes in Computer Science. ISBN 978-3-030-93943-4. Available from URL: [doi:10.1007/978-3-030-93944-1\\_21](https://doi.org/10.1007/978-3-030-93944-1_21)
- [4] CHEN, Lily, Stephen JORDAN, Yi-Kai LIU, Dustin MOODY, Rene PERALTA, Ray PERLNER and Daniel SMITH-TONE. Report on Post-Quantum Cryptography [online]. 2016, 1–3 [cit. 2021-11-17]. Available from URL: <http://dx.doi.org/10.6028/NIST.IR.8105>

---

<sup>5</sup><https://developer.nvidia.com/curand>

# Paillier Cryptosystem Optimisations for Homomorphic Computation

P. Ryšavá<sup>1</sup> and S. Ricci<sup>1</sup>

<sup>1</sup>Brno University of Technology, Technická 12, Brno, 616 00, Czech Republic

E-mail: [xrysav29@vutbr.cz](mailto:xrysav29@vutbr.cz), [ricci@vut.cz](mailto:ricci@vut.cz)

**Abstract**—Homomorphic encryptions can ensure privacy in systems operating with sensitive data. It also allows outsourcing the data processing without the need to disclose the information within. To keep good performance over the growing mass of data, the execution of homomorphic schemes has to be efficient. In this article, we focus on the optimization of the Paillier scheme. This scheme allows the addition of a constant or another ciphertext without decryption of the encrypted values. Since the exponentiation used in the encryption process is time-consuming, we have implemented noise and message pre-computation to avoid time-demanding operations. These adjustments significantly fasten the encryption process, especially using the noise pre-computation.

**Keywords**— Homomorphic Encryption, Paillier Cryptosystem, Secret Sharing

## 1. INTRODUCTION

The Paillier cryptosystem [1] is a probabilistic asymmetric algorithm for public-key cryptography. This scheme has an additive homomorphic property, i.e., allows adding two ciphertexts without corrupting the result. The computation behaves just as if the corresponding plaintexts are added (without needing decryption). Because of this ability, the scheme became very popular, especially for e-voting [3], machine learning on encrypted data [4], and other algorithms that demand high confidentiality even during processing between mutually authenticated parties.

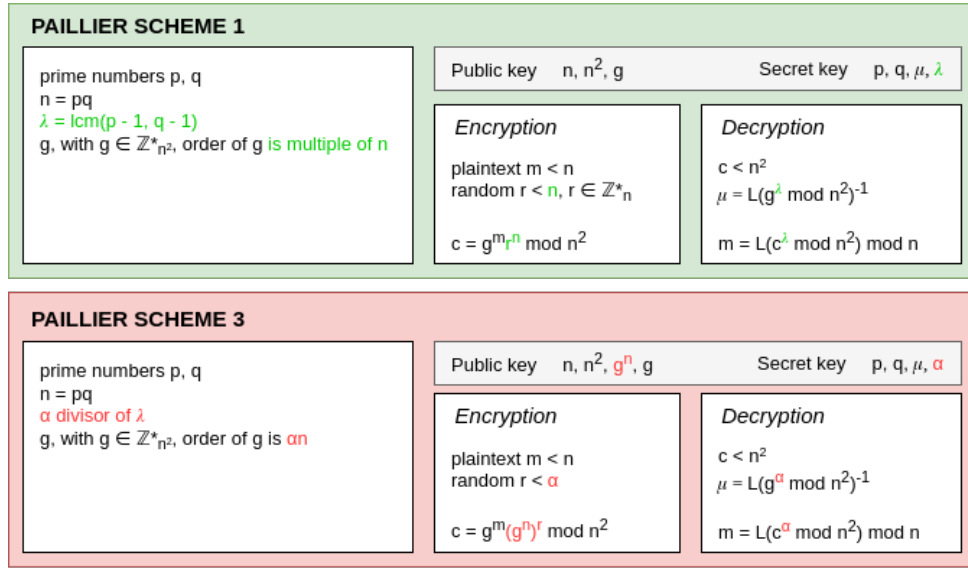
However, a drawback of the Paillier cryptosystem is its efficiency which is a common issue for any scheme based on homomorphic encryption. This article focuses on the Paillier scheme acceleration. We are especially interested in the encryption process. In fact, the encryption phase needs to be applied more times to do computation on encrypted data. Several optimisation proposition has already been made. For instance, Paillier proposed a variant, namely *Scheme 3*, of the original proposal, namely *Scheme 1*, which uses a bound  $\alpha$  on exponentiation power size to speed up the computation. Other improving method can be the use of pre-computation of some values, particularly the exponentiation of either the message  $g^m$  or the noise  $r^n$  as suggested in the article [2]. See Section 2 for more details.

In this work, we implemented *Scheme 1* and *Scheme 3*, and we optimised them with message and noise exponentiation pre-computation. Moreover, we considered the pre-computation of (1) the modular inverse  $\mu$  used during the decryption process, and (2)  $n^2$  and  $(g^n)^r$  used in *Scheme 3*. In particular, we develop an application in the *Visual Studio Code* editor that uses the C/C++ Extension Pack. The back-end layer runs in C programming language. At last, we compared the optimizations through experimental results.

## 2. IMPLEMENTATION

Our implementation uses the C programming language extended with the *OpenSSL* and *cJSON* package. The *OpenSSL* library implements the *BIGNUMBERS* library, which allows computation with larger numbers than the size of the unsigned long long variable type. The library also implements many cryptographic functions, such as the Digital Signature Algorithm (DSA) parameters generation, used in *Scheme 3* during the generation phase. *Scheme 1* and *Scheme 3* algorithms are sketched in Figure 1. Moreover, the *cJSON* library implements procedures simplifying the operations over the JavaScript Object Notation (JSON) files. In the implemented application, JSON files store the pre-computed values and corresponding keys since the JSON structure is simple and allows fast search.

The goal of our implementation is to speed up the cryptosystem procedure. We use the pre-computation of (1) message exponentiation ( $g^m$ ), (2) noise exponentiation where  $r^n$  is computed as  $(g^n)^r$ , and (3) values  $n^2$  and  $\mu$ . Pre-computed message and noise values are stored in a JSON file whereas  $n^2$  and  $\mu$  values are stored directly in the keychain structure since the latter values remain the

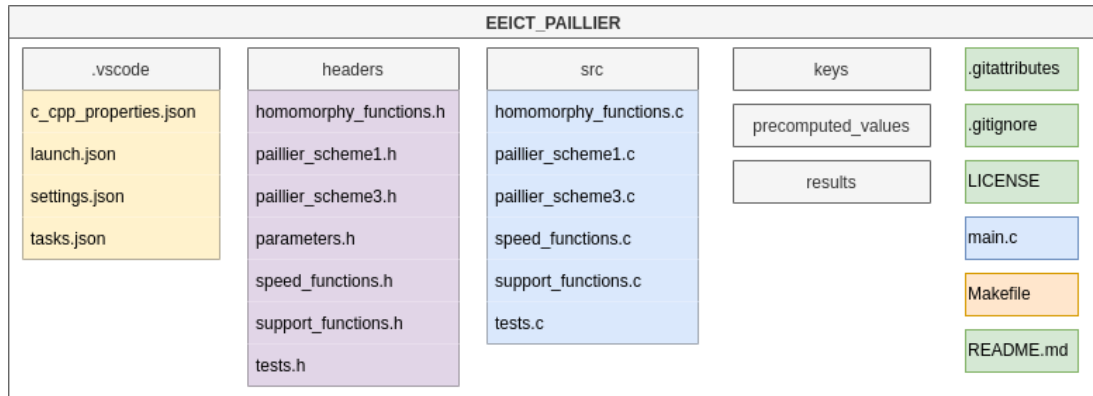


**Figure 1:** Paillier Scheme 1 and Scheme 3 description with highlighted differences in green and red colours.

same for the given keychain. For the purposes of Scheme 3, the  $g^n$  parameter is also precomputed and stored directly in the keychain structure.

## 2.1. Application

In the application development, the *Visual Studio Code* editor (v1.65.0) with the C/C++ Extension Pack (v1.1.0) was used. The directory tree of the implemented application is depicted on Figure 2. When started, the application opens with loading the saved keys from JSON files with prefix *saved\_keys* located in the *keys* sub-directory. If any of the files are missing, a new corresponding keychain is generated and stored for future purposes. After the key-load phase, the algorithm checks if the file containing the pre-computed values exists. If not, the algorithm runs the pre-computation operation with parameters set in the parameters header file. Once the pre-computation is finished, the algorithm waits for the user to continue into a console menu.



**Figure 2:** Directory tree of the application.

The application menu contains two areas named *RUN* and *TESTS*. Inside these sub-menus are listed all available functions. The ran operation is decided by inserting a number in the console. Each implemented scheme can be run either in the plain version, with noise, message pre-computation, or with both message and noise pre-computations. All of the above-stated "modes" can also be launched together. Gained results from the computations are stored in Comma Separated Value (CSV) format in the results sub-directory.

It is important to notice that the amount of the precomputed values can be defined in the parameters header file by setting the parameter *RANGE*. This parameter states that the pre-computation will be applied on values from 1 to the agreed number. The parameter *RANGE* is set by default to 100. With this setting,

the size of files with the pre-computed values is 380.9 kB. When the parameter is equal to 1000, the file-size is about 259.1 kB. The size of the files with the saved keys is 8.4 kB big.

The implemented application is publicly available on GitHub repository<sup>1</sup>.

Scheme 1				Scheme 3			
	Encryption	Decryption	Total		Encryption	Decryption	Total
Plain	2.172961	1.915783	4.088744	Plain	2.530260	2.387416	4.917676
Random	0.036213	2.084682	2.120895	Random	0.096640	2.514155	2.610795
Message	2.137328	1.913080	4.050408	Message	2.414621	2.359763	4.774384
Both	0.008832	2.060453	2.069285	Both	0.025454	2.521757	2.547211

**Table I:** Average time of encryption, decryption and total time in *ms* for 2048-bit length of  $n$ . "Plain" states for in plain mode, "random" for pre-computed noise values, "message" for pre-computed message values, and "both" for both pre-computations.

### 3. EXPERIMENTAL RESULTS

In this section, we compare the efficiency of plain schemes and their optimizations. For the experiments, we used an Intel(R) Core(TM) i7-4510U CPU at 2.00 GHz with 4 cores and Ubuntu 64-bit v20.04.4 LTS operating system. The compiler version is GCC v9.3.0. The experimental results were gained for a DSA modulus  $n$  of 2048-bit long. Encryption results of one message were averaged over 100 iterations and 10 different messages values were considered. In our case, the messages are integers since we are planning to use its homomorphic property as future work.

Table I depicts *Scheme 1* and *Scheme 2* performances averaged over 10 messages, respectively. Based on the gained results, the decryption speed is for all variants comparable. This is due to fact that the decryption operations are limitingly using the pre-computation optimisation. Regarding the encryption, *Scheme 1* is slightly quicker than *Scheme 3* despite expectations. This probably occurs because of the smaller size of the exponentiation base  $r$  in *Scheme 1* with respect to  $g^n$  in *Scheme 3*. The results were obtained with parameter RANGE sets to 100 where the pre-computed message file has size 64.9 kB for both schemes and the pre-computed noise files have size 64.3 and 125.9 kB for *Scheme 1* and *Scheme 3*, respectively.

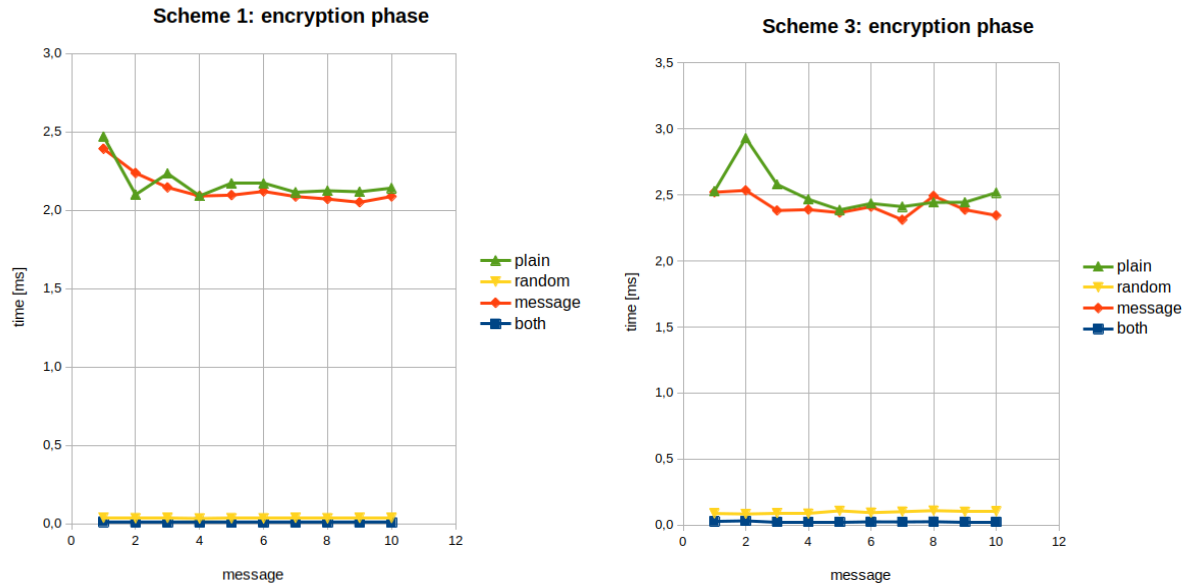
Figure 3 depicts *Scheme 1* and *Scheme 3* computational cost of one message encryption averaged over 100 times iteration of the considered protocol. From the analyses, we can deduce that the most computationally intensive operation is the noise computation. This is the main reason of the drastic speedup when the noise pre-computation is applied to the process.

Optimisation technique	Scheme 1 saved time [ <i>ms</i> ]	Scheme 3 saved time [ <i>ms</i> ]
Noise pre-computation	3.286700	3.660420
$\mu$ pre-computation	2.544364	3.100727
Optimisation technique	Schemes 1 and 3 saved time [ <i>ms</i> ]	
Message pre-computation	0.039340	
$n^2$ pre-computation	0.039340	
$g^n$ pre-computation	13.568727	

**Table II:** Average time in *ms* saved during each optimisation of the considered Paillier schemes.

Table II contains the saved times in average over 10 protocol executions of the considered pre-computations technique. Note that  $\mu$  and  $n^2$  acceleration allows saving time during the whole process since they are applied multiple times in the protocols.

<sup>1</sup> [https://github.com/Norted/EEICT\\_Paillier](https://github.com/Norted/EEICT_Paillier)



**Figure 3:** Encryption overtime of both *Schemes 1 and 3* with their optimisations. "Plain" states for in plain mode, "random" for pre-computed noise values, "message" for pre-computed message values, and "both" for both pre-computations.

## 4. CONCLUSION

Paillier is a homomorphic asymmetric cryptographic scheme that allows addition operations over encrypted values without corrupting the result. However, the scheme presents efficiency drawbacks and its performance needs to be perfected.

In this article, we develop an application in the *Visual Studio Code* editor with C programming language. Our application implements and compares several optimizations of the Paillier scheme. We mainly focused on the encryption process since it is normally applied more times in homomorphic computations, whereas the decryption phase is applied only once. In particular, we considered pre-computation of (1) noise exponentiation and (2) message exponentiation. Moreover, we implemented  $n^2$  and  $\mu$  pre-computations in all protocols.

The experimental results shows that the most time-saving operation is to use the noise pre-computation. Such operation speeds the encryption process significantly. As future work, we plan to apply the Paillier scheme with noise exponentiation pre-computation in our authentication key agreement protocol where it will be used during the sharing-phase in order to maintain the secrecy of the shares.

## REFERENCES

- [1] Paillier P. (1999). *Public-Key Cryptosystems Based on Composite Degree Residuosity Classes*. In: Stern J. (eds) *Advances in Cryptology — EUROCRYPT '99*. EUROCRYPT 1999. Lecture Notes in Computer Science, vol 1592. Springer, Berlin, Heidelberg. DOI: [https://doi.org/10.1007/3-540-48910-X\\_16](https://doi.org/10.1007/3-540-48910-X_16). [Online; accessed 08-March-2022]
- [2] Jost C., Lam H., Maximov A., and Smeets B. (2015). *Encryption Performance Improvements of the Paillier Cryptosystem*. Cryptology ePrint Archive, Report 2015/864. <https://ia.cr/2015/864>. [Online; accessed 08-March-2022]
- [3] Anggriane S. M., Nasution S. M. and Azmi F. (2016). *Advanced e-voting system using Paillier homomorphic encryption algorithm*. International Conference on Informatics and Computing (ICIC), pp. 338-342. DOI: [10.1109/IAC.2016.7905741](https://doi.org/10.1109/IAC.2016.7905741). [Online; accessed 08-March-2022]
- [4] Bost R. and, Popa R. A., Tu S., and Goldwasser S. (2014). *Machine Learning Classification over Encrypted Data*. Cryptology ePrint Archive, Report 2014/331. DOI: <https://dx.doi.org/10.14722/ndss.2015.23241>. <https://ia.cr/2014/331>. [Online; accessed 08-March-2022]

# PTDOS: Tool to support DoS testing

Filip Kamenář<sup>1</sup> and Zdeněk Martinásek<sup>1</sup>

<sup>1</sup>Brno University of Technology, FEEC, Czechia

E-mail: [xkamen19@vut.cz](mailto:xkamen19@vut.cz), [martinasek@vut.cz](mailto:martinasek@vut.cz)

**Abstract**—Nowadays, cyber attacks are a common part of our lives. Targets of attacks are various, such as obtaining sensitive data, network infiltration, system sabotage or denial of service (DoS). This paper deals with the attacks that target denial of service. The main goal is the custom design and development of the ptdos application, which allows testing of the resilience of network infrastructure and web servers against these types of cyber attacks. The application will be freely available and will contain various types of attacks (both flood and logic) to sufficiently test the system. The purpose of the work is to create a quality and simple tool to enhance cyber security.

**Keywords**—Denial of Service attacks, cybersecurity, Python, penterep

## 1. INTRODUCTION

In the modern world of information technology, web applications are becoming increasingly popular, where the user does not need to keep data permanently on his computer, but acquires it from web servers according to current needs. A typical example is video content services, where content is streamed and played directly to the user instead of having first download it to user's storage and then play the video content. For these reasons, great emphasis is placed on the quality of the provided service and in particular, on its availability. It is the availability of services that DoS attacks are focused on, whose main objective is to deny service to the user who wants to consume the content [1]. At that point, the web server providing the service is the target of the attack and the user becomes the victim of the attack. In the beginning, it is important to understand what DoS attacks actually are, how they work, and what types of DoS attacks exist. Any accessible device on the network can be the target of a DoS attack. They are able to operate on any OSI layer [2]. DoS attacks attempt to make the victim's communications with other devices or users inaccessible. This can be achieved by different methods depending on the type of chosen attack. An extended variant of DoS attacks can be distributed DoS attacks, called DDoS. Within DDoS, the attacker acts as a group of mutually cooperating machines, which intensifies the attack. The attacking group of machines is called a botnet.

### 1.1. State of the Art

In order to ensure that provider services continue to operate flawlessly even in the case of various types of DoS attacks, providers undergo regular penetration tests where penetration testers use a simulated attack to identify weaknesses in their solution and then fix them. The penetration tester must be able to effectively perform different types of attacks and thereafter evaluate the impact on the target system. There are currently a plethora of tools available on the market that are capable of conducting DoS attacks (e.g., hping3, trafgen or Apache JMeter). However, a common disadvantage of most existing tools is they focus on only one specific type of attack, or only a certain group of attacks. This can make these existing tools quite inefficient because the penetration tester using these tools must combine different tools to test multiple types of attacks. Another disadvantage of these existing tools is the lack of feedback on the progress of the attack. While the application is able to perform a given attack and monitor its output, e.g. how many data units were sent, it does not provide the penetration tester with any information about how the attack affected the functionality of the target system, which is critical information for evaluating the target system's vulnerability to a given type of attack.

### 1.2. Our contribution

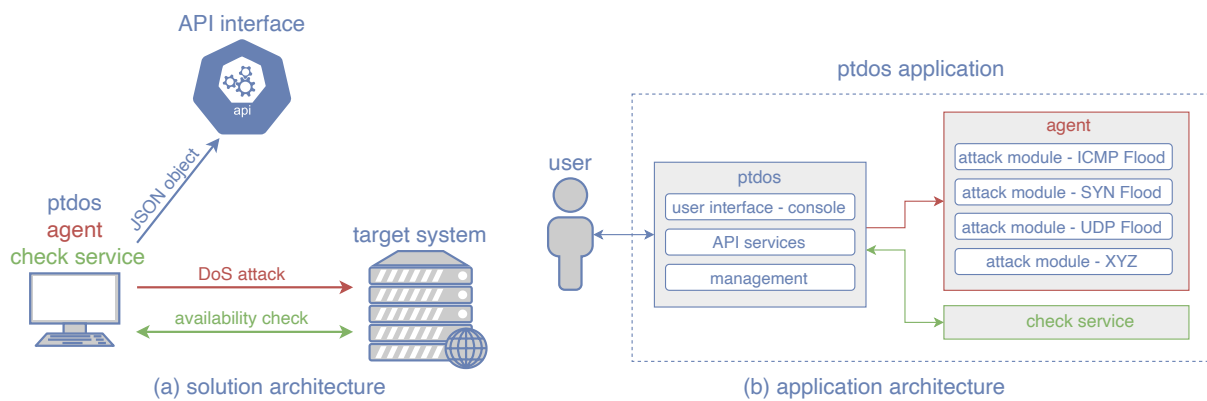
The main contribution of the paper is the design and development of the ptdos application, which enables comprehensive testing of the resilience of network infrastructure and web servers against cyber attacks such as DoS. The application will be freely available and will include various types of attacks (both flood and logic) to sufficiently test the target system. Ptdos will be able to conduct DoS attacks in real time



and at the same time will evaluate the status of target systems. The methodology and selected attacks are chosen on the basis of OWASP [3] recommendations and on the basis of practical experience of penetration testers. Thus, the aim of this work is to create a quality and simple tool to enhance cyber security. The ptdos application was designed and developed as a console application based on the Python programming language. During the development of the application, an innovative and modern plugin (module) architecture was implemented, which ensures easy extensibility of the application in the future. Individual attacks are created as separate plugins, i.e. when creating a new attack plugin it is not necessary to interfere with the original application code. It is only necessary to modify the configuration file, which contains a list of plugins that the application loads at startup. Thanks to this approach, the application is ready for future expansion with new types of DoS attacks. For better usability in the real environment, the application is ready for integration into other systems using JSON objects and passing them to the external API interface. In its current state, the application is ready for integration into the penterep system, which is a comprehensive system used to implement and manage security tests for commercial entities.

## 2. OUR APPLICATION DESIGN AND DEVELOPMENT

The architecture of the application is divided into 3 logical units: ptdos, check service agent. It is a logical division due to the fact that the implemented architecture works in a combined mode, i.e. all 3 parts are operated on one machine, so all logical units are part of one application. The scheme of the solution architecture is shown in section a), Fig. 1. The internal architecture of the application is described in section b), on the same Figure, where the basic sub parts of each logical unit are listed.



**Figure 1:** Solution architecture and application architecture of ptdos.

The purpose of the ptdos part is interacting with the user and controlling all remaining parts of the application. The ptdos includes a console communication interface that processes input data provided by the penetration tester, starts the check service, and sends a command to start a DoS attack to the agent. This is where the power of the implemented plugin architecture is the most evident, because the agent that performs the DoS attack, may contain an unlimited amount of attack modules. The main feature is to implement a new attack module, it is not necessary to modify the original application code, but a configuration file is used which contains information about all available modules. In the control part of ptdos, all available modules are loaded when the application is started. The input data of the ptdos application are, for example, the type of attack, the duration of the attack, the IP/URL address and port of the target system, the size of the data part and etc. The amount and type of input data required varies depending on the selected DoS attack. Once the attack is finished, the output data can be displayed in the console or sent to the API in the form of a JSON object. The output data also contains data from agent and check service.

The check service part is responsible for monitoring the target system and is called before the DoS attack is launched. The monitoring is done via side channel, which is used to separate the control data from the attack data. This check is performed by constantly sending HTTP GET requests to the target system, with a very low frequency of one request per second for the duration of the attack. If the duration of the attack is, for example, 60 seconds, 60 control requests will be sent. In case of the GET request

[illegible]

```
[*] UDP Flood attack started.
[*] udpflood attack started at: Wed Dec 1 19:21:32 2021
[*] Destination: 192.168.0.213, port: 80, attack duration: 60
[*] udpflood attack ended at: Wed Dec 1 19:22:32 2021
[*] Total udpflood packets sent: 9070913
[*] checkservice_total_responses: 58
[*] checkservice_successful_responses: 58
[*] checkservice_failure_responses: 0
[*] checkservice_successful_percentage: 100.0
[*] checkservice_responsetime_ms_avg: 16.498120689655174
[*] checkservice_responsetime_ms_min: 2.641
[*] checkservice_responsetime_ms_max: 46.811
```

returns HTTP status 200<sup>1</sup>, the target system is considered available. In case of any other response, the target system is considered unavailable. The check service saves the availability data of the target system and evaluates it after the monitoring is completed. As part of the evaluation, information about the total number of check packets sent is stored in the output data. This also includes how many of the control packets received successful or unsuccessful response. In addition, in case of successful responses, the minimum, maximum and average response time of the target system is evaluated.

```
python3 ptdos.py -a udpflood -dst 192.168.0.214 -dp 80 -d 600 -dl 2048 --json
```

The application interface is adapted for control from a Linux console. The application is invoked by the command "python3 ptodos", where python3 specifies the use of the Python programming suite in version 3.x.x. Input parameters are then inserted after this command. The individual input parameters vary according to the selected DoS attack. Listing 1 shows an example of calling ptodos application to perform a UDP Flood attack, in which case the target system is a web server with an IP address of 192.168.0.214, which is specified by the parameter -dst. The duration of the attack, specified by the parameter -d, is 600 seconds. The size of the data part of the UDP datagram is 2048 bytes, specified by -dl. When the attack is completed, a JSON object with attack's output data will be created. The creation of JSON object is specified by the --json parameter. The way the data is represented varies depending on whether or not the creation of a JSON object is required. If a JSON object is to be created, i.e., the command contains the parameter --json, it is assumed that the attack is called from the API, so the application in the console does not inform the penetration tester about the current state of the application when performing the attack, but only creates a JSON object after the task is completed and returns it back to the origin API. However, if the JSON object is not required, i.e., the --json parameter is not included in the command, the manual operation of the application is assumed, so the application informs the penetration tester with messages in the terminal about all application states. Once the attack is completed, the result of the

attack with all the output data is printed to the terminal. An example of this terminal print is shown in Fig. 2. The application testing also includes measuring the output data stream of the ptdos application

**Table I:** DoS attacks data stream of selected application, data length is 128 B.

Tool name	UDP Flood	ICMP Flood	SYN Flood
ptdos	209 Mb/s	305 Mb/s	0,052 Mb/s
hping3	112 Mb/s	123 Mb/s	66 Mb/s
trafgen	101 Mb/s	99 Mb/s	69 Mb/s

for UDP Flood, ICMP Flood and SYN Flood attacks. The measured values were compared with the already existing tools hping3 and trafgen. From the measured values, which are displayed in the table I, it can be seen that ptdos is able to generate up to twice the data stream in the case of a UDP Flood attack and up to three times the data stream in the case of an ICMP Flood attack compared to applications hping3 and trafgen. Unfortunately, ptdos does not achieve the intended qualities in the case of the SYN Flood attack, where the generated data stream is very low compared to the other applications as shown in Table I. This is due to the use of the Python scapy library, which as was found out during testing, is not suitable for generating large amounts of data. Therefore, the SYN Flood attack module in the next version of ptdos will have to be redesigned to achieve better results in the generated data stream.

### 3. CONCLUSION

The ptdos application is targeted at IT specialists focused on web application security, to help them with the realisation of penetration and stress testing. The application can be used standalone on local computers or servers, or it can be used by specialists as one of the modules of the complex penterep system, where ptdos is ready for integration. In the penterep integration option, the user no longer uses the local terminal window to control the application, but controls the application using the penterep web interface. The advantage is that it is easy to operate and obtain results, which makes security testing much more efficient. The biggest difference from other available solutions that are able to conduct DoS attacks is the check service, which is able to monitor the status of the target system and evaluate test results, which can be displayed in the terminal window or JSON object, which allows to be further used in other systems using API. A huge advantage of the ptdos application is the modularity it offers, because the individual attacks are created as separate modules and it means that the application is ready for future expansion with new types of DoS attacks which will be discovered. The individual modules do not interfere with the core of the application, which significantly simplifies the implementation of new types of attacks.

### ACKNOWLEDGMENT

The research described in this paper was financially supported by the Technology Agency of the Czech Republic under the grant No. TJ04000456.

### REFERENCES

- [1] CHAO-YANG, Zhang, 2011. DOS Attack Analysis and Study of New Measures to Prevent. *2011 International Conference on Intelligence Science and Information Engineering* [online]. IEEE, 2011, 426-429 [cit. 2022-02-26]. ISBN 978-1-4577-0960-9. Available from: doi:10.1109/ISIE.2011.66
- [2] HADEEL S., Obaid. DoS and DDoS Attacks at OSI Layers. *International Journal of Multidisciplinary Research and Publications* [online]. 2020, 1-9 [cit. 2021-10-11]. ISSN 2581-6187. Available from: doi:10.5281/zenodo.3610833
- [3] TAN, Vincent, Carmen CHEH a Binbin CHEN, 2021. From Application Security Verification Standard (ASVS) to Regulation Compliance: A Case Study in Financial Services Sector. *2021 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*. IEEE, 2021, 69-76. ISBN 978-1-6654-2603-9. Available from: doi:10.1109/ISSREW53611.2021.00046

# Network Media Player with Built-in Class-D Amplifier

M. Liska and M. Štůsek

Department of Telecommunications, Brno University of Technology

E-mail: [xliska19@vutbr.cz](mailto:xliska19@vutbr.cz), [xstuse01@vutbr.cz](mailto:xstuse01@vutbr.cz)

**Abstract**—The article's major contributions summarize theory of the design class-D audio amplifier printed circuit board (PCB) with the possibilities of digital or analog input content. It introduces the possibilities of connecting a digital to analog (DA) converter, audio amplifier, analog filtering and digital volume control. Digital input signal is provided via a Raspberry Pi single-board computer, which is wirelessly connected to receive data in the form of streamed content. Furthermore, it describes the regulation of supply voltages for individual active parts of the printed circuit board. It is possible to connect the output of the amplifier to one pair of stereo speakers and one subwoofer speaker. The article contains a general placement of individual components on the PCB.

**Keywords**—Audio, Amplifier, Filter, Class-D, PCB

## 1. INTRODUCTION

The ever-expanding topic of wireless audio streaming and audio source controlling is an everyday issue of simplifying ordinary life in offices, rooms or other residential buildings areas. The possibilities offered by well established companies are often tailored as proprietary solutions for use with their products, which means that the customer is deprived of the use of their own specific audio set, which would be connectable and controllable from the wireless world. The wireless amplifiers offered are mostly expensive, large in size, incompatible and difficult to connect.

The main idea behind the creation of the amplifier PCB was to make a small, simple device, controllable and connectable with wireless content. For this possibility, we used audio amplifier in (digital) class-D, to which it would be possible to connect content either in analog or digital format. The amplifier outputs were chosen in the form of 2.1 audio for stereo pairs with power up to 30 W at 4  $\Omega$  impedance for each channel, and subwoofer output with power up to 90 W at 4  $\Omega$  impedance. The whole device will be wirelessly connected and managed via Wi-Fi and Bluetooth from the Raspberry Pi single-board computer.

## 2. DEVICE COMPONENTS

The streamed audio content processed by the single-board computer into the digital signal representation first enters the DA converter PCM1791A from Texas Instruments, which was chosen for its main features as 24-bit resolution, sampling frequencies of 128, 192, 256, 384, 512 or 768 kHz with self-correction, user programmability via serial peripheral interface (SPI) or  $I^2C$  bus that can control digital filtering, digital attenuation, noise characteristic and reduction filter [1]. The PCM1791A offers a choice of several digital serial input lines for analog conversion, including  $I^2S$ , supported by the Raspberry Pi. The output signals from the differential balance outputs of the DA converter for each channel are filtered by a low-pass filter of the third order using symmetrical operational amplifiers NE5532 and passive R, C components [1]. As already mentioned, the possibility of digital input content can be omitted and an analog input signal at the line level can be inserted using the jack connector.

The next component on PCB is digital signal volume controller. This controller is also implemented using a single-board computer and since we have to control together two signals for stereo output and two other signals for stereo bass output, the implementation is created using a 4-channel volume controller PGA4311. The use of high-performance internal operational amplifier stages in the PGA4311 volume controller submits low noise and distortion while providing the ability to control up to 600  $\Omega$  loads directly without buffering. Control interface allows for connection to a wide variety of host controllers using the 3-wire serial interfaces, and in addition it support daisy-chaining of conjoint PGA4311 devices [2].

The main part of the whole developed prototype is a Class-D amplifier TPA3244 from the manufacturer Texas Instrument, which states in the catalog sheet the used technology PurePath™ Ultra-HD, for very low

distortion, which should acquire at a power of 1 W and load of 4  $\Omega$  value of total harmonic distortion with noise THD + N = 0.005 %. Its bandwidth is reported to be up to 100 kHz, for high frequency components from High-Definition (HD) sources, and provides an efficiency of up to 94 % [3]. The connection of the amplifier, including its input and output filter part, was realized according to the catalog datasheet TPA3244EVM [4].

## 2.1. Low-Pass Filtration

The input analog signal from the DA converter or jack connector is in the form of stereo audio, for our purposes is necessary to create a bass audio part using analog audio filtering, which will be independently regulated and amplified. In total two signal operations were created - summing the left and right channels into one signal and it's subsequent filtering using RC active filtering by operational amplifiers NE5532. By using the third order of the filter, we ensured a steeper transient response by a decrease of 18 dB per octave, shown in the frequency response in Fig. 1 of the circuit simulation.

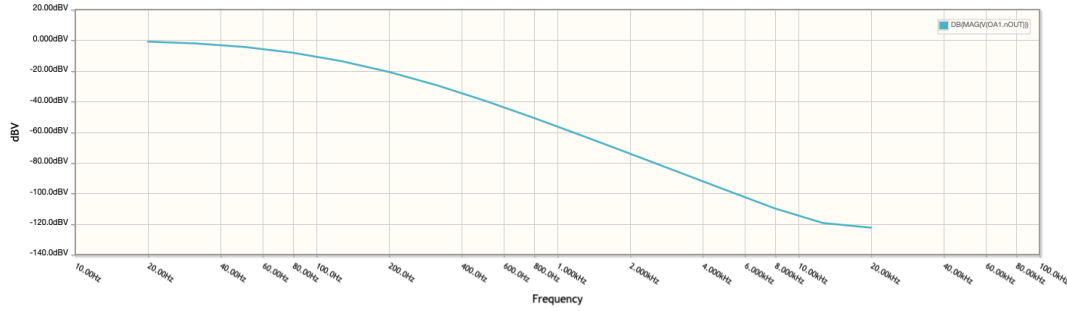


Figure 1: Amplitude-frequency response of LP filter.

The cut-off frequency  $f_c$  with a decrease of 3 dB was determined to be approximately 120 Hz and the values of resistors R and capacitors C were calculated from the equation 1 [5].

$$f_c = \frac{1}{2 * \pi * R * C} \quad (1)$$

This part of the signal operation was inserted before the 4-channel volume controller PGA4311.

## 2.2. Power Supply

Important part of the PCB design was the solution of regulation and stabilization of several supply voltages for individual components. The supply voltage of the whole PCB was chosen to be 30 V with a maximum current of 10 A and is realized using an external switched-mode power supply with power factor correction (PFC) from the company Mean Well. This source is mainly for the TPA3244 amplifier and it is also regulated to power the Raspberry Pi and all components (NE5532, PGA4311, PCM1791A). In the first step, we implemented regulation from a supply voltage of 30 V to 12 V, with an output power of 36 W using a switched regulator LM2576. For its output voltage of 12 V we created a resistance bridge from its feedback, where we chose the resistance of R1 to the value 1 k $\Omega$  and calculated the value of R2 using equation 2 from the catalog datasheet [6]. From this 12 V voltage we supplied the auxiliary circuits of the amplifier TPA3244.

$$R2 = R1 * \left( \frac{V_{out}}{V_{ref}} - 1 \right), \quad (2)$$

where  $V_{ref} = 1,23 \text{ V}$ , a  $V_{out} = 12 \text{ V}$ .

In the second step, we created regulation from our 12 V supply to 5 V. This regulation was divided into two parts due to the separate power supply for the Raspberry Pi with a maximum power of 10 W using the BA50DD0T linear regulator and for others circuits using the NCV8535. The basic converter LM2662 has been added to create negative -5 V for symmetrical power supply of the PGA4311 volume controller.

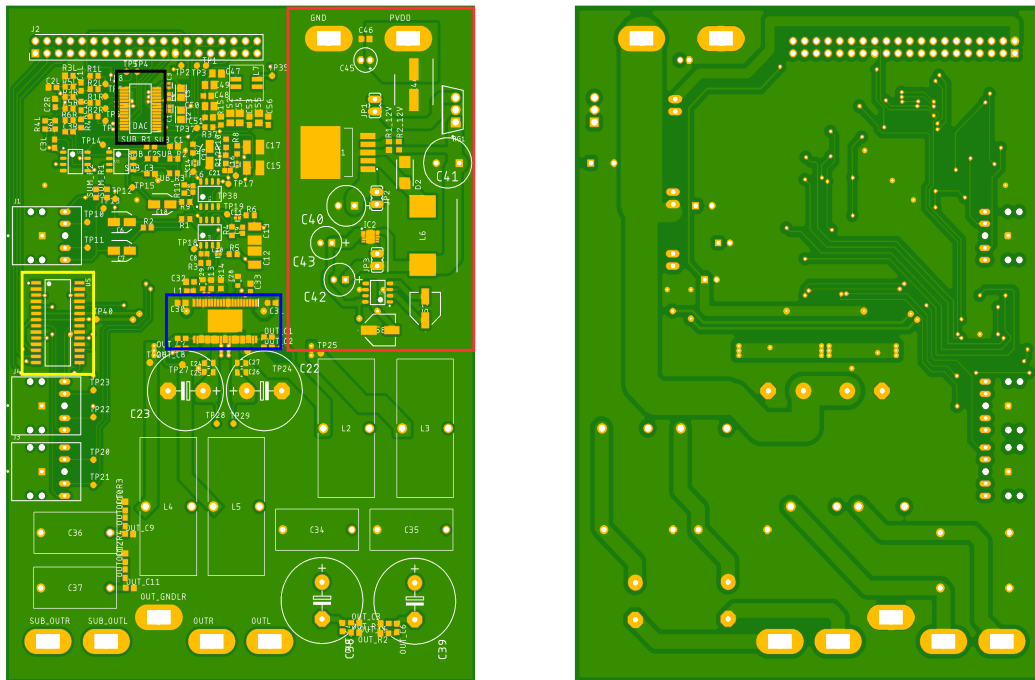
### 2.3. Raspberry Pi Selection

As already mentioned in the article, the acquisition of streamed audio data content and its representation in digital form is mediated by a single-board computer Raspberry Pi, which also provides the platform for controlling the volume of the entire PCB. Raspberry Pi is part of the final product, without which device cannot provide full functionality.

For our purposes was chosen model Raspberry Pi Zero 2 W witch offers 2.4GHz 802.11 b/g/n wireless connectivity and Bluetooth 4.2, along with support for Bluetooth Low Energy (BLE), and modular compliance certification. The device also contains 40-pin GPIO header with support for multiple serial interfaces among others including lines *SPI*, *I<sup>2</sup>C* and *I<sup>2</sup>S*, which we also use [7].

### 3. DESIGN OF PCB

The prototype design of the amplifier's PCB, including all components and connections to the Raspberry Pi pins was created in Eagle software from Autodesk. The overall dimensions of the printed circuit board are 100x140 mm and it was formed on two layers - top and bottom.



**Figure 2:** Visualization of Amplifier PCB.

The power supply management section is located in the upper right corner of the printed circuit board (red part), where the terminals (PVDD, GND) for powering the entire PCB are also located. The TPA3244 amplifier is located in the middle of the printed circuit board, where its upper part serves for input signals and the lower part for output signals (blue part). The PGA4311 digital volume controller is located at the jack input (J1) and output (J3, J4) connectors on the left side of the PCB (yellow part). The digital to analog converter PCM1791A and signal processing section, is located in the upper left corner of the PCB, at the input point (J2) for the GPIO pins from Raspberry Pi (black part).

The Raspberry Pi control is wirelessly implemented using a web application, with the ability to equalize input audio content and the volume control. The volume can be also changed by an externally rotary encoder located on the cover of the device, which is also wired to the Raspberry Pi.

All source codes are in development state, therefore they are not described in this article.



#### 4. CONCLUSION

The article deals with the basic design of a class-D amplifier PCB with the possibility of input in the form of streamed content via Raspberry Pi single-board computer. Content from the Raspberry Pi device can be fed digitally or via a line analog input from a jack connector. The design further includes analog signal processing through filtering, digital volume control using rotary encoder or wireless connection to Raspberry Pi, and amplification. Among other things, it discusses the power supply options of individual components and circuits. The result device is a PCB of amplifier, with the possibility of connecting a stereo speaker pair with a power up to 30W ( $4\ \Omega$  load) for each speaker and one subwoofer speaker with a power up to 90W ( $4\ \Omega$  load). The entire printed circuit board can be connected to different speakers applications depending on the user or customer.

The current state of the device is in the phase of assembly components by surface mounting, for this reason there are no measurement results and source codes in the article.

#### REFERENCES

- [1] Texas Instruments Inc, Datasheet PCM1791A [online]. *24-Bit, 192 kHz Sampling, Advanced Segment, Audio Stereo Digital To Analog Converter*. [ref. 25.11.2021].  
On URL <<https://bit.ly/3oE8yz4>>.
- [2] Burr-Brown Products from Texas Instruments Inc, Datasheet PGA4311 [online]. *4-Channel Audio Volume Control*. [ref. 24.11.2021].  
On URL <<https://bit.ly/3i0YQCY>>.
- [3] Texas Instruments Inc, Datasheet TPA3244 [online]. *60-W Stereo, 110-W peak PurePath™ Ultra-HD Pad Down Class-D Amplifier*. [ref. 25.11.2021].  
On URL <<https://bit.ly/3ya5Zbk>>.
- [4] Texas Instruments Inc, User's Guide TPA3244EVM [online]. *TPA3244 Evaluation Module*. [ref. 25.11.2021].  
On URL <<https://bit.ly/367Vnj9>>.
- [5] PAARMANN, Larry D. *Design and analysis of analog filters: a signal processing perspective*. Boston: Kluwer Academic Publishers, 2001. ISBN 0-7923-7373-1.
- [6] Texas Instruments Inc, Datasheet LM2576 [online]. *LM2576xx Series Simple Switcher® 3-A Step-Down Voltage Regulator*. [ref. 25.11.2021].  
On URL <<https://bit.ly/3ICfqoR>>.
- [7] Raspberry Pi Zero 2 W [online] *Raspberry Pi Zero 2 W* [ref. 8.3.2022]  
On URL <<https://datasheets.raspberrypi.com/rpizero2/raspberry-pi-zero-2-w-product-brief.pdf>>.

# NFT Minting System on Cardano Blockchain

M. Tran<sup>1</sup> and P. Dzurenda<sup>1</sup>

<sup>1</sup>Department of Telecommunications, Brno University of Technology, Brno, Czech Republic

E-mail: [xtranm00@vutbr.cz](mailto:xtranm00@vutbr.cz), [dzurenda@vut.cz](mailto:dzurenda@vut.cz)

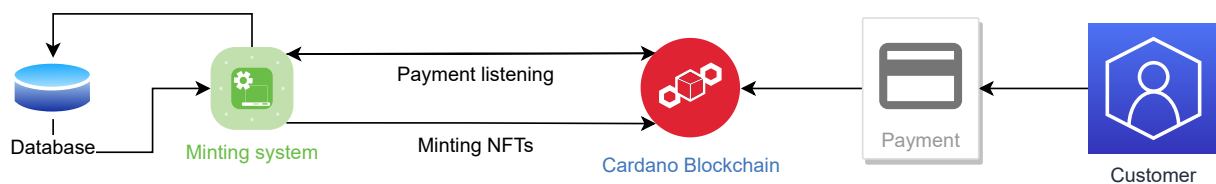
**Abstract**—Nowadays almost everything is digitalized and online. This can cause many security issues for our data. They can be leaked, misused, or even compromised. Therefore, we need technology, which can secure digital assets. The solution to this problem can be blockchain technology and Non-Fungible Tokens (NFT). This article presents our open-source implementation of an automated NFT minting system on the Cardano blockchain. We research the capabilities of NFTs on Cardano and possibilities to create an NFT minting system similar to other successful but proprietary NFT minting systems. Our system is built on Cardano command-line interface commands in combination with Python language for easy handling of the minting process.

**Keywords**—NFT, Non-Fungible Tokens, Minting, Cardano, Blockchain, Python, Cryptocurrency

## 1. INTRODUCTION

In the present world, we need technology that combines safety, transparency, decentralization, and undestroyable proof of ownership of digital assets. Today every digital asset is either saved on just one server or multiple servers. Data of the assets can be changed, deleted, or multiplied by one entity. The owner of digital assets has no guarantee of asset's safety or immutability of their characteristics or uniqueness. This problem can be resolved by blockchain technology and Non-Fungible Tokens (NFT) [1]. A decentralized blockchain with a public digital ledger ensures that the transactions and its data that has been written to a block cannot be changed or deleted. NFT ensures that the asset is unique and irreplaceable. The ownership and characteristics of an asset, for example, digital art or collectibles, can be safely and easily checked on the ledger itself.

In this paper, we implemented an NFT minting system that is fully compatible with the Cardano blockchain and is open-source. This system allows us to set our own policy for NFT minting, define NFTs properties via metadata, and finally, mint and send an NFT to the blockchain. The implemented system is compatible with both testnet and mainnet. The only thing needed to change between them is to adjust one parameter in the minting transaction. Diagram of our system can be seen in Figure 1.

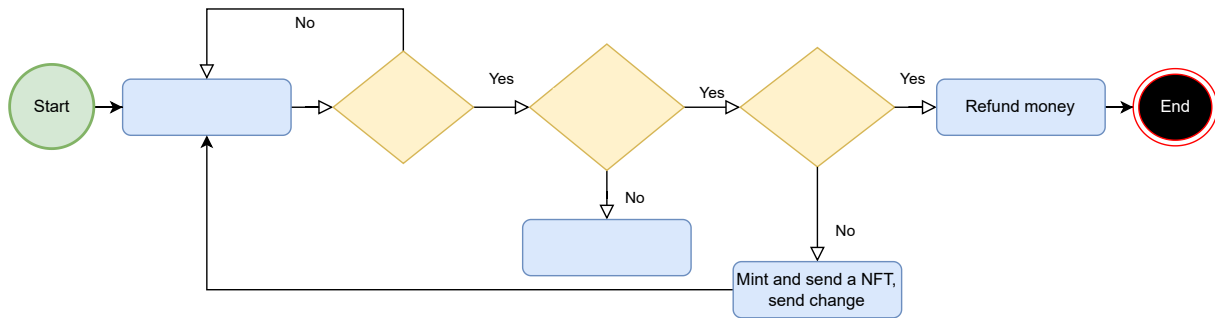


**Figure 1:** General diagram of our NFT minting system.

## 2. NFT MINTING SYSTEM

In Cardano, every interaction with the chain is being done with a transaction through a node. Therefore, the most important prerequisite is a functional Cardano node. Furthermore, we need two Cardano wallets, one cardano-cli type wallet with some ADA to pay for transaction fees, and one Cardano Wallet type wallet for payment listening.

We use Cardano Command Line Interface (CLI) on Cardano node with a combination of Python programming language to create an automated NFT minting system which includes functions for creating policyID, creating metadata from NFT metadata database, the configuration of transactions, payment listening, and refunding change. The system listens for incoming payments and corresponds with creating and sending NFT or refunding a customer. The flow diagram of the system core can be seen in Figure 2.



**Figure 2:** Flow diagram of NFT minting system core.

## 2.1. PolicyID and policy script

PolicyID serves to distinguish NFTs from each other. Value of policyID is derived from a policy script, which is a simple JavaScript Object Notation (JSON) file predefined by Cardano and is used for defining who can mint NFT under the policyID and it also defines a time window when these NFTs can be minted. At the end of the time window, the policy is locked and nobody can mint any NFT under the same policyID ever again. This is the main principle to guarantee the scarcity of an asset. To create a policy script, we firstly generate a policy verification key and signing key. This is done via Cardano node CLI. Next, we generate a hash of the verification key. After that, we define the time window. This is done by getting the current blockchain slot number and adding a number to it. Approximately every second one slot passes [2]. In our case, the time window includes 604800 slots, which corresponds to one week of the minting period. Now, we have everything we need to build a policy script. Our script file is depicted in Listing 1. Finally, we compute the policyID by hashing the policy script. We used Python library subprocess to execute all CLI commands from our Python application itself.

## 2.2. NFT Metadata

Metadata stores policyID and other information such as names of asset and InterPlanetary File System (IPFS) link to the image that will be displayed into an NFT in JSON format. We defined the metadata file as it can be seen in Listing 2.

```

1 policyScriptRaw={
2     "type": "all",
3     "scripts":
4     [{
5         "type": "before",
6         "slot": int(slot)
7     },
8     {
9         "type": "sig",
10        "keyHash": keyHash
11    }]
12 }

```

**Listing 1:** Policy script.

```

1 metadataRaw={
2     "721": {
3         policyID: {
4             assetName: {
5                 "description": description,
6                 "name": name,
7                 "id": int(id),
8                 "image": "ipfs://" + ipfs
9             }
10        }
11    }
12 }

```

**Listing 2:** NFT Metadata.

Because of the IPFS link and custom attributes, we populate the metadata database with those data before starting the payment listening and minting process. We used Comma-Separated Values (CSV) format to do so. From the database, metadata files can be generated and used to mint NFT. We use JSON and pandas library to create metadata files.

## 2.3. Minting and sending NFT

All the communication with the Cardano blockchain is being done by predefined transactions. Therefore, we simply configure the parameters of these transactions through our application. The following actions are being executed with Cardano type wallet. The life cycle of any transaction begins with building it. A building of a minting transaction and its parameters are shown in Listing 3. These parameters can be divided into 4 groups.

The first group defines what net and era we are using. Here, Cardano testnet with the newest era is used. Cardano is using the UTXO model [3], therefore, the transaction input and output have to be there.

This is defined in the second group. These parameters are TXIN (transaction input), TXOUT (transaction output), and change address. Thanks to the usage of the Alonzo era the minting system does not have to mint an NFT to minting wallet and afterward send it to a customer. It can directly mint an NFT to the customer's wallet. It saves time and transaction fees. UTXO hash and index of minting wallet are stored in TXIN. The address where an NFT will be sent, the minimum amount of ADA that has to be included in the transaction, token amount, policyID, and asset name are stored in TXOUT. The remaining assets in TXIN are sent to change address, i.e. the minting wallet address. The third group contains details about minting itself. Parameter mint indicates how many tokens with a particular policyID and asset name will be minted. The next parameter defines what policy script will be used. The last parameter of this group is the metadata JSON file. The fourth group contains information on how many private keys we need to sign the transaction. In the case of our NFT minting system, there have to be two private keys. The last parameter indicates where the built transaction file will be saved.

```

1 cmdTxRaw="cardano-cli transaction build"+\
2   " --testnet-magic 1097911063"+\
3   " --alonzo-era"+\
4   " --tx-in "+xUTXOH+"#"+xUTXOI+\
5   " --tx-out "+address+" "+str(txOutput)+" "+str(tokenAmount)+"
6   "+policyID+"."+assetName+" "+\
7   " --change-address "+addressMint+\
8   " --mint="+str(tokenAmount)+" "+policyID+"."+assetName+" "+\
9   " --minting-script-file "+policyScriptFile+\
10  " --metadata-json-file "+metadataFile+\
11  " --invalid-hereafter "+str(slotNumber)+\
12  " --witness-override 2"+\
13  " --out-file "+txMintRawFile

```

**Listing 3:** Minting build transaction.

The minting transaction needs to be multisigned with two separate private keys (i.e., two wallets). It is because we need permission for using ADA to pay transaction fees, permission for deploying the policy script, and being able to mint under the policyID. Finally, we submit the transaction to the blockchain. The actual mint of NFT happens when the block, where the submitted transaction was placed, gets created on the blockchain.

## 2.4. Payment listening

The most suitable tool for listening for incoming payments is Cardano Wallet. That is because this type of wallet focuses only on wallet functionality thus it provides a higher level of features. For example, this wallet saves easy to read history of all transactions of a particular wallet. For this reason, this wallet is used as a public receiving wallet for the customer to pay for goods. To retrieve the history of payments just a simple cardano-wallet GET request through the command line is needed. The NFT minting system repeatedly fetches a list of transactions and adds new incoming transactions to a database of payments. The system then reads every transaction one by one and determines if the amount of ADA received meets the price of NFT or not.

## 2.5. Sending ADA

ADA is being sent from Cardano Wallet. Here only one POST request command with three parameters is necessary: 1) sender address, 2) header parameter is always the same and defines the content type, and 3) data consist of a passphrase which acts as a security element, address of the receiver, and amount of ADA.

## 3. EXPERIMENTAL RESULTS

The NFT minting system was tested in an environment that simulates real-world usage. There are three possibilities of the outcome. If the amount that the customer sends meets the price of an NFT or is greater then: 1) the system mints, 2) sends NFT, and 3) sends back the change to the customer. Otherwise, the customer is refunded. All of them were tested with expected outcomes. For testing purposes, the customer used a third-party wallet called Nami wallet. In Figure 3, there is a screenshot of an easily readable minting transaction taken from cardanoscan.io, from Nami wallet, and the screenshot of minted NFT from Nami wallet. We can see that NFT parameters: 1) transaction ID, 2) date, 3) asset name, 4) policyID, and 5) asset fingerprint match.

## Transaction Details

Transaction Hash	20e5520da1c6e14cf3c277ecf8a516690ce4f0e2f8d0c0239fae1f7419844850
Block	3263091
Assurance	High 105571 Confirmations
Epoch/Slot	182 / 397691
Absolute Slot	48652091
Timestamp	01/24/2022 11:48:27 AM a month ago
Total Fees	0.187281
Certificates	0
Total Output	241.377018

Transaction ID

20e5520da1c6e14cf3c277ecf8a516690ce4f0e2f8d0c0239fae1f7419844850

Transaction Extra

Minting

1.500000 tA

1 Asset

MOKNFT #002

x 1

Send

As Avatar

Policy 95abef13f28d8785a92177376d7038203a895c709e81e09a674a47a8

Asset asset1mjrps5yjtqju589p8vtykna6ffl4pr673lqr

Policy Id	Asset Name	Fingerprint	Mint Amount
95abef13f28d8785a92177376d7038203a895c709e81e09a674a47a8	moknft002	asset1mjrps5yjtqju589p8vtykna6ffl4pr673lqr	1

**Figure 3:** Minting transaction from cardanoscan.io and Nami wallet and minted NFT displayed in Nami.

There were minimum delays (ms) between incoming transactions from customers and minting transactions from the NFT minting system at the time of testing. However, the time between sending money and receiving NFT depends on the blockchain's load. It could be milliseconds or even hours. Because of single thread usage, there were no collisions in databases and transactions.

## 4. CONCLUSION

In this paper, we present our NFT minting system. We demonstrate how the Cardano command-line interface in combination with Python language can be used for handling the minting process. The NFT minting system is fully compatible with the Cardano blockchain. It supports settings of policyID for NFT minting, setting of NFT metadata, minting, and sending NFT to the blockchain. Experimental results show that our minting system does work and automatically handles all the possible outcomes with minimum time delay.

## REFERENCES

- [1] PARHAM, Arsalan a Corinna BREITINGER. *Non-fungible Tokens: Promise or Peril?* [online]. 13 Feb 2022, 7 [cit. 2022-03-11]. Available from: doi: <https://doi.org/10.48550/arXiv.2202.06354>
- [2] *Cardano Developer Portal* [online]. Zug, Switzerland: Cardano Foundation, 2022 [cit. 2022-03-11]. Available from: <https://developers.cardano.org/>
- [3] CHAKRAVARTY, Manuel, James CHAPMAN, Kenneth MACKENZIE, Orestis MELKONIAN, Jann MÜLLER, Michael PEYTON JONES, Polina VINOGRADOVA a Philip WADLER. *Native Custom Tokens in the Extended UTXO Model* [online]. October 2020, 20 [cit. 2022-03-11]. Available from: <https://iohk.io/en/research/library/papers/native-custom-tokens-in-the-extended-utxo-model/>

# FPGA Digital Circuit for up to 400 Gbps Transfers over Ethernet

D. Kondys<sup>1</sup> and D. Smékal<sup>1</sup>

<sup>1</sup>Brno University of Technology, Czech Republic

E-mail: [xkondy00@vut.cz](mailto:xkondy00@vut.cz), [smekald@vut.cz](mailto:smekald@vut.cz)

**Abstract**—Network cards with a hardware acceleration feature are a popular solution for meeting the ever-increasing demands for throughput in high-speed networks. Utilizing the FPGA (Field Programmable Gate Array) chips as the hardware acceleration elements, this paper presents a generic and highly modular digital circuit for FPGA that manages the transfer of data in form of Ethernet frames at rates reaching up to 400 Gbps. To achieve this, the proposed digital circuit takes advantage of the Ethernet intellectual property (IP) blocks in high-end FPGAs from Intel. By first implementing and fine-tuning it for data rates up to 100 Gbps, the next step is expanding it to reach data rates up to 400 Gbps. The created digital circuit will then be used in the FPGA design for the XpressSX AGI-FH400G network card (among others) created by companies CESNET a.l.e and REFLEX CES. Even though the target data rate is 400 Gbps, this paper focuses on the first step, which is the utilization of the Intel Ethernet hard IP blocks to reach 100 Gbps.

**Keywords**—400 GbE, Agilex, CESNET, Ethernet, FPGA, Intel, NDK, Stratix 10

## 1. INTRODUCTION

To be able to transmit higher and higher volumes of data over the network, high-end network cards utilize hardware acceleration elements. These are usually ASIC (Application Specific Integrated Circuit) or FPGA chips. They can perform certain tasks, like the transmission and reception of packets, instead of a processor so it has more time to work on other things.

The reprogrammability of FPGAs (among other things) makes them ideal for prototyping and research activities. This is one of the main reasons why they are used in high-end network cards. For high throughput speeds (10 Gbps and above), FPGAs often contain hard IP blocks that manage the functions of Ethernet sublayers. This eliminates the need for custom implementations saving many FPGA resources and development time in the process. However, proper configuration is necessary to fully take advantage of these hard IP blocks.

Many FPGA-based high-throughput applications are created for one particular Ethernet speed (like Limago [1]). In these cases, the functions of Ethernet sublayers can be provided by a certain hard IP block with a certain configuration (like in Limago) or it can be custom-implemented like [2], which is also for one Ethernet speed only. As the NDK (Network Development Kit) platform [3] aims to support different types of user applications (more in section 2), a single Ethernet speed would not be sufficient.

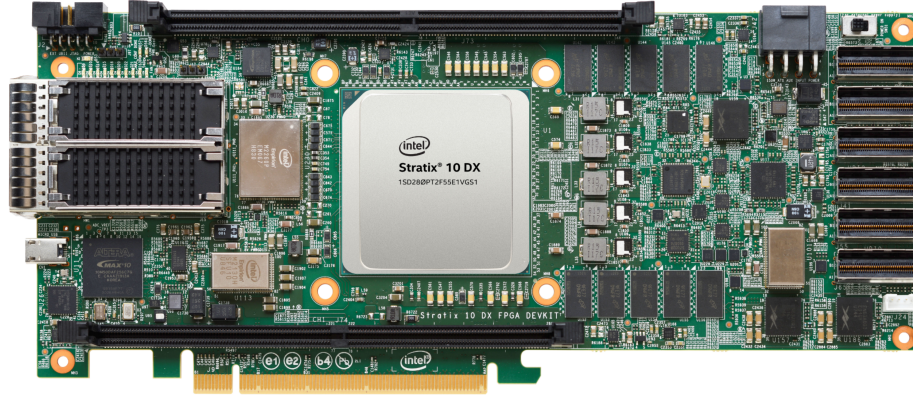
This paper, therefore, presents a generic and highly modular digital circuit called the Network Module that utilizes the available hard IP blocks to connect the serial transceivers of the FPGA with the NDK platform. Thanks to its genericity and modularity, it provides easy scaling from 10 up to 400 Gbps. The first step is to utilize the hard IP blocks on the Intel Stratix 10 DX FPGA to reach Ethernet speeds up to 100 Gbps.

## 2. TECHNOLOGIES

Three network cards are important for this paper. The first and the most important is the Intel Stratix 10 DX FPGA Development Kit, displayed in Fig. 1. This is where the first implementations of the Network Module were tested and fine-tuned. The second network card is the Intel Agilex I-Series FPGA Development Kit. This network card is used for early debugging of the Network Module and other parts of the FPGA design built on top of the NDK platform. Finally, there is the XpressSX AGI-FH400G network card, which is currently in manufacture. This is the target network card that will contain the Network Module as a part of its equipped firmware (FPGA design).

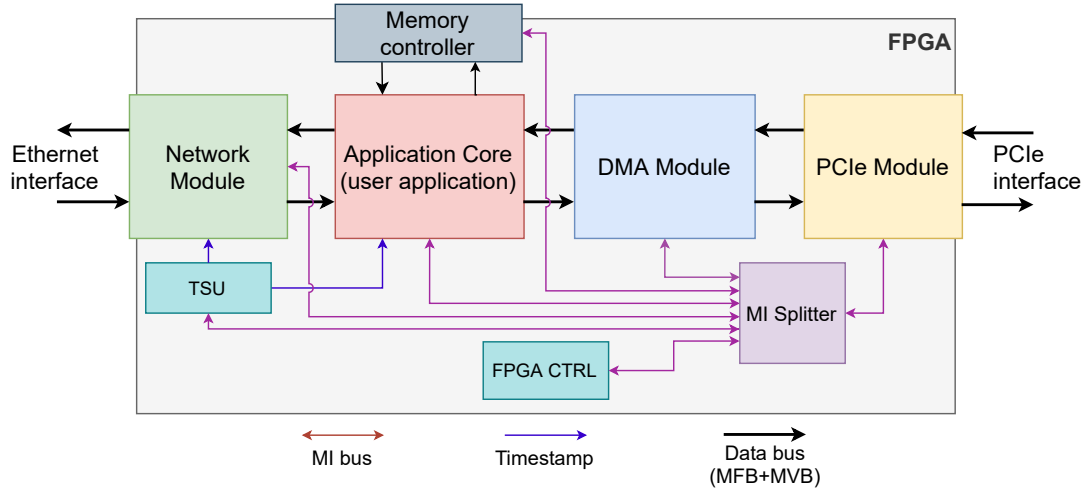
Another thing that was already mentioned is the NDK platform. It contains many components for FPGA, some of which were utilized in the Network Module. It also contains complete modules that





**Figure 1:** The Intel Stratix 10 development kit.

were built on top of the NDK platform as well (like the 400 Gbps DMA Module [4]) and together they form the basic firmware (shown in Fig. 2) for many FPGAs across multiple vendors. It enables the transfer of frames from the Ethernet interface to the RAM of the host PC via the PCIe interface and vice versa. The Application Core is where the user can implement an arbitrary custom application.



**Figure 2:** Overview of the basic firmware built on top of the NDK platform.

### 3. INTEL HIGH-END FPGAS AND THEIR ETHERNET HARD IP BLOCKS

The Stratix 10 DX FPGA has been around for some time, meanwhile Agilex is quite new. They both contain some transceiver tiles (the E- and F-tile) with the support of the IEEE 802.3 standard for Ethernet (revision 2015 and 2018, respectively). These transceiver tiles can be utilized by instantiating the E- or F-tile hard IP blocks in an FPGA design.

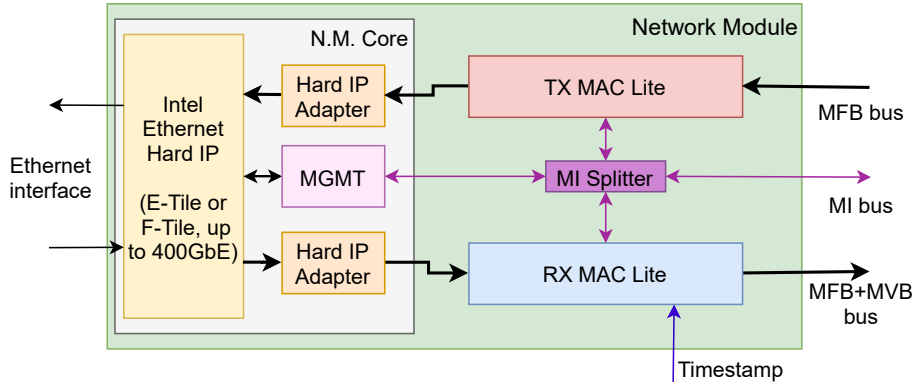
E-tiles are a part of Stratix 10 DX as well as of some Agilex devices. Among the configuration options of the E-tile hard IP are the Ethernet speed and the number of Ethernet channels in the form of *channel\_count*  $\times$  *data\_rate*. Some of the options are: 4 $\times$ 10 GbE, 4 $\times$ 25 GbE and 1 $\times$ 100 GbE (for transceivers in NRZ mode). This is why the maximum data rate on Stratix 10 is 100 Gbps (per one optical port). This and more can be found in the given user guide [5].

F-tiles are one of the new features of the Agilex FPGAs and therefore are not a part of any of the previous FPGAs (like the Stratix 10 DX). It supports the same options as the E-tile as well as many others, e.g. 2 $\times$ 40 GbE, 8 $\times$ 50 GbE, 2 $\times$ 200 GbE and 1 $\times$ 400 GbE. This implies the possibility to reach data rates up to 400 Gbps, so that is why these tiles (and FPGA) are used for the XpressSX AGI-FH400G network card. Both of these hard IP blocks have their unique client interface.

#### 4. THE NETWORK MODULE

The Network Module exchanges Ethernet frames with the rest of the FPGA design according to the MFB (Multi-Frame Bus) protocol, which specifies how the frames are transferred over the NDK platform. It was designed to be modular (for easy scaling and development), generic (e.g. support different data rates) and to allow access to its local CSR (Control and Status Registers). The basic architecture of the Network Module is displayed in Fig. 3.

The key part of the architecture is the Network Module Core. It contains one of the generated (and appropriately configured) Ethernet hard IP blocks, also adapters to connect its unique client interface to the MFB bus and a management unit (MGMT) for access to its CSR. This means that the inner structure of the Network Module Core depends on the selected hard IP block. However, the rest of the Network Module is not affected by this and remains the same.



**Figure 3:** The basic architecture of the Network module.

The TX and RX MAC Lites provide a buffer (and therefore some form of backpressure), a stat unit that calculates statistics about processed frames and a crossing between different clock domains. The local CSR are controlled through read and write requests over the MI (Memory Interface) bus.

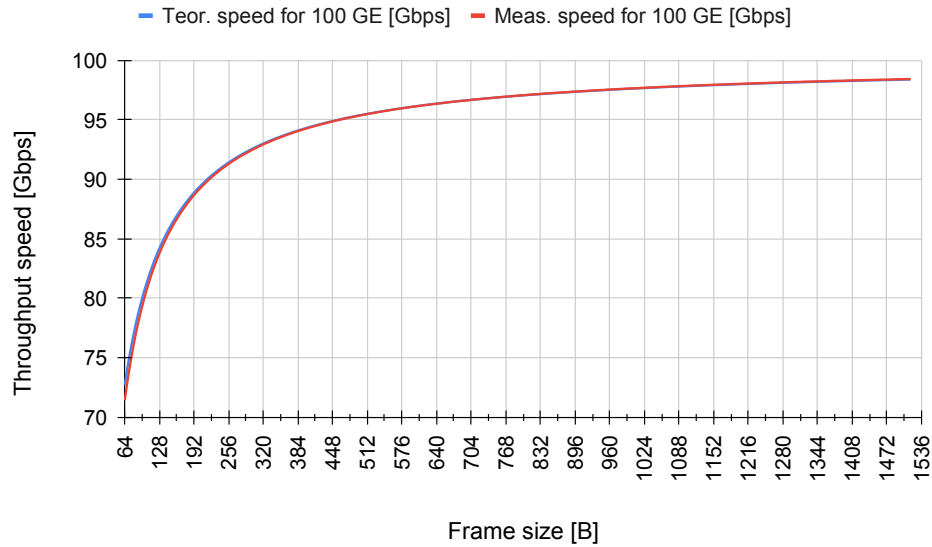
For the Network Module to be generic, it must support both types of the mentioned hard IP blocks and all of their possible options in the form of a number of channels and the data rate. From the point of view of the channels, this means instantiating some components (e.g. the adapters and MAC Lites) as many times as there are channels and splitting and merging them from/into one stream. From the point of view of the data rate, this means adjusting the width of the buses. For multi-port configurations, every component of the Network Module has to be instantiated that many times again.

#### 5. TESTS AND EVALUATION

Even though both architectures of the Network Module (E- and F-tile) have been implemented, only the one with the E-tile hard IP has been tested. The tests were executed in hardware on the Intel Stratix 10 DX FPGA Development Kit where the two ports were connected in a loopback by an optical cable. Ethernet frames were generated before the TX MFB side of the Network Module, passed through the TX part of the Network Module out to the cable and back in again, through the RX part of the Network Module and out to a component that measured the throughput speed (data rate).

The throughput speed was measured for frames of different lengths and put into comparison with the maximum theoretical throughput speed on Ethernet. The speed was measured for lengths from 64 to 1518 B as these are the common Ethernet frame lengths. The Network module can however deal with jumbo frames (frames larger than 1518 B) for lengths up to 16383 B. The results are displayed in Fig. 4 for the 1×100 GbE configuration of the Network Module. The measured throughput speed is virtually identical to the maximum speed across all frame lengths. Similar results were obtained for 4×25 GbE and 4×10 GbE configurations.

Another form of verification of the Network Module was based on statistics collected from the MAC lites. The statistics allowed for the detection of lost frames by comparing the number of transmitted and received frames, counting erroneous frames (e.g. with incorrect Frame Check Sequence) and discarded frames and more. Due to the collected statistics and the measured throughput speed being in accordance with expected values, the Network Module is considered functional.



**Figure 4:** The throughput for 1×100 GbE Network Module.

## 6. CONCLUSION

The objective of this paper was to implement the Network Module that would enable the rest of the FPGA design to communicate with the serial Ethernet interface at data rates up to 100 Gbps. This was achieved by utilizing the E-tile hard IP blocks on selected Intel FPGAs. With the E-tile architecture of the Network Module proved functional, it enables user applications to communicate over Ethernet at speeds 1×100, 4×25 or 4×10 Gbps. The network cards must however contain the mentioned hard IP blocks. Furthermore, due to the modular and generic design of the Network Module, different types of Ethernet hard IP blocks even from different vendors can be added easily, making it even more universal.

This is important, as the Network Module targets data rates up to 400 Gbps which cannot be achieved with the currently utilized hard IP blocks. Due to the modularity and genericity of the Network Module and also the lessons learned with the E-tile architecture, the F-tile architecture allowing throughput speeds up to 400 Gbps needed rather minor adjustments. This architecture is currently under testing and the first results look positive. After verifying its proper function, the Network Module will be used as a part of the FPGA design for the XpressSX AGI-FH400G network card (as soon as it becomes available).

## REFERENCES

- [1] RUIZ, M., D. SIDLER, G. SUTTER, G. ALONSO and S. LÓPEZ-BUEDO. *Limago: An FPGA-Based Open-Source 100 GbE TCP/IP Stack* [online]. 2019 [cit. 2022-03-24]. Available at: [doi: 10.1109/FPL.2019.00053](https://doi.org/10.1109/FPL.2019.00053)
- [2] TOYODA, H., G. ONO and S. NISHIMURA. *100GbE PHY and MAC layer implementations* [online]. 18 March 2010 [cit. 2022-03-24]. Available at: [doi: 10.1109/MCOM.2010.5434377](https://doi.org/10.1109/MCOM.2010.5434377)
- [3] Liberouter / Cesnet TMC group. *Network Development Kit* [online]. [cit. 2022-03-11]. Available at: <https://www.liberouter.org/ndk/>
- [4] KUBÁLEK, J., J. CABAL, M. ŠPINLER a R. IŠA. *DMA Medusa: A Vendor-Independent FPGA-Based Architecture for 400 Gbps DMA Transfers* [online]. 2021 [cit. 2022-03-24]. Available at: [doi: 10.1109/FCCM51124.2021.00045](https://doi.org/10.1109/FCCM51124.2021.00045)
- [5] Intel Corporation. *E-tile Hard IP User Guide: E-Tile Hard IP for Ethernet and E-Tile CPRI PHY Intel® FPGA IPs* [online]. 04.08.2021 [cit. 2021-11-27]. Available at: <https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/ug/ug20160.pdf>

# Impact of devices connected to powerline on selected BPL topologies

Michal Mahút<sup>1</sup> and Petr Musil<sup>1</sup>

<sup>1</sup>Department of Telecommunications, Brno University of Technology, Brno, Czech Republic

E-mail: [xmahut00@vutbr.cz](mailto:xmahut00@vutbr.cz), [xmusal56@vutbr.cz](mailto:xmusal56@vutbr.cz)

**Abstract**—The following article deals with the impact of device originated noise on BPL (Broadband over powerline) communication for different topologies of BPL modems. In household scenario, every device connected to power line has an impact on transmission spectrum and limits throughput. Different topologies and two types of devices were tested and evaluated. The aim of this article is to determine the appropriate connection of BPL modems into the powerline with the lowest possible decrease of data throughput.

**Keywords**— BPL modem, PLC, noise level analysis, powerline, throughput, topology

## 1. INTRODUCTION

BPL is technology which operates in the 1.8 – 500 MHz and reaches data rates at the physical layer from Mbps to Gbps [1, 2]. Several standards for broadband over power lines as HomePlug AV, HD-PLC, ITU-G.hn were developed [3]. BPL communication is promising technology to transmit information in cases, where we cannot use other technologies such as optical or wireless approach. Nowadays it is possible to transmit information via BPL with high throughput up to 800 Mbps according to [4]. Communication over powerline is affected by noise that is produced by devices connected into powerline. The main objective is to find a capable topology for connecting BPL modem to keep the impact of devices on communication as low as possible. Devices which fed and running are main source of noise in the frequencies where BPL operate. In addition to noise, communication is also affected by impedance of device and its power cord. This measurement focuses on the real conditions, therefore typical office devices were connected into the powerline cord in close proximity to the BPL modem. The noisiest device is a desktop PC according to [5] which has power spectral density up to -110 dBV/Hz in the frequency range below 4 MHz. Such noise might have time-variant characteristic with period equal to the mains period, which might have effect on throughput fluctuations. As [6] describes situations, when vendors state that their products fulfill Electromagnetic Compatibility requirements EMC - that means that it conforms to noise emission standards, led to (50 % drop) of data throughput of BPL modems. According to [7] the maximum PSD (power spectral density) of BPL is below 30 MHz. Due to this fact we need to ensure interference minimization for optimal transmission conditions.

## 2. MEASUREMENT

For these measurements 2 BPL modems were chosen, prototype modem (D1) and commercial modem (D2). Both devices utilize G.hn protocol and operate on similar bandwidth, D2 works on bandwidth, which is 6 MHz higher than bandwidth used for D1. Modems support MIMO communication, which enables higher throughput in network. An overview of the used devices is displayed in Table I. Modem D1 was used as a tool for acquisition of noise levels in the operating frequency range.

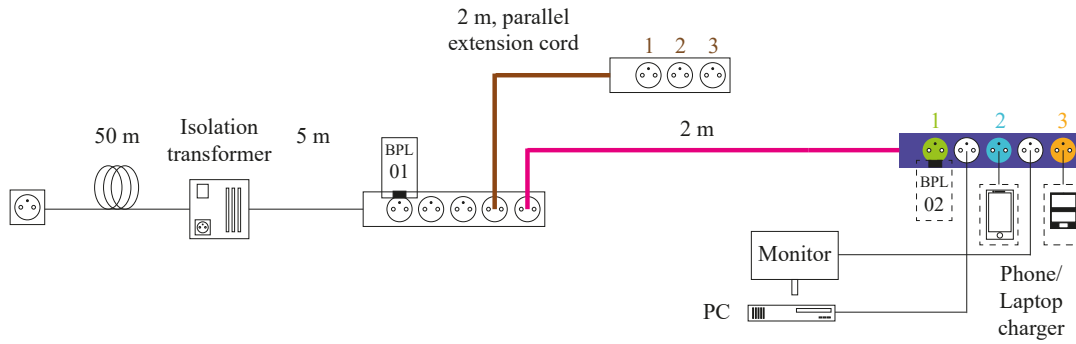
**Table I:** BPL modems.

Label	BPL Modem	Standard	Bandwidth	PHY throughput MIMO
			[MHz]	[Mbps]
D1	Prototype	G.hn	2 – 80	1200
D2	Commercial	G.hn	2 – 86	2000

50 m long extension cord connects the measured topology with the plug. Isolation transformer is used for minimizing the impact of power line interference. The 5 m long extension cord with first BPL modem is connected to isolation transformer. This described part of the layout is similar for all measurements to ensure the most efficient measurement conditions. All measured values come from position BPL02 (second).

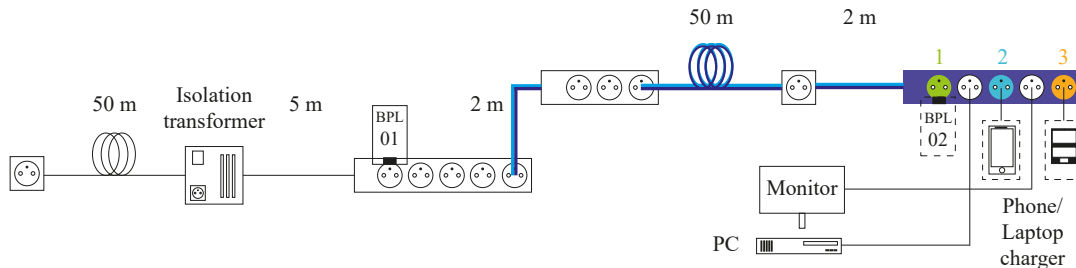
The first measured topology is shown in the Figure 1. Two 2 m extension cords are connected to the 5 m one. The second BPL02 modem is at first connected in parallel extension cord (brown), which is in this case used for measuring BPL data throughput. The second extension cord is used for powering PC, monitor, mobile phone and laptop charger. These interfere with communication and simulate noise on powerline in household scenario. Starting condition (0 in throughput Figure 5) was measured when all devices were unplugged from second extension cord.

Second extension cord (magenta) was used for measurement of the noise and throughput on BPL modem, when second-BPL02 was connected in different positions: 1-green, 2-blue (Phone charger was moved to position 1), 3-orange (Laptop charger was moved to position 2), 0-purple (all devices were unplugged). The yellow and magenta cables represent the throughput measurements in Figure 5 and Table II for different positions. The coloured numbers and sockets represent position of BPL modem for Figure 3, when the noise level was measured.



**Figure 1:** Topology with 2 m extension cord.

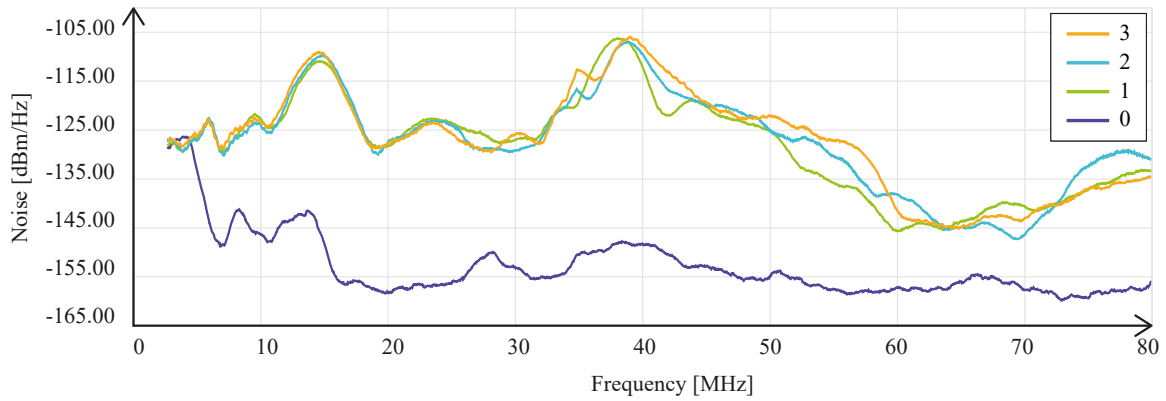
The second topology uses the connection with isolation transformer described above. Topology contains 2 m, 50 m and 2 m extension cords, to which devices with their interference and second BPL modem are connected. The measurement method used on the last extension cord and colour description are identical to the method used on the first topology, results of the noise impact are displayed in Figure 4. Cyan and dark blue cables were used for the throughput measurements in Figure 5 and Table II for different positions.



**Figure 2:** Topology with 54 m extension cord.

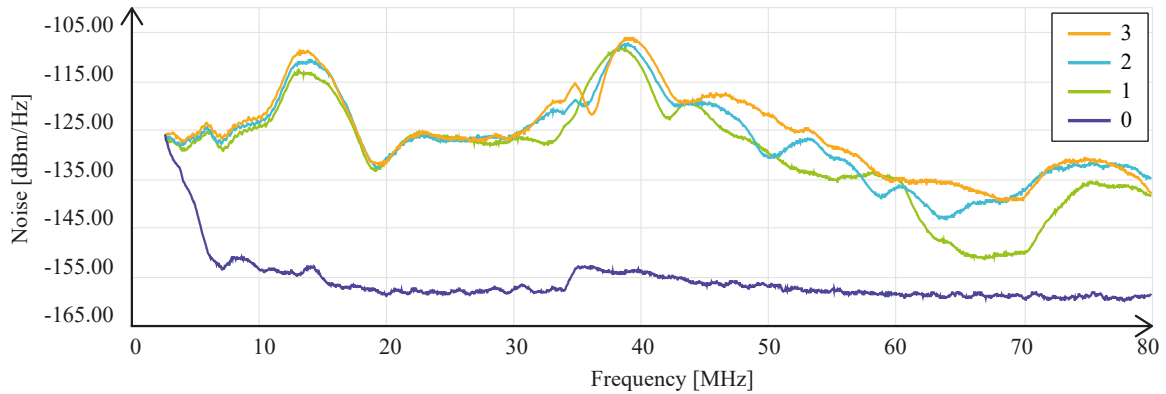
### 3. RESULTS AND DISCUSSION

Graph captured in Figure 3 shows noise level dependence on frequency from first topology measurement, when results are measured on second extension cord. 0-purple curve represents measurement when BPL modem was in position 1 and devices were unplugged. 1, 2, 3 curves represent change of position of the devices and BPL in extension cord as shown in 1. In Figure 3 we can see that after the devices are connected, noise is increased by 10 – 15 dB. Green curve represents position 1 in cord with noise level lower than other two positions.



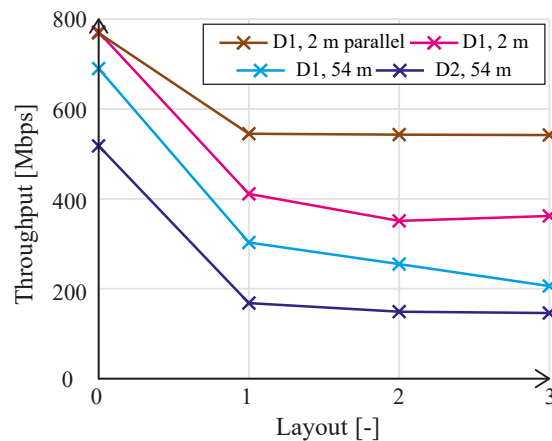
**Figure 3:** Graph of the noise level for topology with 2 m extension cord.

Noise levels shown in Figure 4 represents measurement of the second topology. Curve 0 has flatter characteristic than in previous measurement (Figure 3), because 50 m cord might have filtered out more noise coming from the powerline. Noise level in the frequency band above 45 MHz is higher than in previous measurements. This is due to the lower BPL power on these frequencies. As in the previous measurement, the lowest noise level can be seen in position 1.



**Figure 4:** Graph of the noise level for topology with 54 m extension cord.

Figure 5 shows comparison of throughput according to position of BPL modem and devices in extension cord for first and second topology. Measured values can be found in the Table II. Brown and magenta curves are for D1 in first topology. Cyan and dark blue for D1, D2 devices and second topology. Layout 0 represents cord without connected devices with only BPL modem. 1, 2, 3 are positions in the cord.



**Figure 5:** Throughput – position in cord.



**Table II:** Throughput for different layouts and devices.

<b>BPL Modem</b>		<b>D1</b>	<b>D1</b>	<b>D1</b>	<b>D2</b>
<b>PL length [m]</b>		2	2, parallel	54	54
		<b>Throughput [Mbps]</b>			
<b>Layout</b>	without devices (0)	771	769	690	518
	1	411	545	303	168
	2	351	543	255	149
	3	362	542	206	146

#### 4. CONCLUSION

As described before two topologies were measured. From result of the measurement with 2 m extension cord (1), it can be concluded that use of parallel extension cord with BPL modem connected and second extension cord for devices is better. Parallel connection has 29 % throughput drop. In case of BPL modem connected in the second cord beside the working devices is the throughput drop much higher at 47 %. Significant discovery is when the usage of the parallel cord is not possible. In this case position 1 in the second cord shows the best results, others positions show throughput drop of at least 15 %. An overview of the throughput change is displayed in the Table II. The reason for decrease in throughput position 2 and 3 is increase of noise level, which can be visible in Figure 3.

Second measurement (2) contains comparison of real use cases with two BPL modems D1, D2 at 54 m powerline with impact of connected devices. From Table II and Figure 5 is visible the large decrease in throughput influenced by connected active devices. In position 1 the decrease in throughput is 56 % on D1 and 68 % for BPL D2 compared to the case, when no devices are connected. From the results it is visible, that the position 1 has noticeably lower throughput drop compared to other two measured positions. Throughput drop is 16 % for D1 and 11 % for D2.

#### REFERENCES

- [1] P. Mlýnek, M. Rusz, L. Beneš, J. Sláček, and P. Musil, "Possibilities of Broadband Power Line Communications for Smart Home and Smart Building Applications", *Sensors*, vol. 21, no. 1, 2021, doi [10.3390/s21010240](https://doi.org/10.3390/s21010240).
- [2] L. T. Berger, A. Schwager, and J. J. Escudero-Garzás, "Power Line Communications for Smart Grid Applications", *Journal of Electrical and Computer Engineering*, vol. 2013, pp. 1-16, 2013, doi [10.1155/2013/712376](https://doi.org/10.1155/2013/712376).
- [3] G. Lopez, J. Matanza, D. De La Vega, M. Castro, A. Arrinda, J. I. Moreno, and A. Sendin, "The Role of Power Line Communications in the Smart Grid Revisited: Applications, Challenges, and Research Initiatives", *IEEE Access*, vol. 7, pp. 117346-117368, 2019, doi [10.1109/ACCESS.2019.2928391](https://doi.org/10.1109/ACCESS.2019.2928391).
- [4] R. Fujdiak, J. Slacik, M. Orgon, P. Mlynek, J. Misurec, J. Hallon, and J. Halgos, "Investigation of Power Line Communication and Wi-Fi Co-existence in Smart Home", *2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 1-4, 2018, doi [10.1109/ICUMT.2018.8631197](https://doi.org/10.1109/ICUMT.2018.8631197).
- [5] M. Antoniali, F. Versolatto, and A. M. Tonello, "An Experimental Characterization of the PLC Noise at the Source", *IEEE Transactions on Power Delivery*, vol. 31, no. 3, pp. 1068-1075, 2016, doi [10.1109/TPWRD.2015.2452939](https://doi.org/10.1109/TPWRD.2015.2452939).
- [6] A. S. de Beer, A. Emleh, H. C. Ferreira, and A. J. H. Vinck, "Influence of LED tubes on the throughput of an indoor broadband PLC channel", *2017 40th International Conference on Telecommunications and Signal Processing (TSP)*, pp. 251-254, 2017, doi [10.1109/TSP.2017.8075980](https://doi.org/10.1109/TSP.2017.8075980).
- [7] S. Battermann and H. Garbe, "Influence of modern broadband inhouse PLC transmission on short-wave reception", *2015 IEEE International Symposium on Electromagnetic Compatibility (EMC)*, pp. 283-288, 2015, doi [10.1109/ISEMC.2015.7256173](https://doi.org/10.1109/ISEMC.2015.7256173).

# Literature Overview of Fault Tolerant Electrical Machines

V. Sizonenko, O. Vitek

Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech Republic

E-mail: [vitaliy.sizonenko@vut.cz](mailto:vitaliy.sizonenko@vut.cz), [viteko@vut.cz](mailto:viteko@vut.cz)

**Abstract**—In this paper, a literature overview of fault tolerant electrical machines (FTEM) is presented. Various electrical machine designs are shown, their advantages as well as design and technology challenges are briefly discussed. It was found that one of the most popular types of FTEM is permanent magnet synchronous machine (PMSM) with fractional slot concentrated winding. In order to define boundaries of starting parameters for the design of FTEM, several important aspects of design of electrical machines are also provided.

**Keywords**—motor, generator, fault tolerant electrical machine (FTEM), permanent magnet synchronous machine (PMSM), fractional slot concentrated winding, tooth-coil winding.

## 1. INTRODUCTION

Modern requirements for reliability of safety critical electrical drive systems, e.g. electrical vehicles, aerospace applications, wind power generation applications, and technological limitations during manufacturing of electrical machines provide solid foundation for the development of fault tolerant electrical machines. The inter turn faults in the stator coils of an electrical machine can sometimes occur despite precise quality control of its manufacturing process. Surge voltage tests of the coils before their connection have own limitations and are not capable to detect all potential defects of inter turn insulation. Amplitude increase of the test voltage pulse for detection of small defects in the insulation can damage it and cause a flashover between the first and the last turn of the coil. There are more potential aspects which can lead to a failure of a winding, e.g. weak joints between rotor bars and short-circuit rings of induction motors, partial discharges even in low-voltage machines caused by frequency converters, different issues with temperature detectors, harsh environment conditions. On the other hand, significant advances in technology of power electronics and frequency converters allow to overcome these challenges by applying of wider set of winding arrangements, magnetic circuit designs and control methods of electrical machines. An excellent example of such a development is switched reluctance machine (SRM), which has cheap and rugged construction, but demanding requirements for both its intelligent control and electromagnetic calculations, taking into account saturation effects of the magnetic circuit as a function of rotation angle. Still, such a machine has its limitations in some applications as well. For example, the usage of SRM in fuel-filled pumps would lead to a significant increase of hydraulic losses, and consequently to the reduction of the efficiency of the fuel pump [1]. Further, if such an important demand as high torque density applies, SRM will concede to permanent magnet machine (PMSM), which is shown in [2]. This requirement is especially relevant e.g. for electric aircraft propulsion applications, which have been gaining popularity in recent years. Consequently, due to the high performance characteristics of PMSM, this type of electrical machines has received high attention among other fault tolerant electrical machines (FTEM), which is illustrated e.g. in the literature reviews [3] and [4]. Inherently, PMSMs have own design and technological challenges, such as impossibility of deexcitation of the machine, poor default fault tolerant capabilities, risk of demagnetization of permanent magnets (PM) due to their overheating or short circuit in the winding, gluing and bondage of PMs to the rotor, limited speed range, high cost. Accordingly, cooling capabilities of a PM FTEM must be able to manage with excessive heating of permanent magnets during turn-to-turn short circuit as well. Thereby, all above mentioned factors demand careful attention during design of FTEM for the particular application. This paper will provide some achieved results in the development of FTEMs and several important aspects of their design.

## 2. EXISTING SOLUTIONS

Different measures listed in literature to meet fault tolerant requirements, reduction of losses and torque ripple are represented in this chapter. Solutions below are intended to introduce designer to the existing achievements and help to select, adopt and develop appropriate methods for design of relevant FTEM.

A four-phase fault tolerant permanent magnet machine for an engine fuel pump was presented in [5]. High specific output for this machine has been achieved with help of fuel cooling and PMs arranged into Halbach array. The four-phase motor has six poles rotor and four stator coils, wound around four of the eight stator teeth. It was demonstrated, that the drive continued to produce rated output even in the presence of faults in the machine.

Torque ripple and converter power rating of a duplex 3-phase machine are significantly lower compared to a duplex 2-phase machine, which has been shown in [6]. However, a duplex 2-phase machine has a higher average torque capability. Also, an effective technique based on search coils has been presented for detection of winding short-circuits.

Surface PM machines with different multi-layer FSCWs and different slot/pole combinations were analyzed for turn-to-turn faults in [7]. Among other conclusions it has been shown that turn-to-turn fault between two different phases generates lower fault current than turn-to-turn fault within the same phase.

Comparison of fault tolerant permanent magnet motors with five different topologies was presented in [8]. These are one surface-mounted topology, one surface-insert topology and three interior topologies (spoke-type, conventional and V shape). It has been shown that each topology has a potential for use in suitable applications.

Detailed design, tests and analysis of a 175 kW main power fault-tolerant embedded PM generator for aerospace applications was presented in [9]. Before the design of the generator, a literature review was provided, including the guides and references for the choice of the feasible number of phases, stator slots and rotor poles. After test results it was shown, that the generator can deliver the full power of 175 kW over the wide 4:1 speed range with regulated DC bus voltage.

An example of a water-cooled five-phase fault tolerant 20-slot/18-pole permanent magnet synchronous machine with single layer fractional slot concentrated winding for electric vehicles was investigated in [10]. Among other results it was shown, that after segmentation of PMs, eddy current losses in them are decreased by more than 50%.

For the given rotor diameter and armature current density, fault tolerant PMSM is shorter, has larger slot areas and higher number of turns per slot in comparison to a non-fault tolerant PMSM, which was shown in [11]. Further, the advantages of short-circuit ratio value = 1 and a shorter stack length were discussed.

An asymmetrical six-phase salient-pole synchronous generator for the aircraft power system was proposed in [12]. The pulsating torque of such a machine is reduced due to the fact that the 5th and 7th harmonic components of the MMF wave have no influence on the air-gap flux thanks to the asymmetric phase displacement in the winding of the machine.

A 5-phase fault tolerant outer-rotor PM machine with flux barriers in rotor yoke was designed in [13]. The armature reaction magnetic field in such a machine is greatly reduced without affecting the main pole magnetic field. As a result, eddy current losses in the PMs are reduced as well.

Different rotor structures of fault tolerant PMSM are discussed from the perspective of fault tolerance capability in [14]. Irreversible demagnetization of the PMs during short-circuit fault is avoided by the decrease of short-circuit current due to the increased self-inductance, using appropriate air-gap length and stator slot size. The effective limitation of the short-circuit current by the proposed rotor structure was demonstrated with help of simulation.

An integrated modular motor drive system with three, four, symmetric and asymmetric six phase winding configurations was examined for fault tolerance capability in [15]. It was shown, that symmetric six-phase machine is the most promising configuration by means of fault tolerance and redundancy among all configurations.

A comparison of optimized tooth-coil-winding modular permanent magnet synchronous machines with U-shape and H-shape stator segments for fault-tolerance applications was presented in [16]. It was shown that U-shape machines provide minimum mutual inductance and develop the highest mean torque and efficiency, which makes them suitable for high-torque and fault-tolerant applications. Instead, H-shape machines showed lower torque ripple and are less appropriate for fault-tolerance applications.

### 3. DESIGN CONSIDERATIONS

The following requirements for the fault tolerant electrical machines have been noted in literature as one of the most important: higher number of phases, implicit limiting of fault currents, complete electric isolation between phases, magnetic isolation between phases, effective thermal isolation between phases, physical isolation between phases [3]. It has also been shown that the fractional slot concentrated windings are one of the most popular solutions in applications requiring high fault tolerant capability [17], with the advantages of high power density, high efficiency, short end turns, high slot fill factor particularly when coupled with segmented stator structures, low cogging torque, flux weakening capability [3].

The values of linear current density  $A$  and air gap flux density  $B_\delta$  are very important in the design process of an electrical machine, drafting the size of its rotor at the very beginning. Further, the value of current density  $J$  defines amount of the copper required for the delivery of the requested power. Often, FTEMs have unusual environmental conditions and duty cycles, e.g higher output power during limited period of time, operation under one or several faulty phases, operation under extremely low or high temperatures, etc. That means the electric loadings must be considered individually for each specific application. To illustrate this diversity, electric loadings given in [1], [11], [14], [18], [19] references are summarized in the Table I, where  $P$  is the output power,  $D_r$  is the rotor diameter and  $l$  is an active length. If the value of  $A$  is not given explicitly in a reference, it can be estimated using the formula:

$$A = \frac{2 \cdot I \cdot N \cdot m}{\pi \cdot D_\delta}, \quad (1)$$

where  $I$  is the current in a turn,  $N$  is the total amount of turns in a phase,  $m$  is the number of phases,  $\pi D_\delta$  is the circumference of the air gap.

**Table I:** Electrical loadings of some fault tolerant PMSMs

Ref.	$A$ , [kA/m]	$J$ , [A/mm <sup>2</sup> ]	$P$ , [kW]	$D_r \times l$ , [mm]	Winding	Notes
[1]	174	18.7	15	$68 \times 40$	6-phase	fuel-filled
[11]	$\approx 200$	10, 15, 20	30	$150 \times 100$	poly 3-phase	no info on cooling
[14]	36	22.5	20	$127 \times 100$	5-phase	$T_{amb} = -183^\circ\text{C}$
[18]	36.5	4.4	1	$71 \times 90$	dual 3-phase	
[19]	5.5	11	14.7 kVA	$120 \times 70$	3-phase	$k_{Cu} = 0.15$

One of the core challenges for the design of FTEMs is to combine fault tolerant capabilities of the machine with the required torque density, which constitute a direct contradiction to each other. For example, open slots close to rectangular shape and unequal teeth width can contribute to the higher torque density of the machine, whereas semi-closed slots or closed slots can significantly increase phase inductance of the machine and reduce rotor eddy current losses.

The short circuit current of a synchronous machine is defined by its d-axis inductance, which includes magnetizing and stator leakage inductances  $L_d = L_{md} + L_{s\sigma}$ , and rotor leakage inductances. Magnetizing inductance can be calculated acc. to the equation (2):

$$L_{md} = \frac{\mu_0}{\pi} \cdot \frac{m N_{ef}^2}{p^2} \cdot \frac{D_\delta l'}{\delta_{ef}}, \quad (2)$$

where  $\mu_0 = 4\pi \cdot 10^{-7}$  H/m is the vacuum permeability,  $m$  is the number of phases,  $N_{ef}$  is the number of effective turns,  $p$  is the number of pole pares,  $D_\delta$  is the diameter in the middle of the air gap,  $l'$  is the equivalent stator length,  $\delta_{ef}$  is the effective air gap, taking into account PM and iron length along the flux line. If magnetic circuit saturates too much during the short circuit, the increase of the effective air gap and decrease of the magnetizing inductance can lead to the unfavorable high value of the short circuit current. This must be taken into account especially during design of FTEM.

Stator leakage inductance can be divided into several components:

$$L_{s\sigma} = L_{\delta} + L_u + L_{tt} + L_{ew} + L_{sq}, \quad (3)$$

where  $L_{\delta}$  is the air gap leakage inductance,  $L_u$  is slot leakage inductance,  $L_{tt}$  is tooth tip leakage inductance,  $L_{ew}$  is end winding leakage inductance and  $L_{sq}$  is skew leakage inductance. Rotor leakage inductances depend on the design of the rotor of the machine and must be considered individually. While increase in leakage inductance can limit short circuit current, it can also contribute to the decrease of power factor and efficiency of the machine. Low leakage components of the synchronous inductance allow to obtain larger torque from the machine. In this regard tooth-coil windings can offer another advantage in e.g. limiting of the end winding leakage inductance, by decreasing of end winding length and its wrapping closely to the electrical sheet stack. This is especially relevant for the machines with the short stacks, since their end-winding leakage inductance contributes considerably to the total inductance. Finally, some applications imply long-term operation during a fault, which means that the fault state leakage inductance can be considered as a new rated leakage inductance.

#### 4. CONCLUSION

This paper summarized some achieved results in the development of FTEMs and provided several important aspects of their design listed in the literature, such as choice of the number of phases, slots and poles, special stator and rotor designs, usage of PMs. It was shown that diversity of applications implies unique approaches to the design of FTEMs, in order to fulfill all the requirements to the fault tolerance, torque density and operational modes of a particular drive system. Among different variables defining design of the machine in the beginning, especially careful attention should be paid to the choice of its air gap diameter. For the purpose of higher torque density and fault tolerant capability one approach can be to choose the longest air gap diameter possible at the first stages of the design process, keeping active length of the machine and its end winding length within reasonable limits. This measure allows to choose between broader number of stator slots, defining their shape in the most optimal way and potentially apply broader number of winding arrangements. Finally, it is a task for the designer to find an optimal balance between fault tolerant capabilities of the machine and its torque density, required for the given application.

#### ACKNOWLEDGMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports under institutional support and BUT specific research programme (project No. FEKT-S-20-6379).

#### REFERENCES

- [1] F. R. Ismagilov et al., "Design of a Six-Phase Fault-Tolerant Electric Motor for an Aircraft Fuel Pump," 2019 Fourteenth International Conference on Ecological Vehicles and Renewable Energies (EVER), 2019, pp. 1-6, [doi: 10.1109/EVER.2019.8813531](https://doi.org/10.1109/EVER.2019.8813531).
- [2] A. G. Jack, B. C. Mecrow and J. A. Haylock, "A comparative study of permanent magnet and switched reluctance motors for high-performance fault-tolerant applications," in IEEE Transactions on Industry Applications, vol. 32, no. 4, pp. 889-895, July-Aug. 1996, [doi: 10.1109/28.511646](https://doi.org/10.1109/28.511646).
- [3] A. M. EL-Refaie, "Fractional-Slot Concentrated-Windings Synchronous Permanent Magnet Machines: Opportunities and Challenges," in IEEE Transactions on Industrial Electronics, vol. 57, no. 1, pp. 107-121, Jan. 2010, [doi: 10.1109/TIE.2009.2030211](https://doi.org/10.1109/TIE.2009.2030211).
- [4] W. Zhao, L. Xu and G. Liu, "Overview of permanent-magnet fault-tolerant machines: Topology and design," in CES Transactions on Electrical Machines and Systems, vol. 2, no. 1, pp. 51-64, March 2018, [doi: 10.23919/TEMS.2018.8326451](https://doi.org/10.23919/TEMS.2018.8326451).

- [5] B. C. Mecrow et al., "Design and testing of a 4 phase fault tolerant permanent magnet machine for an engine fuel pump," IEEE International Electric Machines and Drives Conference, 2003. IEMDC'03., 2003, pp. 1301-1307 vol.2, [doi: 10.1109/IEMDC.2003.1210407](https://doi.org/10.1109/IEMDC.2003.1210407).
- [6] J. Chai, J. Wang, K. Atallah and D. Howe, "Performance Comparison and Winding Fault Detection of Duplex 2-Phase and 3-Phase Fault-Tolerant Permanent Magnet Brushless Machines," 2007 IEEE Industry Applications Annual Meeting, 2007, pp. 566-572, [doi: 10.1109/07IAS.2007.91](https://doi.org/10.1109/07IAS.2007.91).
- [7] M. R. Shah, A. M. EL-Refaie and K. Sivasubramaniam, "Analysis of turn-to-turn faults in surface PM machines with multi-layer fractional-slot concentrated windings," 2008 18th International Conference on Electrical Machines, 2008, pp. 1-4, [doi: 10.1109/ICELMACH.2008.4799990](https://doi.org/10.1109/ICELMACH.2008.4799990).
- [8] Q. Chen, G. Liu, L. Sun, Y. Jiang and J. Yang, "Comparison of five topologies rotor permanent magnet motors with improved fault-tolerance," 2013 IEEE International Symposium on Industrial Electronics, 2013, pp. 1-5, [doi: 10.1109/ISIE.2013.6563819](https://doi.org/10.1109/ISIE.2013.6563819).
- [9] A. M. EL-Refaie, M. R. Shah and K. Huh, "High power-density fault-tolerant PM generator for safety critical applications," 2013 International Electric Machines & Drives Conference, 2013, pp. 30-39, [doi: 10.1109/IEMDC.2013.6556125](https://doi.org/10.1109/IEMDC.2013.6556125).
- [10] P. Zheng, Y. Sui, Z. Fu, P. Tang, F. Wu and P. Wang, "Investigation of a five-phase 20-slot/18-pole PMSM for electric vehicles," 2014 17th International Conference on Electrical Machines and Systems (ICEMS), 2014, pp. 1168-1172, [doi: 10.1109/ICEMS.2014.7013664](https://doi.org/10.1109/ICEMS.2014.7013664).
- [11] H. Zhang, O. Wallmark and M. Leksell, "An iterative FEA-based approach for the design of fault-tolerant IPM-FSCW machines," 2015 17th European Conference on Power Electronics and Applications (EPE'15 ECCE-Europe), 2015, pp. 1-7, [doi: 10.1109/EPE.2015.7309335](https://doi.org/10.1109/EPE.2015.7309335).
- [12] M. Alnajjar and D. Gerling, "Six-phase electrically excited synchronous generator for More Electric Aircraft," 2016 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), 2016, pp. 7-13, [doi: 10.1109/SPEEDAM.2016.7525938](https://doi.org/10.1109/SPEEDAM.2016.7525938).
- [13] J. Zheng, W. Zhao, J. Ji, J. Zhu, C. Gu and S. Zhu, "Design to Reduce Rotor Losses in Fault-Tolerant Permanent-Magnet Machines," in IEEE Transactions on Industrial Electronics, vol. 65, no. 11, pp. 8476-8487, Nov. 2018, [doi: 10.1109/TIE.2018.2807363](https://doi.org/10.1109/TIE.2018.2807363).
- [14] Y. Chen and B. Liu, "Design and Analysis of a Five-Phase Fault-Tolerant Permanent Magnet Synchronous Motor for Aerospace Starter-Generator System," in IEEE Access, vol. 7, pp. 135040-135049, 2019, [doi: 10.1109/ACCESS.2019.2941447](https://doi.org/10.1109/ACCESS.2019.2941447).
- [15] G. H. Bayazit, M. Uğur and O. Keysan, "Fault Tolerance Capabilities of Three, Four and Six-Phase Configurations of a 24 Slot Modular PMSM," 2019 IEEE 13th International Conference on Power Electronics and Drive Systems (PEDS), 2019, pp. 1-6, [doi: 10.1109/PEDS44367.2019.8998851](https://doi.org/10.1109/PEDS44367.2019.8998851).
- [16] E. Pérez et al., "Comparison of Optimized Fault-Tolerant Modular Stator Machines with U-shape and H-shape Core Structure," 2021 IEEE Energy Conversion Congress and Exposition (ECCE), 2021, pp. 4254-4259, [doi: 10.1109/ECCE47101.2021.9596020](https://doi.org/10.1109/ECCE47101.2021.9596020).
- [17] N. Bianchi, M. D. Pre, G. Grezzani and S. Bolognani, "Design considerations on fractional-slot fault-tolerant synchronous motors," IEEE International Conference on Electric Machines and Drives, 2005., 2005, pp. 902-909, [doi: 10.1109/IEMDC.2005.195829](https://doi.org/10.1109/IEMDC.2005.195829).
- [18] M. Barcaro, N. Bianchi and F. Magnussen, "Faulty Operations of a PM Fractional-Slot Machine With a Dual Three-Phase Winding," in IEEE Transactions on Industrial Electronics, vol. 58, no. 9, pp. 3825-3832, Sept. 2011, [doi: 10.1109/TIE.2010.2087300](https://doi.org/10.1109/TIE.2010.2087300).
- [19] F. R. Ismagilov, V. Y. Vavilov, V. I. Bekuzin, V. V. Ayguzina and D. Y. Permin, "Analysis of Stator Cooling Methods of Fault-Tolerant Electric Machines," 2020 Fifteenth International Conference on Ecological Vehicles and Renewable Energies (EVER), 2020, pp. 1-6, [doi: 10.1109/EVER48776.2020.9243120](https://doi.org/10.1109/EVER48776.2020.9243120).



# Design of a High-speed Solid-rotor Asynchronous Generator for an Expansion Turbine

D. Pribulla and O. Vitek

Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech republic

E-mail: [xpribu02@vut.cz](mailto:xpribu02@vut.cz), [viteko@vut.cz](mailto:viteko@vut.cz)

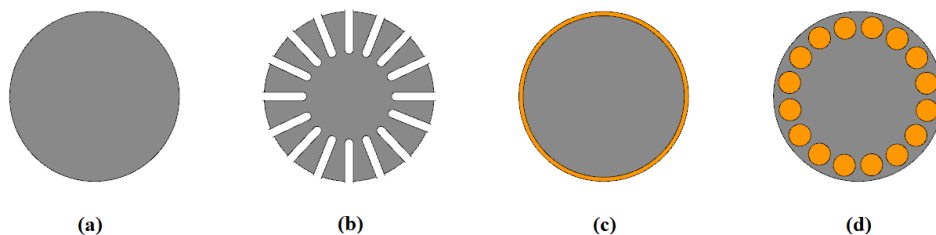
**Abstract**—In this paper, an electromagnetic design of a slitted solid-rotor high-speed generator is presented. Recommendations for an analytical preliminary design were summarized. The design was further optimized using 2-D finite element method. The power of the machine is 6.15 kW at 150 000 rpm and 2% slip. An efficiency of 90% was reached, despite of high eddy losses in the machine rotor. Our design may prove to be a cost-effective alternative to the permanent magnet synchronous machine in the cryogenics application after the thermal and mechanical analysis.

**Keywords**—Solid rotor, high-speed, slitted, design, FEA

## 1. INTRODUCTION

Cryogenics and gas liquefaction systems are well established fields of application for high-speed electrical machines. They involve compressors and turbines that require very high speeds to operate effectively. One of the critical components is an expansion turbine, where the energy of the gas is converted into the mechanical work. Coupling the turbine with electric generator enables to remove this energy from the system, thus preventing overheating. Increased efficiency of the process is another benefit. The design specification for this application was  $150\,000\text{ min}^{-1}$  and 10 kW. Based on the overview in [1] and the parameter  $\text{rpm} \cdot \sqrt{\text{kW}}$ , synchronous permanent magnet machine (PMSM) was previously selected and designed in [2]. The machine has very high efficiency, and meets the requirements, but the manufacturing of the machine rotor and sleeve is extremely complicated and expensive. Thus the second suitable option, an induction machine (IM) was chosen.

There are four main rotor variations (Figure 1) to be considered, when designing IM for a high speed application, namely solid steel rotor, solid steel rotor with axial slits, solid steel rotor coated with copper and squirrel cage rotor.



**Figure 1:** Common high-speed solid-rotor cross sections. (a) Smooth. (b) Slitted. (c) Coated. (d) Caged.

They are sorted with increasing electromagnetic performance with solid steel rotor gaining the worst efficiency and power factor. The mechanical properties go in an inverse order with solid steel rotor having the highest rigidity and resistance to centrifugal forces, thus achieving the highest circumferential speed. In this application a slitted solid steel rotor type was selected. It's electromagnetic performance is better than that of the solid rotor. The slits act as a barrier for the eddy currents flowing in the tangential direction and help the magnetic field to penetrate deeper into the rotor. Moreover, it is still a machinable form of a single piece of material, without the need of a special manufacturing technology. The downside is a slightly lower mechanical rigidity as the slits act as stress concentrators and increase windage losses.

## 2. MACHINE DESIGN

Optimal design of IM with slitted rotor is challenging, because magnetic and electric field share the same volume. It is impossible to define integration paths and simplify the problem, like in the caged IM. Methods used for smooth rotor, utilizing multi-layer transfer matrix [3] can not be used, because of the complex geometry of the slotted rotor. Nowadays, a finite element analysis (FEA) is commonly used to evaluate and optimize the performance of the machine. Solid rotor IM design was analyzed in detail in [4] and [5] and further optimized in [6]. Their design recommendations were considered in the preliminary design.

The stator of the machine was designed using the calculation process presented in [7], as a standard, caged IM. Relatively low efficiency of the machine is expected, so the air gap power of the stator was increased to 12 kW to be able to cover the losses in rotor. In case of a solid rotor machine, the air gap length has to be significantly larger, in comparison with conventional squirrel cage rotor. Actual value is a compromise between high magnetization current and the eddy losses at the surface of the rotor. The stator bore was fixed at 37 mm due to compatibility with previous design and mechanical properties of the rotor. Because of the small bore diameter there are 12 stator slots to obtain reasonable tooth width. Long, narrow stator teeth would lack mechanical rigidity and pose manufacturing problems. Machine was designed as two-pole, in order to minimize iron losses in the stator sheet stack. With 150 000 rpm, the no-load supply frequency is 2500 Hz, thus special care must be taken when selecting the material. The thickness of the sheet must be kept below the depth of penetration of the changing magnetic field. Solid rotor IM are known for high eddy-current losses caused by higher-order harmonics in the rotor. They must be suppressed by all means. Because of this, a two-layer winding with 5/6 winding pitch was used. This has been proven in [4] to have the lowest higher-order harmonics in the air gap. The permanence harmonics, caused by stator slotting are suppressed by increasing the air gap length and adjusting the slot opening width.

There are only few rotor dimensions that need to be considered. According to [5], the optimal electromagnetic properties are reached when depth of the slit is 60% of the rotor diameter, but in practice it has to be limited to about 50% due to the mechanical constraints. The width of the slot should be as thin as possible, but during the manufacturing the thickness of the available machining tools is the limit. Tool thickness generally raises with its length and achievable depth of the slit. The number of the slits  $Q_r$  must be matched with number of stator slots  $Q_s$  in order to avoid vibrations, synchronous and asynchronous torques. The selection criteria can be found in [7]. To avoid synchronous harmonic torques, it is recommended that

$$Q_r < 1.25Q_s . \quad (1)$$

For a high speed machine with fluid bearing, mechanical vibration is the biggest concern. It was found, that odd numbers of slits have high unbalanced forces [6]. These can destroy the bearings. The slot harmonics and synchronous torques will drop out the slot numbers matching conditions (2) and (3):

$$Q_r = 6Q_s \pm 2p , \quad (2)$$

$$Q_r = 6pg \pm 2p , \quad (3)$$

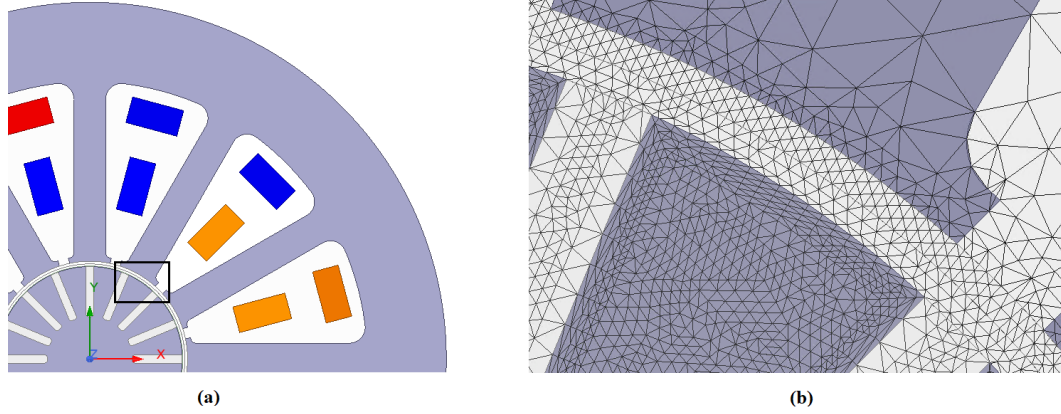
where  $g$  is any integer and  $p$  number of poles. Due to small number of stator slots, all but 16 slots were disqualified, but 14 and 18 were also taken into consideration because of reasonable geometry. The last factors to be considered are length of the end-rings and the resistivity of the material. An overview of parameters of the preliminary machine design are shown in the Table I. The rotor material was provided by the customer. Operational temperature of the rotor and the material resistivity was estimated at 200 °C in this application.

**Table I:** Preliminary electromagnetic design of the PMSM generator.

Rated voltage $U$	[V]	400	Air gap $\delta$	[mm]	1
Stator winding connection		Star	Rotor diameter $D_r$	[mm]	35
Number of coil turns $N$		28	Rotor length $l_r$	[mm]	50
Single wire cross section $S_{cs}$	[mm <sup>2</sup> ]	3	Slot width $W_{sl}$	[mm]	1.5
Stator material		NO 10	Rotor material		Carbon steel
Stator core length $l$	[mm]	45	Rotor material resistivity	[MS/m]	3.5

### 3. DESIGN OPTIMIZATION

In order to optimize the number and depth of slots and to evaluate the performance of the final design, a parametric 2-D FEA model of the machine was created in Ansys Maxwell software (Figure 2a). The simulation was transient with eddy current calculation activated in the rotor volume. The rotor speed was constant and machine windings were supplied with a variable frequency sine voltage sources. The time step was set to  $1\mu s$ , which is enough to capture the effect of at least first eight space harmonics in the air gap. The mesh must be sufficiently refined in the air gap, and especially on the surface of the rotor, since the higher-order harmonics have low depth of penetration. The detail of the mesh is shown in Figure 2b.

**Figure 2:** (a) Machine model (b) Mesh detail

The effect of end rings can not be modeled with 2-D model. The resistivity of rotor material is increased by a coefficient  $k_{Russel}$  to take the resistance of the end rings into account.

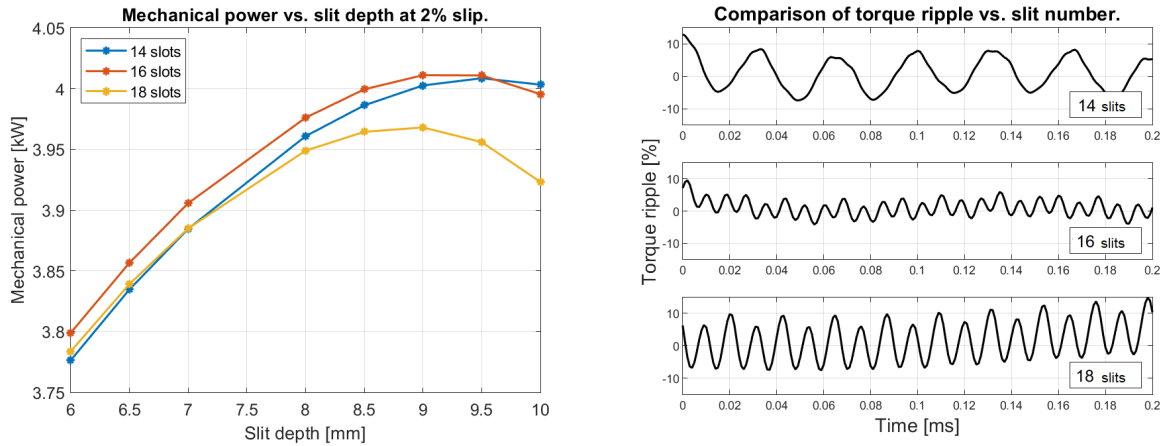
$$k_{Russel} = 1 - \frac{2\tau_p}{\pi l_r} \tanh\left(\frac{\pi l_r}{2\tau_p}\right), \quad (4)$$

where  $\tau_p$  is the pole pitch. The rotor resistivity was corrected to 1.29 MS/m.

### 4. RESULTS

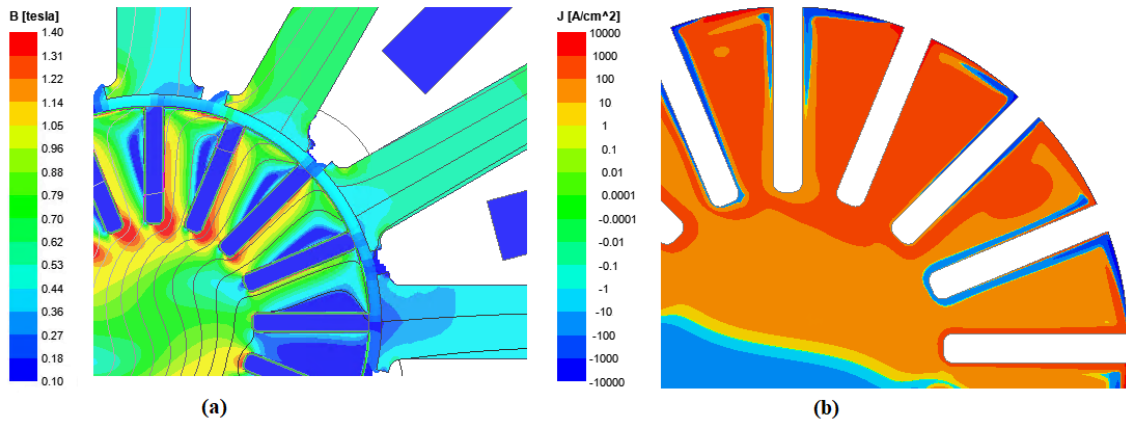
A series of model variations with slit depth ranging from 6 - 10 mm (34 - 57 %) for each number of slots were simulated. Slip of the machine was fixed to 2%. The optimization criteria was maximum electromagnetic torque and mechanical power, that can be removed from the system. Results of the optimization are shown in Figure 3. Maximum power was removed with 16 slit and 9 mm slit depth. The performance of the 14 slit rotor is comparable and could be better choice from mechanical point of view. The 18 slit rotor performs much worse and achieves the peak power at lower slit depth. This might be due to magnetic saturation at the root of the slit. Relatively wide slits take up too much from the rotor cross-section. Another selection aspect was torque ripple of the rotor. The findings are in accordance with the recommendations (1) – (3). The 16 slit rotor has about a half ripple compared to other two options. Thus, the 6 slit and 9 mm depth combination was chosen as the final design.

Magnetic flux density distribution of the selected design is shown in Figure 4a. Whole magnetic



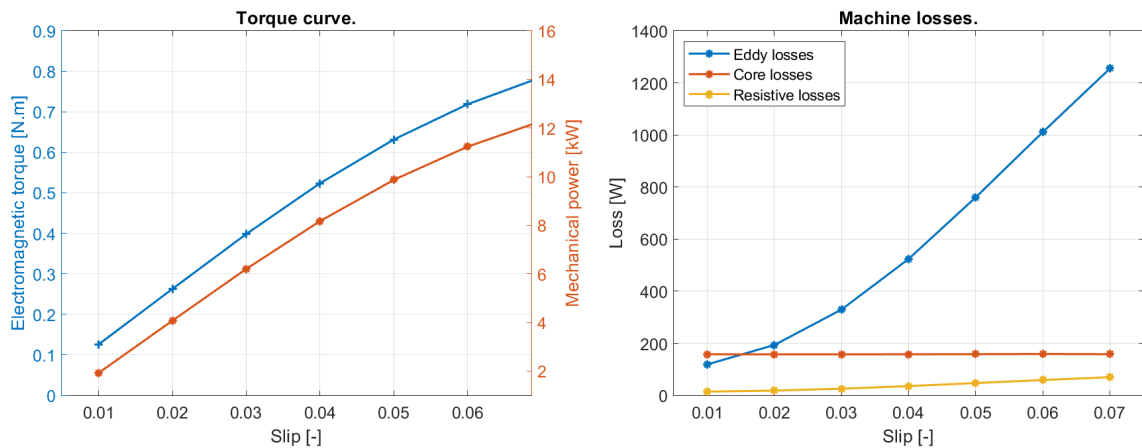
**Figure 3:** Optimization results at 2% slip: (a) Power optimization (b) Torque ripple

circuit is designed to have low saturation (below 0.8 T) to limit the losses. Figure 4b shows a very strong localization of the eddy currents on the rotor surface and explains the need of mesh refinement in this area.



**Figure 4:** Final design, fields distribution: (a) magnetic flux density (b) eddy current density

The torque and losses of the designed machine as a function of slip are shown in Figure 5. The machine would achieve the specified power at 5% slip, but the rotor losses seem excessive. Because of this the nominal slip was set to only 3%.



**Figure 5:** Characteristics of the selected machine design.

Overview of performance characteristics at this set-point are listed in Table II. Analytical equations assuming smooth rotor surface in [7] were used to calculate the windage losses. Based on the findings in [8], the roughness coefficient was raised to 2.5 to consider the effect of the slits.

**Table II:** Rated performance parameters of the designed induction generator at 3% slip.

RMS phase current $I_s$	[A]	18.7
Rotor eddy current loss	[W]	329
Windage loss $P_p$	[W]	98
Iron core loss $P_{Fe}$	[W]	158
Copper loss $P_{Cu}$	[W]	26
Efficiency $\eta$	[%]	90.5
Mechanical power $P_{max}$	[W]	6 150

## 5. CONCLUSION

An induction machine with slitted solid rotor may be a suitable and economical solution for an expansion turbine application. It was shown, that the preliminary, analytical and guessed design was very close to the optimized result. From the electromagnetic point of view, the machine is capable of removing specified mechanical power from the cryogenic circuit with a very high efficiency. The main culprit, which may reduce the performance of the machine, is that the vast majority of losses is dissipated into the rotor. This could be solved by increasing air gap length as a tradeoff between the rotor losses and magnetization current and copper losses. Only the thermal analysis and prototype testing will show, if the liquefied gas can provide sufficient rotor cooling capacity. It may be enhanced by a flow analysis and optimal skewing of the rotor slits. A mechanical analysis is necessary to validate the rotor material selection, calculate sufficient rotor teeth strength and critical speed of the whole rotating assembly.

## Acknowledgment

This research work has been carried out in the Center for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports under institutional support and BUT specific research program (project No. FEKT-S-20-6379).

## REFERENCES

- [1] D. Gerada, A. Mebarki, N. L. Brown, C. Gerada, A. Cavagnino, and A. Boglietti, "High-Speed Electrical Machines: Technologies, Trends, and Developments", in *IEEE Transactions on Industrial Electronics*, 2014, vol. 61, no. 6, pp. 2946-2959.
- [2] D. Pribulla, "Design Of A High-Speed Generator For A Helium Expansion Turbine", in *Proceedings I of the 27st Conference STUDENT EEICT 2021: General papers*, 2021.
- [3] J. Pyrhönen, "The high speed induction motor: Calculating the effects of solid rotor material on machine characteristics", Dissertation thesis, Lappeenranta, 1991.
- [4] J. Hupponen, "High-Speed Solid-Rotor Induction Machine: Electromagnetic Calculation and Design", Dissertation thesis, Lappeenranta, 2004.
- [5] T. Aho, "Electromagnetic Design of a Solid Steel Rotor Motor for Demanding Operation Environments", Dissertation thesis, Lappeenranta, 2007.
- [6] N. Uzhegov, "High-Speed Solid-Rotor Induction Machine: Calculation program", Dissertation thesis, Lappeenranta, 2012.
- [7] J. Pyrhönen, T. Jokinen, and V. Hrabovcová, *Design of rotating electrical machines*, 2nd edition. Chichester: Wiley, 2014.
- [8] J. Saari, "Thermal Analysis of High-speed Induction Machines", Dissertation thesis, Helsinki, 1998.

# MATHEMATICAL MODEL OF RCCB SENSITIVE TO PULSATING DIRECT CURRENTS

J. Kaser<sup>1</sup> and J. Valenta<sup>1</sup>

<sup>1</sup>Doctoral Degree Programme, Power Electronics and Energy, FEED BUT, Brno University of Technology,  
Czech Republic

E-mail: [xkaser00@vut.cz](mailto:xkaser00@vut.cz), [valentaj@vutbr.cz](mailto:valentaj@vutbr.cz)

**Abstract**—This study is focused on a residual-current circuit breaker (RCCB) and its eventual operation malfunction due to direct current components. This problem is associated with magnetic phenomena inside its transformer core. These phenomena are described in detail in the magnetic equations used in a mathematical model created in Simulink. The model incorporates the comparison of individual transformer B/H loop and verifies its reaction to the various residual currents and the correct operation of residual current circuit breaker to achieve results.

**Keywords**— Residual-current circuit breaker, Mathematical model, Simulink program, Magnetic phenomena, Pulsating direct current

## 1. INTRODUCTION

The electrical protections ensure a fundamental safety in the power electrical engineering and installations and guarantees a fast interruption in a required time. It must fulfil the conditions of correct protection (selectivity, maximum line resistance, sensitivity, speed and next) against an overload, a short-circuit current and a thermal or force stress. Furthermore, a minimum risk to electrical shock is very important and essential. Therefore, it is necessary to detect and interrupt a direct and indirect contact in a few milliseconds. Basically, the residual-current device reacts on this fault. However, there are some problems with the tripping as shown reference 7. In this case, this protection may not detect a pulsating or a smooth direct current from the rectifiers (one way, six-pulse and next). The main issue is that the transformer core of device is saturated. Thus, it is useful to choose the correct type of residual-current breaker as described in chapter 2.

## 2. RESIDUAL-CURRENT DEVICE

According to references 1, 2 and 3, the residual-current circuit breaker (RCCB) or residual-current device (RCD) is a protection to detect a leakage (residual) fault current and ensures safety against electrical shocks caused by the indirect and direct contacts in a grounded installation. As a result, the RCCB indicates all earth fault currents, including the earth short-circuit currents.

The RCCB disconnects the circuit whenever the current exceeds a rated residual sensitivity. This protection contains a sum current transformer that measures and compares the instantaneous vector values of currents passing through the working wires (L and N) for the closure of loop. As already mentioned, the RCCB is connected without earth wire (PE). In the normal conditions, the sum of the current vector values is zero. When an earth fault occurs and the rated residual current sensitivity is exceeded, this device interrupts a circuit. The induced secondary current creates the magnetic field of coil of the sensitive relay which effects against the magnetic field of permanent magnet. This magnet drops out and then disconnects the mechanical contacts.

The type of transformer core and its sensitivity are important for next results in chapters 4.

**AC** - Residual sinusoidal alternating current.

**A** - Residual sinusoidal alternating current, pulsating direct current and smooth current to 6 mA.

**F** - Same as type A but it utilizes for frequency up to 1 kHz.

**B** - Residual sinusoidal alternating current, pulsating direct current and smooth current over 10 mA.



### 3. MAGNETIC TERMS AND EQUATIONS

The model of transformer core of residual current device utilizes the magnetic phenomena and their basic equations to create the loops for complete closure. The necessary parameters are magnetic flux  $\phi$  of core and magnetic flux linkage  $\psi$ , inductance  $L$  and mutual inductance  $M$ , magnetic flux density  $B$  and magnetizing field strength  $H$  to create B-H loop, magnetization current  $i_0$  to establish this magnetic flux  $\phi$ . The basic magnetic explanation and the knowledge are obtained from references 4, 5 and 6.

In the ferromagnetic core, the magnetic flux  $\phi$  is the product of the magnetic flux density  $B$  and the perpendicular area of core  $S$  that it penetrates. The magnetic flux is the measure of total number of induction lines passing through this area.

$$\Phi = B S \cos\alpha = B S \quad (3.1)$$

The coil of transformer has several turns, so this flux is a magnetic flux linkage and is marked  $\psi$ . Furthermore, this flux  $\psi$  is expressed as the product of inductance  $L$  and the total current flowing  $I$  in the circuit.

$$\Psi = \Phi N = L I \quad (3.2)$$

The magnetic flux density  $B$  is expressed as the product of the magnetic field strength of core  $H$ , the vacuum permeability  $\mu_0$  and material permeability  $\mu_r$  (ferromagnetic materials). The permeability is the measure of magnetization that the material obtains in reaction to the magnetic field in the loop. The B-H loop determines the core type mainly according to remanence  $B_r$ .

$$B = \mu H = \mu_0 \mu_r H \quad (3.3)$$

In the electromagnetism, an inductance  $L$  of coil is 1 H where induces a voltage 1 V if a current change itself 1 A per second. The inductance expresses dependence between the induced voltage and current change. In this equation, a core dimension as a length  $l$  and area  $S$ , permeability  $\mu$  and the number of turns  $N$  are important for a whole expression.

$$L = \frac{N^2 \mu_0 \mu_r S}{l} \quad (3.4)$$

The mutual inductance  $M$  expresses the proportionality of induced voltage in the first coil that produces the magnetic field in the second coil.

$$M = \frac{N_1 N_2 \mu_0 \mu_r S}{l} \quad (3.5)$$

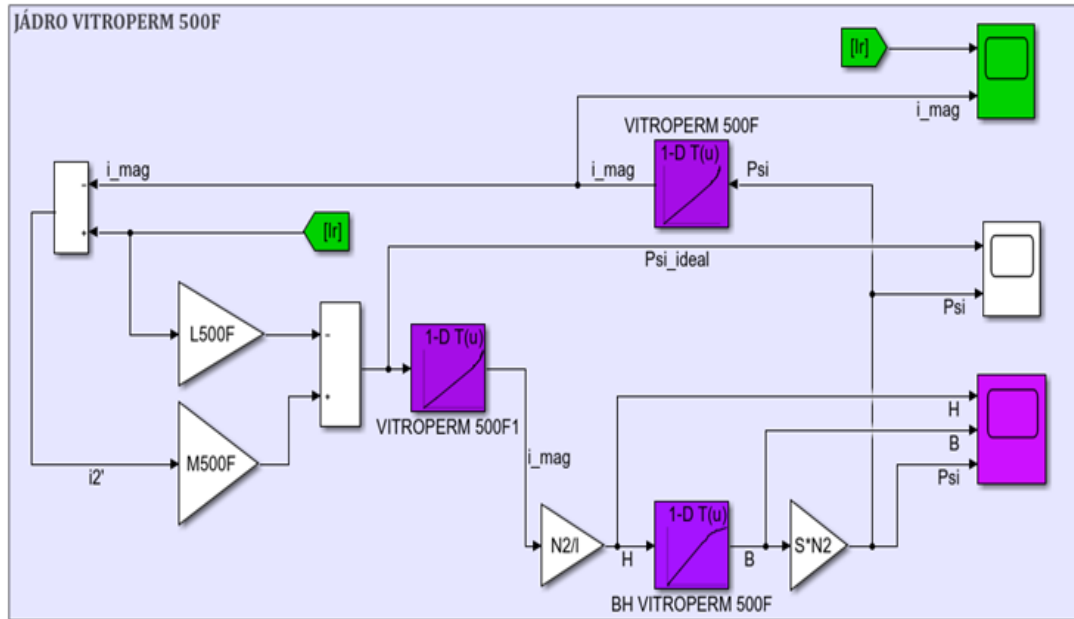
### 4. MATHEMATICAL MODEL OF RCCB IN SIMULINK

The research has 2 parts. The first is a magnetic model and second is a fault current loop with the pulsating direct current 30 mA controlled by the thyristors in the  $0^\circ$ ,  $90^\circ$  and  $135^\circ$ .

This magnetic model consists of the fundamental equations from chapter 3 to create the final loop of transformer core of RCD. The loop includes the magnetic flux linkage  $\Psi$  from equation 3.2. and the current carrying in the turns ( $i_1$  and  $i_2$ ) and through the core ( $i_0$ ). The current of secondary turn  $i_2$  transfer to the primary side to correct results. The mutual inductance  $M$  create the magnetic field from induced side as shown in 4.1. and a stray inductance is neglected due to toroid core.

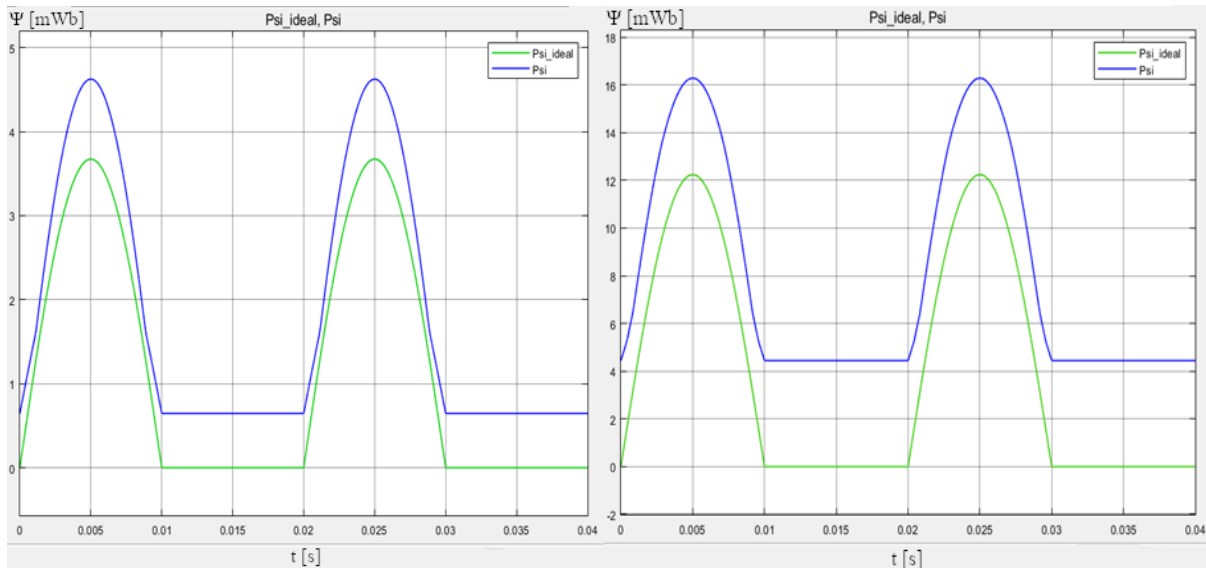
$$\Psi = L i_1 + M i_2' \quad (4.1)$$

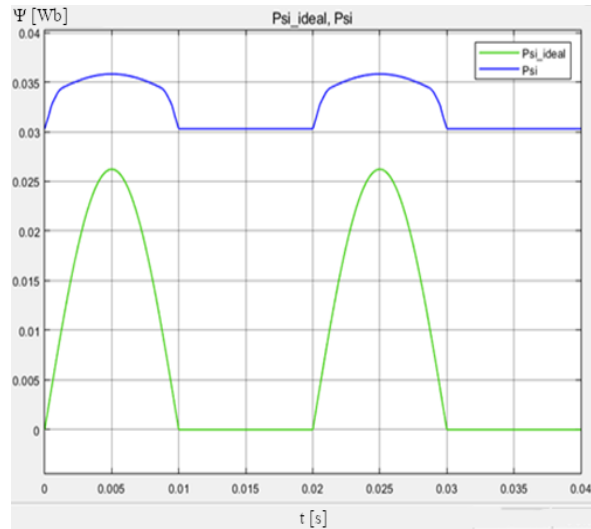
$$i_0 = i_1 - i_2' \quad (4.2)$$



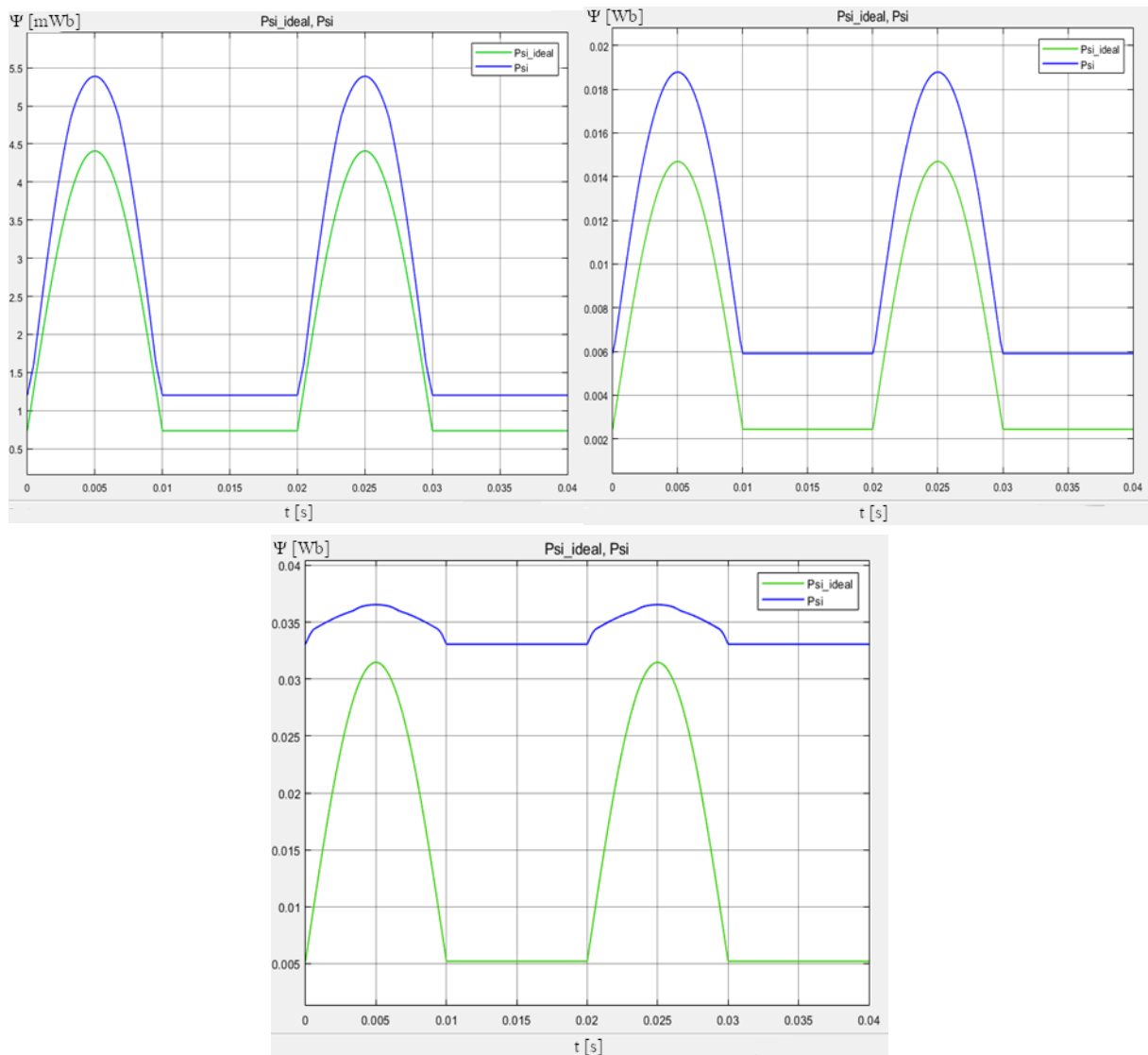
**Figure 1:** Transformer core of Vitroperm 500F

The research verifies three different materials for transformer core. Each material has a specific B-H loop with magnetic features. In the Figure 1 is Vitroperm 500F, other loops of the materials look like same. The green arrow ( $I_r$ ) is the residual pulsating direct current. The purple squares show the B-H loop and  $\Psi$ - $i_0$  characteristics of core. The results draw the comparison of magnetic flux linkage  $\Psi$  of all 3 materials and are measured by the oscilloscopes. The simulations are for pulsating direct current 30 mA controlled in  $0^\circ$  in the Figure 2 and even with smooth direct current 6 mA in Figure 3.





**Figure 2:** Simulation of magnetic flux  $\Psi$  a) Vitroperm 500F b) Vitroperm 800F c) Material C



**Figure 3.** Simulation of magnetic flux  $\Psi$  with smooth direct current 6 mA a) Vitroperm 500F b) Vitroperm 800F c) Material C

## 5. RESULTS AND DISCUSSION

The residual-current circuit breaker is a fundamental protection, and correct type of core and its sensitive must fulfil the conditions. The residual pulsating direct current 30 mA flow through the mathematical model and the obtained results in figure 2a and 2b show the successful simulations of Vitroperm 500F and 800F due to the low remanence ( $B_r$ ) their B-H loops. The remanence of Vitroperm 500F is the lowest of all materials (0,008 T), so the circuit have no significant saturation of core (each Psi blue course). The similar simulation has the Vitroperm 800F. The remanence is raised (0,055 T) but an overall course is not different from ideal as shown in figure 3b. It is only shifted along the y-axis, so the inductive voltage of secondary winding is high enough.

The third material C have the highest remanence (around 0,37 T) and the simulation is significant deformed as shown figure 3c (Psi blue course) and the saturation occurs. The magnetic flux density B comes up to 0,45 T that exceeds the linear values of B-H loop. The magnetic flux linkage is four times less than ideal and the induced voltage would not be sufficient. Therefore, this material utilizes only for type AC.

In the next analysis in Figure 3, the smooth direct current 6 mA raises a saturation probability. However, the Vitroperm 500F and 800F withstood these conditions and could handle even higher smooth direct current. So, these materials fulfil the standard conditions for type A. The material C is even more deformed (Figure 3c) and there is high permanent magnetic field.

In this research, the model of transformer core is simplified. The magnetic flux linkage must be derived for the correct form of induced voltage. However, this operation represents a high numerical error for the Simulink. So, a load is not in this model and the current transformer is short-circuited. The load and induced voltage would change slightly the secondary current  $i'_2$  and magnetization current  $i_0$ . The next simplification is the neglect of stray inductance as shown in the equation 4.1 and winding resistance for its minimum value. On the secondary side is a toroid core (number of turns is 795), so the stray inductance would be minimum, too.

As shown in the figure 2, the residual pulsating direct current is controlled in  $0^\circ$ . However, the thyristors of fault loop can control current in  $90^\circ$  and  $135^\circ$  according to the norm and this loop even produce the smooth direct current to increase the saturation of core. The next courses as the magnetic flux density B and this controlling in the next steps are not contain in this article.

## ACKNOWLEDGMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports under institutional support and BUT specific research program (project No. FEKT-S-20-6379).

## REFERENCES

- [1] Měření při revizích elektrických instalací [online]. ILLKO, 2020 [cit. 2019-12-30]. Source: [https://www.illko.cz/images/dokumenty/mereni\\_proudovych\\_chranicu.pdf](https://www.illko.cz/images/dokumenty/mereni_proudovych_chranicu.pdf)
- [2] RCCB/RCBO function and types. ElectGo [online]. Indonesia: Copyright ElectGo, 2018 [cit. 2021-11-16]. Source: <https://www.electgo.com/rccb-2/>
- [3] Residual Current Devices – RCCBs. ABB [online]. Stockholm: Copyright ABB, 2021 [cit. 2021-11-16]. Source: <https://new.abb.com/low-voltage/products/system-pro-m/residual-current-devices/rccb>
- [4] Magnetism. Physics of Non-destructive Evaluation [online]. USA: Iowa State University, 2021 [cit. 2021-11-16]. Source: <https://www.nde-ed.org/Physics/Magnetism/index.xhtml>
- [5] Circuit Theory. Electrical Classroom [online]. Copyright Electrical Classroom, 2021 [cit. 2021-11-16]. Source: <https://www.electricalclassroom.com/>
- [6] Magnetic Flux Density. ScienceDirect [online]. Brno: VUT, 2021 [cit. 2021-11-16]. Source: <https://www.sciencedirect.com/topics/engineering/magnetic-flux-density>
- [7] Nuisance tripping of residual current circuit breakers. Electric Power Systems Research [online]. 2014, 106, 180-187 [cit. 2021-11-29]. ISSN 0378-7796. DOI: 10.1016/j.epsr.2013.07.020

# Influence of the magnet shape on the rotor eddy-current losses in the outer rotor permanent magnet synchronous motor

P. Klíma<sup>1</sup>, and O. Vitek<sup>1,2</sup>

<sup>1</sup>Brno University of Technology, Czech Republic

<sup>2</sup>Supervisor

E-mail: [xklima28@vut.cz](mailto:xklima28@vut.cz), [viteko@vut.cz](mailto:viteko@vut.cz)

**Abstract**—This paper is focused on the comparison of the permanent magnet's shape and magnetization direction influence on the rotor eddy-current losses in the 3 kW 30 000 rpm high-speed outer rotor permanent magnet synchronous motor. The analysis uses the results of the finite element simulations. The harmonic analysis of the air-gap flux density was done to better understand the difference between the studied magnet shapes.

**Keywords**—PMSM, eddy-current losses, rotor shape, fast Fourier transform

## 1. INTRODUCTION

The improvements in the quality of materials used in electrical machines and the development in power electronics have brought an increase in the utilization of high-speed electrical machines. Their main advantage is the high power density and the possibility to directly connect the load with the machine without the need for a gearbox. This leads to less complex drives and more miniaturized systems. [1]

The high-speed permanent magnet machines with an inner rotor need to consider the effects of the centrifugal forces on the magnets. The magnets are therefore usually placed inside a metal or carbon fiber sleeve. This complicates the manufacturability and the thermal state of the rotor. The permanent magnet machines with an outer rotor utilize the fact that the permanent magnets have high compressive strength and the rotor yoke act as the holding sleeve. The cooling of the outer rotor is also better. [2]

The high frequency of the high-speed machines brings not only a high power density but also a high loss density. The stator steel needs to be as thin as possible to reduce the eddy-current core losses. The eddy-currents also affect the stator winding and could cause a substantial increase of losses in comparison with lower frequency machines. The stator winding wire therefore needs to have a small diameter or to consist of more parallel conductors of a small diameter [3]. The permanent magnets and the rotor yoke are also subjects of the eddy-current losses. These are generated by the high-order harmonics in the air-gap. These losses increase the temperature of the magnets, which decreases their strength and could, in the worst case, cause their demagnetization.

## 2. ANALYZED MACHINE

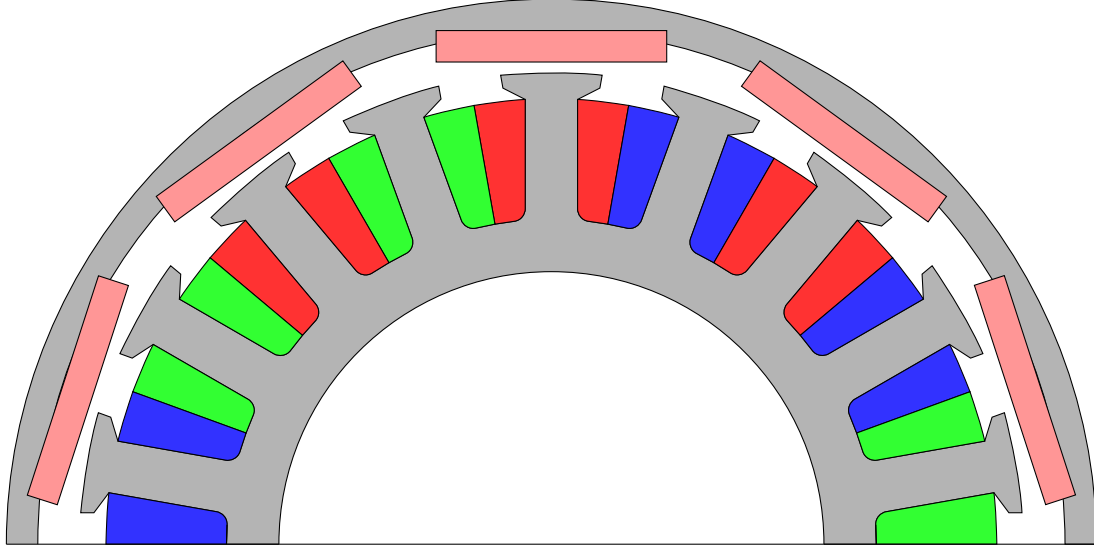
The studied machine is an outer rotor permanent magnet synchronous motor. It's main parameters are in Table I.

**Table I:** Main parameters of the motor

Name	Unit	Value
Rated speed	min <sup>-1</sup>	30 000
Rated power	W	3 000
Number of stator slots	-	18
Number of poles	-	10
Rotor outer diameter	mm	52
Active length of the motor	mm	44
Radial air-gap length	mm	0.5

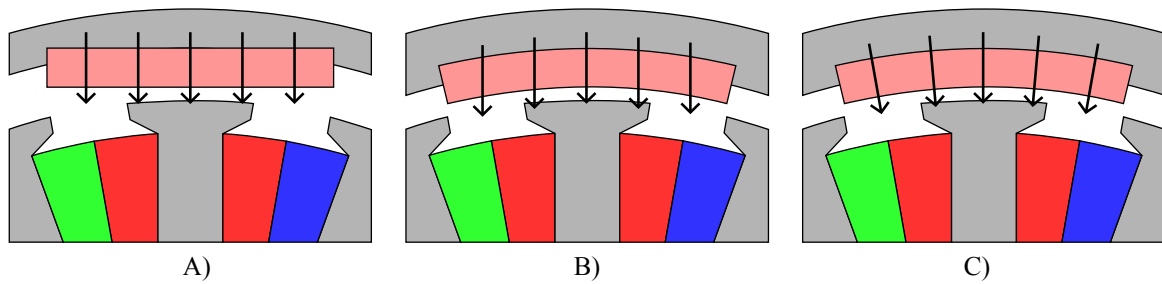
The stator steel is made from NO-20 steel to mitigate the effects of the high rated frequency on the core losses. The stator winding is made of copper wire made of parallel strands of a small diameter, therefore the eddy effects in the winding are not calculated. The rotor material is high tensile steel with good mechanical and electromagnetic properties. The neodymium surface permanent magnets are used to provide a high magnetic coercivity and a high remanent flux density.

The electromagnetic properties of the motor were calculated using the transient finite element method with the step of 1/400 of the period, which should be enough to properly calculate the eddy-current losses [4]. The model with rectangular magnets is shown in Figure 1. The model includes the air on the outer side of the rotor to properly calculate the leakage flux in that area. To produce the required power, the motor was supplied by the three-phase sine current in the most torque per ampere mode. The temperature of the magnets was set to 120 °C.



**Figure 1:** 2D motor model used in finite element method analysis.

The studied shapes and magnetization directions of the magnets are shown in Figure 2. The motor dimensions and other parameters were kept the same for the different magnet shapes. The change in the radial air-gap length ( $\delta$ ) was done by keeping the rotor dimensions the same while decreasing the outer diameter of the stator and the stator slot cross-section.



**Figure 2:** Analysed shapes and magnetization directions of the magnets. A) rectangular magnets with vertical magnetization, B) radial magnets with vertical magnetization and C) radial magnets with radial magnetization.

### 3. SIMULATION RESULTS

The value of the effective current was set for the motor to produce the 3 000 W, while the rotation speed of the rotor was kept at a constant 30 000 rpm.  $\delta = 0.5$  mm and  $\delta = 1.0$  mm was studied to compensate the increased losses for the radial magnets. But the increase of  $\delta$  decreases the stator outer diameter and the copper slot area, which causes higher current density and losses in the winding.



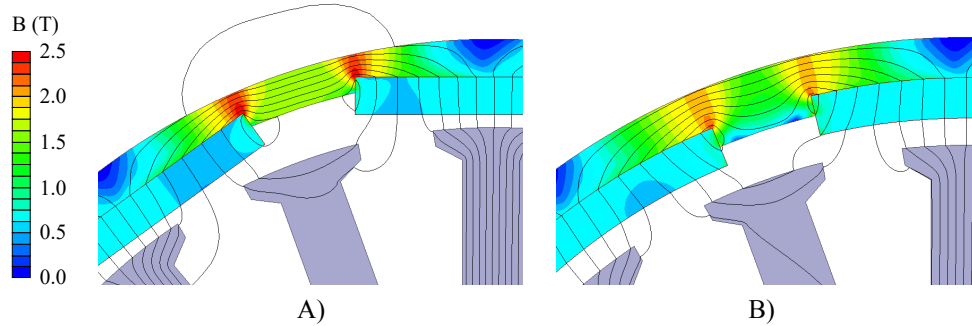
### 3.1. Motor losses and rotor state

The resulting losses are shown in Table II. The results show bigger rotor eddy-current losses for the radial magnets with both magnetization directions with  $\delta = 0.5$  mm. The main reason for this is a bigger magnetic flux resulting from lowering the rotor saturation near the end of rectangular magnets. The radial magnets don't have the rotor yoke bottleneck which is one of their geometrical advantages.

**Table II:** Current and losses in the motor for different magnet shapes, magnetization directions and  $\delta$ , output power 3 kW at 30 000 rpm

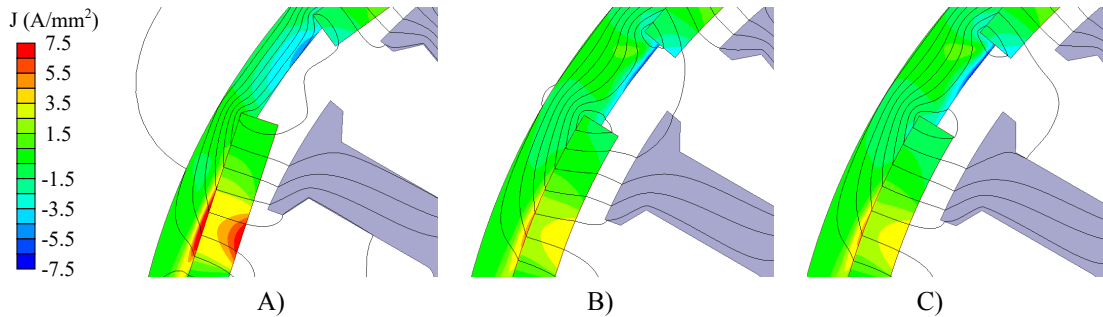
Magnet Shape	Magnetization Direction	Air-Gap Length	Current	Losses			Electromagnetic Efficiency
				Rotor Eddy	Core	Copper	
-	-	mm	A	W	W	W	%
Rectangular	Vertical	0.5	8.85	90.3	77.4	108.1	91.6
		1.0	10.10	40.3	56.3	156.1	92.2
Radial	Vertical	0.5	7.79	142.2	102.8	83.7	90.1
		1.0	8.71	58.6	73.8	114.9	92.4
	Radial	0.5	7.55	157.7	110.2	78.7	89.6
		1.0	8.33	62.2	80.1	105.1	92.4

This phenomenon can be seen in Figure 3.  $\delta = 1.0$  mm was used to effectively lower the rotor eddy-current losses, but this is done at the cost of increasing the stator copper losses. This is the case mainly because of a slightly smaller stator copper cross-section and a higher leakage flux.



**Figure 3:** Flux density distribution for A) rectangular magnets with  $\delta = 0.5$  mm and B) radial magnets with vertical magnetization with  $\delta = 1.0$  mm.

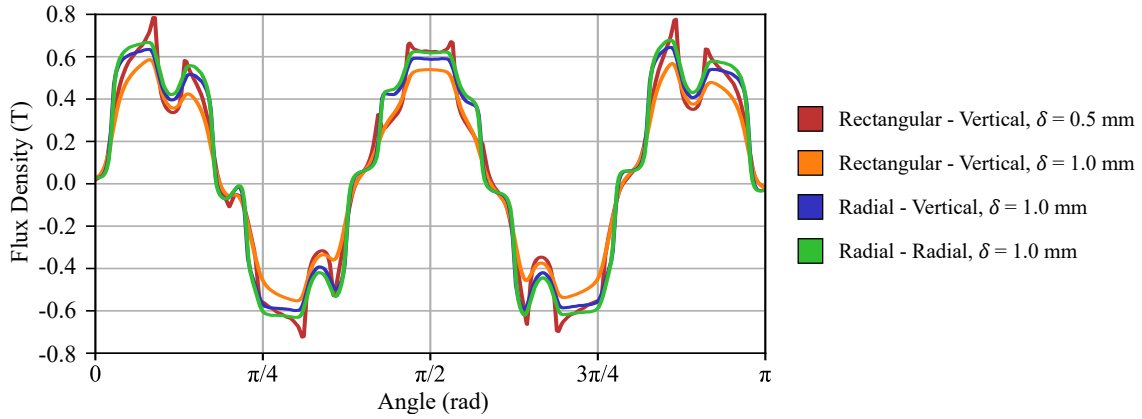
Figure 4 shows the current density in the permanent magnets and the rotor yoke. The amplitude of the eddy-currents for the rectangular magnets with  $\delta = 0.5$  mm reaches up to  $7.5 \text{ A/mm}^2$ . The lower eddy-current values are clearly visible for the radial magnets with  $\delta = 1.0$  mm, especially in the magnets.



**Figure 4:** Eddy currents in the rotor for A) rectangular magnets with  $\delta = 0.5$  mm, B) radial magnets with vertical magnetization with  $\delta = 1.0$  mm and C) radial magnets with radial magnetization with  $\delta = 1.0$  mm.

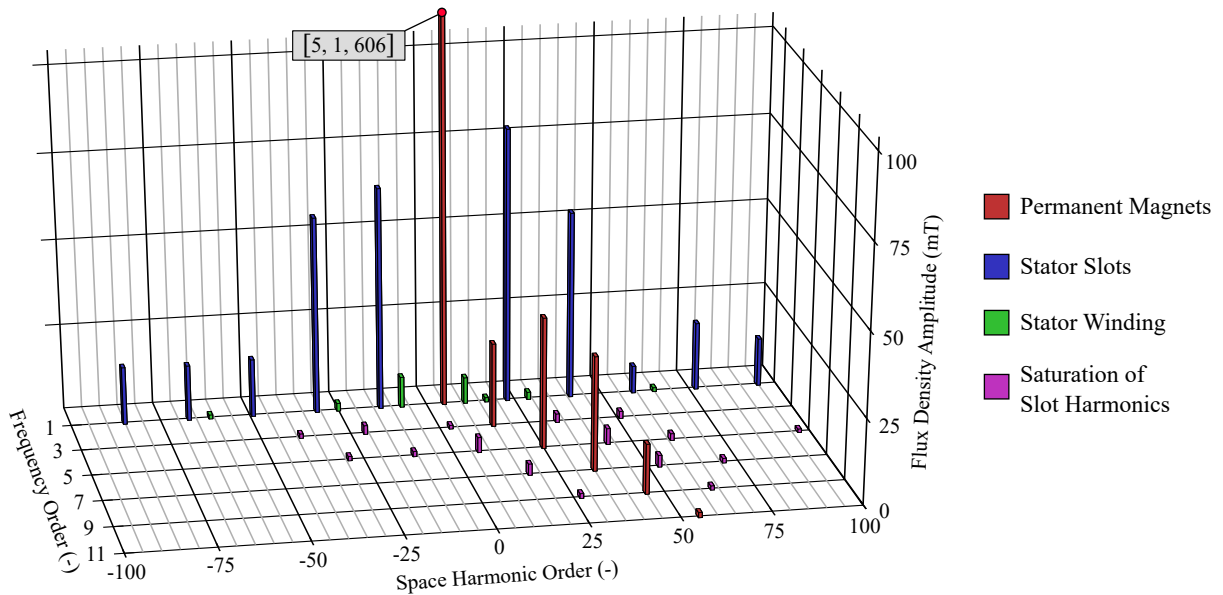
### 3.2. Harmonic analysis of air-gap flux density

To better understand the influence of the magnet shape and magnetization on the motor, the analysis of the air-gap flux density was done. In Figure 5 the radial flux density in the air-gap 0.25 mm from the rotor at the rated operation can be seen. The 0.25 mm distance from the rotor (middle of the air-gap for  $\delta = 0.5$  mm) was chosen to properly evaluate the effect of the harmonics on the rotor losses for the different air-gap lengths.



**Figure 5:** Flux density in the air-gap 0.25 mm from the rotor different magnet shapes and magnetization directions.

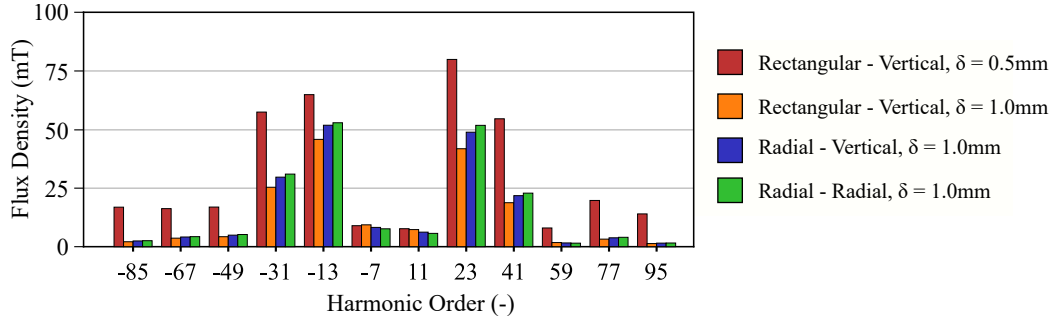
The 2D Fast Fourier Transformation presented in [5] was done to properly separate the source of the flux density harmonics. Figure 6 shows the space and frequency harmonics from the stator point of view. The harmonics are separated by their color according to their main source. The first order stator slot space harmonics 23rd and -13th are also generated by the stator winding but their main source is the stator slots. The fundamental 5th space harmonic amplitude is 606 mT and is not shown to the full extent to preserve the readability of the figure. Only the harmonics with an amplitude higher than 1 mT are shown.



**Figure 6:** Space and frequency harmonics of the flux density in the air-gap 0.25 mm from the rotor for the rectangular magnets at rated operation with highlighted sources. The fundamental 5th harmonic is not shown in its full extent.

From the rotor point of view, only the 1st frequency order harmonics are the source of the eddy-current losses and are further studied. The saturation harmonics are predominantly caused by the rotor and should therefore not contribute to the rotor losses [5]. Figure 7 shows the chosen air-gap harmonics comparison

for different shapes and  $\delta$ . The fundamental 5th harmonic for the rectangular magnets with  $\delta = 0.5$  mm reaches 606 mT, for  $\delta = 1.0$  mm 529 mT, for radial magnets with vertical magnetization and  $\delta = 1.0$  mm 619 mT and for radial magnets with radial magnetization and  $\delta = 1.0$  mm 650 mT. The model with longer air-gap benefits from lower high-order harmonics. The rectangular magnets have slightly lower slot harmonics than the radial magnets. The radial magnets with radial magnetization have slightly higher slot harmonics than the radial magnets with vertical magnetization.



**Figure 7:** High-order space harmonics of the 1st frequency order of the flux density in the air-gap 0.25 mm from the rotor for different magnet shapes and  $\delta$ .

#### 4. CONCLUSION

According to the results, the main advantage of the radial magnets in comparison to the rectangular magnets is an elimination of the bottleneck in the rotor yoke. This increases the magnetic flux for this shape of magnets which allows an increase in the air-gap length from 0.5 mm to 1.0 mm while keeping the supplied stator current nearly the same. The harmonic content of the flux density in the air-gap is mainly influenced by the length of the air-gap while the shape of the magnets has not shown a strong influence on it.

The radial magnets should be used when possible. In the analyzed machine, they reduced the rotor eddy-current losses from 90.3 W for the rectangular magnets to 58.6 W and 62.2 W for the vertical and radial magnetization respectively. The efficiency increased from 91.6 % to 92.4 % while the needed supply current was slightly lowered. The difference between the vertical and radial magnetization is marginal.

#### ACKNOWLEDGMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE).

Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports under institutional support and BUT specific research programme (project No. FEKT-S-20-6379).

#### REFERENCES

- [1] D. Gerada, A. Mebarki, N. L. Brown, C. Gerada, A. Cavagnino, and A. Boglietti, "High-Speed Electrical Machines: Technologies, Trends, and Developments," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 6, pp. 2946-2959, 2014.
- [2] S. Li, Y. Li, W. Choi, and B. Sarlioglu, "High-speed electric machines: Challenges and design considerations," *IEEE Transactions on Transportation Electrification*, vol. 2, no. 1, pp. 2-13, 2016.
- [3] P. Klima, "Copper Losses In Parallel Wires And Litz Wire In Permanent Magnet Synchronous Machine," in *Proceedings I of the 27st Conference STUDENT EEICT 2021: General papers*, pp. 529-533, 2021.
- [4] J. Pyrhönen, T. Jokinen and V. Hrabovcova, *Design of Rotating Electrical Machines*. 2nd edition Chichester, U.K.: Wiley, 2013
- [5] C. Di, I. Petrov, and J.J. Pyrhonen, "Extraction of rotor eddy-current harmonic losses in high-speed solid-rotor induction machines by an improved virtual permanent magnet harmonic machine model," *IEEE Access*, vol.7, pp. 27746-27755, 2019

# Axial Channel Effect on Leakage Reactance in Split-Winding Transformer

M. Mrajca and R. Cipín

Brno University of Technology, Czech Republic

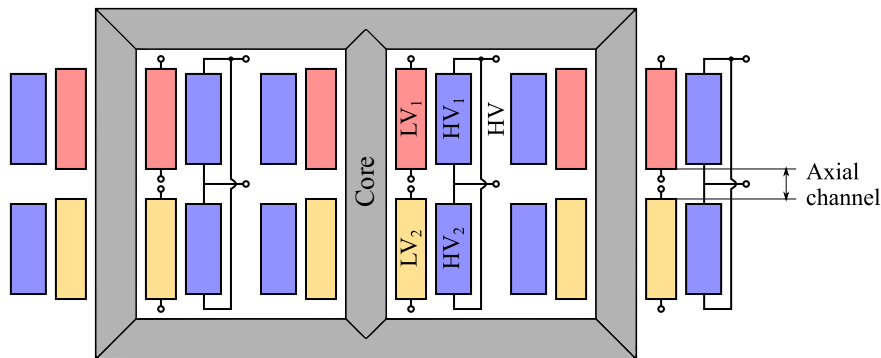
E-mail: xmrajc01@vut.cz, cipin@vut.cz

**Abstract**—Transformer short-circuit impedance is one of the most important output parameters in the design process. Its value depends mainly on leakage inductance, which is different for one or both shorted output windings during short-circuit test of split-winding transformer with input windings working in parallel. In this paper, an analysis of split-winding transformer with changing axial channel between winding pairs and its effect on leakage reactance has been performed for both shorted outputs. The effect is evaluated with analytical calculations, 3D and 2D finite element method simulations on a typical 5300 kVA three-phase oil transformer. The results show that the channel can't be neglected in short-circuit calculation and the geometry of the transformer can be simplified to 2D with minimal influence on results.

**Keywords**—split-winding transformer, leakage inductance, short-circuit impedance, axial channel, finite element method, FEM, simulation, analytical calculation

## 1. INTRODUCTION

Split-winding transformers are widely used in subway and traction networks to power supply motors of separated single undercarriages or as a standby power plant transformer, which usually consists of two (or more) pairs of coaxial windings arranged on the same core leg, while a lower voltage (LV) winding is closer to a core and a higher voltage (HV) winding is located on the outer side as illustrated in the Figure 1. For the above applications, all inner windings operate as separate outputs and independent sources of load, and all outer windings are connected in parallel as one input [1]. Inputs and outputs can be switched, for instance, in a block transformer united with more generators [2].



**Figure 1:** Three-phase split-winding transformer arrangement

For both usages, the short-circuit should be tested with one separate output winding shorted while powering both inputs connected in parallel. This test should be carried out for both separated output windings. Furthermore, a short-circuit test must be performed for all outputs shorted simultaneously.

For all described test procedures, a tolerance of  $\pm 10\%$  is applied to the required impedance value based on the standard [3]. From experience, it is possible to neglect the active component - winding resistance - in short-circuit impedance, so the impedance depends only on the leakage inductance. The split-winding transformer usually has a high value of short-circuit impedance, which can limit the starting current of the motor or the fault current [4].

In the past, many articles have discussed the calculation of leakage reactance based on analytical equations or simulations [5], [6], nevertheless, they focus mainly on one shorted output winding with disconnected

terminals of the second output, and none of these publications considers the effect of the axial channel between pair of windings on the short-circuit impedance for both shorted output windings. For example, [2], [7] and [8] deal with an equivalent circuit for the split-winding transformer with four windings for operation modes occurring during short-circuit test without reflecting the axial channel effect in the case of shorted output windings and these papers count with the fact the top and bottom winding pairs do not affect each others leakage magnetic field, so it behaves like the channel is infinite.

## 2. ANALYZED TRANSFORMER

For the case study, a three-phase split-winding hermetically sealed oil transformer with four windings was chosen. The material of the winding conductor is copper and the core consists of high permeability grain-oriented silicon steel H075-23L. HV windings are shorter than LV windings and they are mutually axially centered, it means shortenings on both outer winding ends are the same. The electrical and mechanical parameters of the analyzed transformer are listed in Table I.

**Table I:** Parameters of the analyzed transformer

Winding	HV <sub>1</sub>	HV <sub>2</sub>	LV <sub>1</sub>	LV <sub>2</sub>
Rated power	2650 kVA	2650 kVA	2650 kVA	2650 kVA
Rated voltage	23000 V	23000 V	1280 V	1280 V
Line current	66.52 A	66.52 A	1195.30 A	1195.30 A
Connection	Y	Y	d	y
Number of turns	578	578	53	31
Winding height	500 mm	500 mm	485 mm	485 mm
Internal diameter	433.2 mm	433.2 mm	311.3 mm	311.3 mm
External diameter	599.5 mm	599.5 mm	415.2 mm	415.2 mm
Core diameter	304 mm			
Height of the core window <sup>a</sup>	1035 mm			
Leg pitch	609 mm			
Short-circuit impedance	8%			

<sup>a</sup>Height of the core window for zero axial channel between top and bottom windings, final height change during parametric analysis, while the distances of winding ends from core yokes are constant.

## 3. LEAKAGE INDUCTANCE ANALYTICAL CALCULATION

The leakage inductance for a pair of concentric windings unaffected by other windings  $L$  is determined in the mentioned papers [2], [6] and [5] by the well-known following formula:

$$L = \frac{\mu_0}{H_w} \cdot K_{Rog} \cdot C \cdot N^2 \cdot \left( \frac{B_{w1}}{3} + B_{ch} + \frac{B_{w2}}{3} \right), \quad (1)$$

where  $\mu_0$  is vacuum permeability,  $H_w$  express the mean winding height,  $N$  is the number of turns,  $B_{w1}$  and  $B_{w2}$  are the widths of the concentric windings and  $B_{ch}$  represents the radial channel between them. Then  $K_{Rog}$  is the Rogowsky factor that is related to the flux fringing in the top and bottom of the windings and can be governed based on the winding geometry as:

$$K_{Rog} = 1 - \frac{B_{w1} + B_{ch} + B_{w2}}{\pi \cdot H_w}, \quad (2)$$

where variable  $C$  in equation 1 is the mean circumference of the winding pair calculated as follows:

$$C = \pi \cdot \frac{D_{w1,e} + D_{w2,i}}{2}, \quad (3)$$

where  $D_{w1,e}$  and  $D_{w2,i}$  are the external and internal diameters of the inner winding 1 and outer winding 2. The final leakage inductance for the analyzed short-circuit case is according to the equivalent circuit by [2], [7] and [8] the arithmetic mean of the results for pairs of upper and lower windings.

#### 4. ANALYTICAL METHOD INCLUDING AXIAL CHANNEL EFFECT

In the case of both short-circuited output symmetrical windings, the behavior of the transformer can be simplified to a two-winding transformer, for which can be found the formula in [9] for the calculation of leakage inductance  $L_{gap}$  taking into account the axial channel effect in the middle of the winding heights. The formula is a modification of (1) with replaced  $H_w$  and  $K_{Rog}$  by equivalent magnetic field lines length  $l_{eq}$  as follows:

$$L_{gap} = \frac{\mu_0}{l_{eq}} \cdot C \cdot N^2 \cdot \left( \frac{B_{w1}}{3} + B_{ch} + \frac{B_{w2}}{3} \right), \quad (4)$$

where the calculation applies for the equivalent magnetic field lines  $l_{eq}$ :

$$l_{eq} = 2 \cdot H_w + q \cdot \left( 2 - e^{-\frac{H_{ch}}{q}} \right), \quad (5)$$

where  $q$  is computed as a third of the width of the concentric winding pair:

$$q = \frac{1}{3} (B_{w1} + B_{ch} + B_{w2}). \quad (6)$$

The impedance is usually stated as a percentage value  $u_k$  in the short-circuit test report. Therefore, it is better to recalculate the resulting inductance in this form considering the neglected winding resistance:

$$u_k = \frac{Z_k}{Z_n} = \frac{\omega \cdot L_r \cdot I_n}{U_n} = \frac{2\pi \cdot f \cdot L_r \cdot S_n}{3 \cdot U_n^2}, \quad (7)$$

where  $Z_k$  and  $Z_n$  are short-circuit and rated transformer impedances,  $\omega$  means angular frequency,  $I_n$  and  $U_n$  are rated branch current and voltage (depends on phase connection). Then  $f$  is the nominal frequency and  $S_n$  is the rated power.  $L_r$  is generally all calculated and simulated leakage reactance in this article.

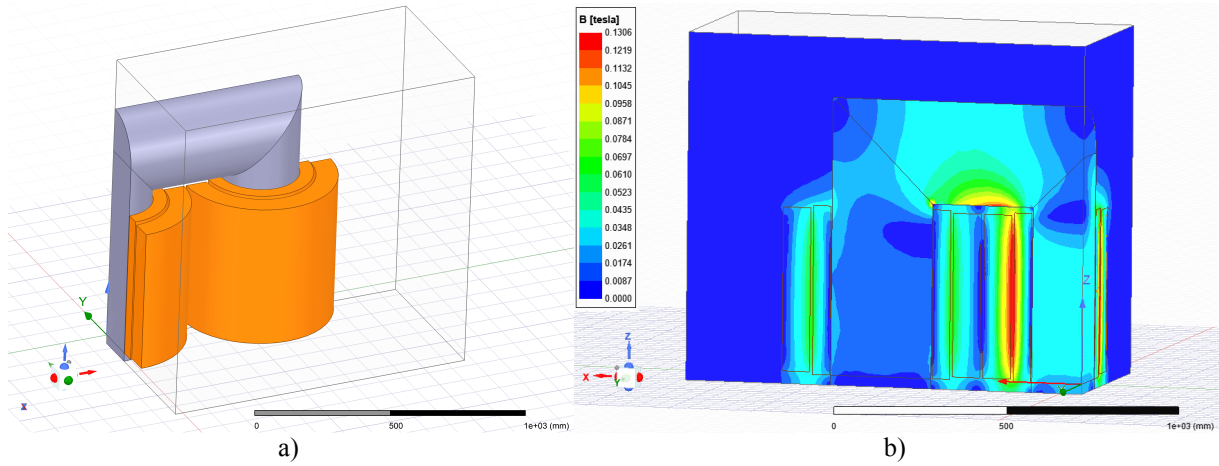
#### 5. 3D SIMULATION METHODOLOGY

At first, the leakage inductance of the split-winding transformer is calculated and analyzed using a 3D nonlinear magnetostatic finite element method (FEM) simulation. The magnetostatic mode is sufficient for leakage inductance determination, as the leakage flux density moves on the linear part of the core lamination B-H curve, as was proved in the transient testing simulation. Only an eighth of the transformer model is simulated to maximally reduce the simulation time demand using model symmetry. The core leg and yoke cross-section are considered fully circular, omitting the stepped shape for the same purpose. Windings are simplified to hollow cylinders without concluding the actual turns of conductors.

Changeable cylinders representing part of the core legs are inserted into the model to perform a parametric analysis of the axial channel between the windings. The normal flux symmetry boundary is attached to the face of the border region, representing the horizontal cut of the transformer created on the bottom of the changeable cylinders. The region without parametric extension has dimensions of 1200 mm × 600 mm × 1000 mm and the rest of the faces cutting through the transformer has a tangential flux symmetry boundary. The surrounding faces of the region are elected with the insulating boundary owing to the fringing magnetic field direction around the transformer, see Figure 2a).

The current in the middle leg phase has to be in amplitude to ensure the symmetry of the simulation, so the current in both phases of the side legs is minus half the amplitude value. Faces of cut hollow cylinders representing stranded windings are set to the current excitation of the magnitude mentioned above gained with the number of turns, so the values of LV and HV on the same core leg are identical only with opposite directions.





**Figure 2:** a) Eighth of the model for 3D FEM analysis b) Resulting magnetic flux density obtained from 3D simulation with excitation on all phases

There is no need for manual refinement of the mesh grid for all simulation approaches used, adaptive meshing brings the fastest results with a convenient error in the setup with a minimal error limit of 0.01%. The matrices coupling windings of the same phase current excitations are calculated after the simulation last pass to obtain the leakage inductance  $L_{FEM}$ . The resulting flux density field is portrayed in the Figure 2b). The obtained value must be recalculated for the complete model  $L_{full}$  using coefficient  $n$ , whose value is 8 for the middle phase and 4 for the side, and with the square number of turns  $N$ :

$$L_{full} = L_{FEM} \cdot n \cdot N^2. \quad (8)$$

Secondly, the next type of simulation analysis considers the leakage flux without affecting the surrounding phases, so the same 3D model is used and the parametric simulation is performed twice for the middle and side core limbs, each with excitation on only one phase with amplitude of the line current.

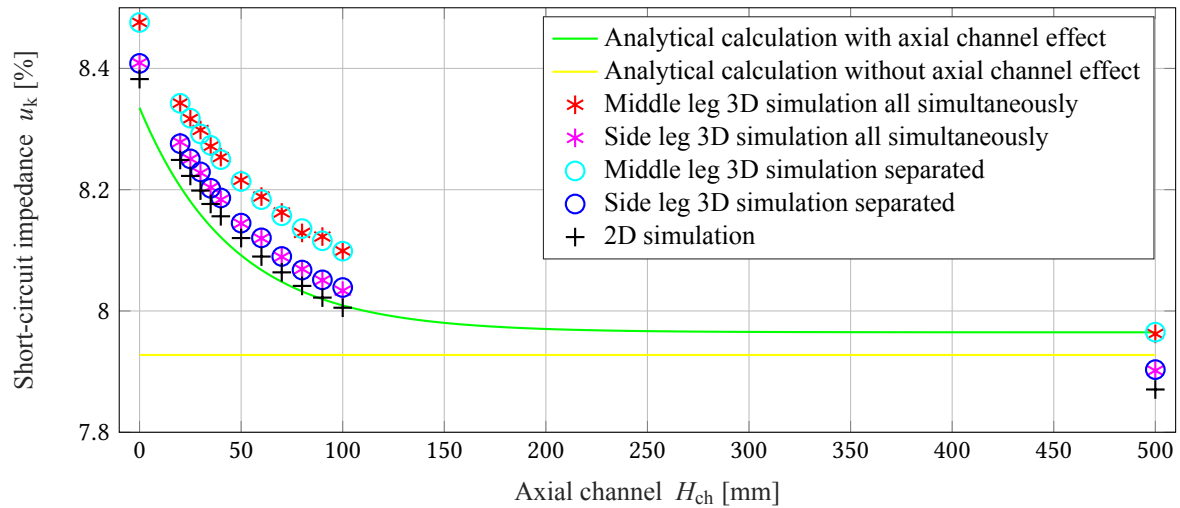
## 6. 2D SIMULATION METHODOLOGY

The influence of the yoke on the leakage flux is entirely ignored in this approach, so a 2D simulation with the cylindrical axis rotated around the z-axis can be used with the same result for the middle and side core limbs. The 2D model is based on the face from the previous 3D model cutting through the middle leg with the same conditions, except that the surrounding border is set newly with a balloon boundary. The obtained leakage inductance must be recalculated using (8) with  $n = 2$  as half of the phase is simulated.

## 7. RESULTS

Short-circuit impedance data obtained from analytical calculations and FEM simulations depending on the height of the axial channel are plotted in the Figure 3. As can be seen, the analytically calculated data are displayed as curves to show how the percentage impedance behaves according to the channel height of 0 mm to 500 mm. The points in the FEM simulations on these extreme height values are not relevant according to the transformer design; they are captured due to the analytical formula outside the presumed axial channel range verification.

As the useful range for the transformer design can be considered heights from 20 mm to 100 mm, as a result of the necessary insulation at the top and bottom of the winding and oil flow. The average decrease in impedance is approximately 0.25% in this range for all methods. However, the impedance value differs for the 20 mm channel from the original calculation approach without channel around 0.35% as the arithmetic mean for all methods, which is almost half of the allowed testing tolerance for the required 8% impedance. Taking into consideration the tolerances of the dimensions of the windings in the manufacturing process, neglecting the axial channel between the windings during the short-circuit impedance design procedure can be very risky, as the manufactured transformer will not meet the requirements of the customer. Possible reconstruction can cost additional costs in materials and manpower.



**Figure 3:** Analytically calculated and simulated short-circuit impedance depending on the axial channel height

The influence of phases has little difference from the results in the 3D simulation based on the Figure 3. The effect of the yoke makes only a minor variance in the middle and side phases, similarly to that in the 2D simulation case. All executed FEM methods provide sufficient results, though the 2D model and analytical calculation in chapter 4 bring indisputable time savings.

## 8. CONCLUSION

This paper describes the split-winding transformer with all outputs shorted short-circuit impedance calculation and points out insufficient methods in previous publications neglecting the axial channel effect. It is proved that the channel between the windings should be included in the design process based on the new analytical method, 3D and 2D FEM simulations.

## ACKNOWLEDGMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports under institutional support and BUT specific research programme (project No. FEKT-S-20-6379).

## REFERENCES

- [1] D. D. Chase and A. N. Garin, "Split winding transformers," *Transactions of the American Institute of Electrical Engineers*, vol. 53, no. 6, pp. 914–922, 1934, doi: [10.1109/T-AIEE.1934.5056749](https://doi.org/10.1109/T-AIEE.1934.5056749).
- [2] J. Kopeček, "Split-winding transformer equivalent circuit," (in Czech), *Elektrotechnický obzor*, vol. 60, no. 1, pp. 1–5, 1971.
- [3] *IEC 60076-1: Power transformers - Part 1: General*, 3rd ed., Geneva, Switzerland, 2011.
- [4] C. Yang, Y. Ding, H. Qiu, and B. Xiong, "Analysis of turn-to-turn fault on split-winding transformer using coupled field-circuit approach," *Processes*, vol. 9, no. 8, p. 1314, 2021, doi: [10.3390/pr9081314](https://doi.org/10.3390/pr9081314).
- [5] D. Azizian, M. Vakilian, J. Faiz, and M. Bigdeli, "Calculating leakage inductances of split-windings in dry-type traction transformers," *Transactions on Electrical Engineering*, vol. 10, 2012/02/01.
- [6] Y. Li, L. Li, Y. Jing, F. Han, S. Li, and F. Zhang, "Calculation and analysis of leakage magnetic field and short-circuit impedance in split-winding transformer," in *2013 International Conference on Electrical Machines and Systems (ICEMS)*, 2013, pp. 2263–2266, doi: [10.1109/ICEMS.2013.6754557](https://doi.org/10.1109/ICEMS.2013.6754557).
- [7] J. Kulda, "Equivalent circuit of a general transformer with four windings," *EtzArchiv: Archiv für elektrische Energietechnik der etz (Elektrotechnische Zeitschrift)*, vol. 5, no. 8, pp. 261–263, 1983.
- [8] P. Stuchl, L. Doležal, A. Zajíc, and O. Weinberg, "Performance of transformers with split winding under nonstandard operating conditions," *CIGRE 2000*, no. 12–103, pp. 1–5, 2000.
- [9] W. Knaack and H. Schwaab, "Additional inductance in transformer," (in German), *Archiv für Elektrotechnik*, vol. 32, no. 7, pp. 470–482, 1938, doi: [10.1007/BF01660155](https://doi.org/10.1007/BF01660155).

# An Analysis of Conical Homopolar Electrodynamic Bearing

D. Růra and J. Bárta

Brno University of Technology, Czech Republic

E-mail: [david.rura@vut.cz](mailto:david.rura@vut.cz), [bartaj@vut.cz](mailto:bartaj@vut.cz)

**Abstract**—The electrodynamic bearings (EDB) have a great perspective to substitute the traditional reluctance type electromagnetic bearings (EMB) used for supporting the rotor and keep it stable during the whole range of operation. Hybridization plays an important role in the development of magnetic bearing (MB) because combining for example axial and radial bearings into one single device results in a cost-effective solution usually smaller in size that has advantages when implemented in the motor design. In this paper, a hybrid arrangement that combines radial and axial EDB into a single device is presented. The most perspective low-loss design called homopolar EDB adopted with the traditional magnet-pole shoe arrangement is studied in the conical arrangement of rotor and stator. The results show great progress in the performance of MB as well as an easy way to balance between radial and axial load capacity.

**Keywords**—EDB, eddy current, electrodynamic bearing, induced currents

## 1. INTRODUCTION

Magnetic bearings are an outstanding replacement for traditional ball bearings. The oil-free and maintenance-free operation is the key to success in the industry. Implementing such a complicated device in the drive system is quite expensive, thus just specific applications are suitable and can revenue the initial costs within a short period. Good examples for adopting magnetic bearings in drives are turbomachinery applications such as pumps, circulators and compressors, drives for harsh environment, cardiac devices etc [1].

EMB uses electromagnets for its operation, thus besides the hardware, advanced amplifiers, as well as position sensors, are required [1]. Sensor-less control can be adopted as well but it is still considered as an experimental functionality. These are making the MB system unreachable for most simple solutions and manufacturers prefers the traditional approach to drop the costs and stay competitive on the market. The only motors suitable for adoption with MB are high-speed motors. In the high-speed operation, the friction in the ball bearings is high enough to cause serious damage and later need for bearing replacement. On the other hand, passive magnetic bearings do not require any sensors and amplifiers but they are facing low stiffness and low damping issues. Both parameters are required high enough to prevent instability in the system. Passive magnetic bearing (PMB) can be stacked from passive PMB block as a result it can perform similarly to EMB [2] but force/volume ratio is still higher for EMB.

The key to success is the stable levitation in the whole range of operating speeds of the rotor. Nevertheless, the rotor has five degrees of freedom hence, a pair of bearing for radial support and a single one for axial support is required. Another option is a bearing combining the properties of radial and axial magnetic bearing in a single device. These bearings are usually marked as hybrids. In this paper, a hybrid electrodynamic bearing that supports the rotor simultaneously in both radial and axial directions type is presented.

This paper is structured as follow. In the first section principles of operation of EDB are discussed as well as a few arrangements from literature. This is followed by key parameters definition that describes EDB. In the next section, the details about studied geometry are presented followed by the section in which finite element (FE) simulation is discussed. The last but not least, results are presented and further discussed. All gained knowledge and outputs from this paper are summarized in conclusion.

## 2. PRINCIPLE OF OPERATION

The idea of eddy current magnetic bearing is not new and there were numerous attempts in the 70s-90s [2],[3]. Research had been initiated with designs similar to today's high-speed machines with solid rotor [4]. Nevertheless, as it will be shown later, those geometries faced a problem of high copper losses in the rotor. Over the years the research reached a goal to produce low-loss EDB. Previously, the rotor spun in the static field (similar to the frozen field in the induction motor) but the field polarity under the neighbouring teeth was in the opposite direction and thus rotor faced high magnetic field change that led to high load capacity but simultaneously to high losses if the rotor spun without eccentricity around the central point. The currents are induced even rotor is not off-centred.

The low-loss bearing arrangement [3] came from the idea to relocate the varying magnetic field from the rotor periphery to the axial direction. In this arrangement (rotor centred) induced currents are the same, thus cancelling each other and literary, in case of zero current in the circuit, the losses are zero as well. When the rotor is off-centred, induced currents are not anymore the same, and reacting force pushing the rotor back to the centre is developed.

The working principle of EDBs is based on Ampere's law. According to this law, in the conductor inserted into space where an alternating magnetic field is present, the currents are induced in the conductor. This induced current creates the magnetic field around the conductor that acts against the outer field resulting in force interaction. EDB is usually made from two parts - the part in which a magnetic field is generated - actively or passively and conductor. Both can be static or rotating parts but the performance varies according to the configuration [5]. The drawback of the induced current in the conductor is its phase shift. It means that the current is not in phase with developing eccentricity in the rotor thus just part of the induced current is actively used for centring the rotor. The rest of the current develops extra drag force in the rotor system that causes instability issues.

The key parameters of the magnetic bearing are stiffness and damping. They are part of the equation motion of the harmonic oscillator that describes the force balance and motion in the vibrating system.

$$M\ddot{x} + C\dot{x} + Kx = f(t), \quad (1)$$

where  $M$  is a mass,  $x$  actual position,  $C$  is a damping,  $K$  is stiffness and  $t$  is a time.

Stiffness affects the natural frequency of the system. In a sense of EDB, the high stiffness is demanded to ensure a fast return to the central position. In the case of low stiffness, the bearing can multiply the eccentricity easily over a few spins that leads the rotor to come in contact with the stator. To prevent the system of such destruction, touchdown (landing or safety) bearings are added to keep the rotor away from the stator [1].

In the rotating system the stiffness  $k$  can be calculated as follow [3]:

$$k = -\frac{dF_B}{d(\Delta x)} \quad (2)$$

where  $F$  is a force and  $\Delta x$  is an offset from the centre. As mentioned above, the force has two components - lift and drag. The same way lift and drag stiffness can be derived.

Damping  $C$  is acting as an energy dissipator. Damping causes eccentricity amplitude to decay over time. In-plane damping and perpendicular damping describe the same lift-drag phenomenon. It can be calculated as:

$$C = -\frac{dF_B}{dv}, \quad (3)$$

where  $v$  stands for velocity.

The parameters for further examination are marked as lift because only the lift part leads to pure levitation.

### 3. STUDIED GEOMETRY

This paper deals with the conical-shaped EDB. As mentioned earlier, only a pair of conical EDBs are needed to support the rotor in all 5 degrees of freedom. The arrangement is based on the [3],[6] in which the author uses a pair of axially magnetized magnets acting against each other. The magnetic flux of both magnets meets at their interface and is radially outgoing from the magnet interface. Then, it goes around the magnet and returns to the opposite side.

To reduce the reluctance in magnetic flux paths, pole shoes could be used. One is located at the interface of magnets and the second at both ends as is shown in Fig. 1. To concentrate the flux path on the inner or outer periphery (if inner or outer rotor). The pole shoes can be narrower and coincide on the side where are higher magnetic flux is demanded. Nevertheless, for simplification, this design attribute is suspended in this work hence magnets and pole shoes vary just in thickness.

There are two options for arranging the static part. The first one is when magnets are located perpendicularly to the conical rotor surface and the second one is when magnets are located perpendicularly to the rotor axle. The first arrangement would be hard to manufacture due to the hollow conical shape of magnets but in the second arrangement, just ring axial magnets with chamfered edges are needed. For this paper, the second arrangement has been selected.

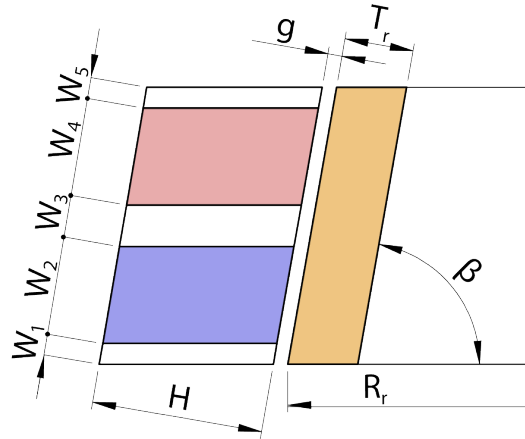


Figure 1: Cross section of conical EDB.

### 4. FE ANALYSIS

FE analysis of eddy currents is not an easy task. As well as for any other FE analysis the accuracy of the simulation is dependent on the mesh. As in [7], [3] for eddy current analysis the hexahedral mesh is most convenient. Unfortunately, the used software does support only tetrahedral mesh. With a small modification, the tetrahedral mesh can work as good as hexahedral mesh. The advantage of hexahedral mesh is that it naturally creates concentric layers in the rotor. Due to the skin effect, eddy currents flow right under the surface and with higher frequency the skin depth drops. Thus, at very high frequencies the skin depth is very thin compared to the rotor thickness and is recommended to have at least 3 layers of mesh in the skin depth.

In the following section, the results from FE analysis are presented followed by a discussion.

### 5. RESULTS

FE analyses have been simulated to obtain performance parameters of conical EDB. In the Tab. I are listed obtained from the analysis. Three  $\beta$  angles have been selected and the rest of the parameters follows the design guide from [3].

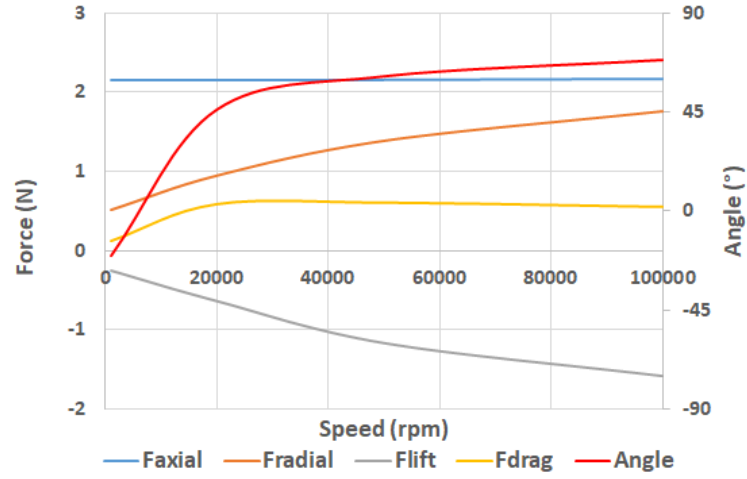
The performance of the conical EDB for one selected  $\beta$  is shown in Fig. 2. Only selected parameters are shown, the rest of them are listed in the Tab.I.

**Table 1:** Resulted data from FE analysis for three  $\beta$ , dimensions of analyzed EDB.

$\beta(^{\circ})$	65				75				85			
Speed (x1000 rpm)	100	50	20	1	100	50	20	1	100	50	20	1
$F_{axial}$ (N)	2.17	2.16	2.15	2.15	1.54	1.53	1.52	1.52	0.42	0.42	0.42	0.42
$F_{radial}$ (N)	1.76	1.39	0.95	0.52	2.03	1.60	1.07	0.46	2.54	2.08	1.54	1.04
$F_{lift}$ (N)	-1.58	-1.16	-0.63	-0.25	-1.83	-1.33	-0.69	-0.34	-2.00	-1.46	-0.76	-0.40
$F_{drag}$ (N)	0.56	0.61	0.59	0.12	0.81	0.84	0.78	0.12	1.21	1.19	1.08	0.16
$\alpha(^{\circ})$	68.65	61.24	46.20	-20.05	65.57	57.21	41.49	-42.03	57.12	50.17	34.93	-14.41
$K_{radial}$ (kN/m)	17.6	13.9	9.5	5.2	20.3	16.0	10.7	4.6	25.4	20.8	15.4	10.4
$C_{radial}$ (Ns/mm)	168	265	453	4957	194	305	512	4425	242	397	735	9909

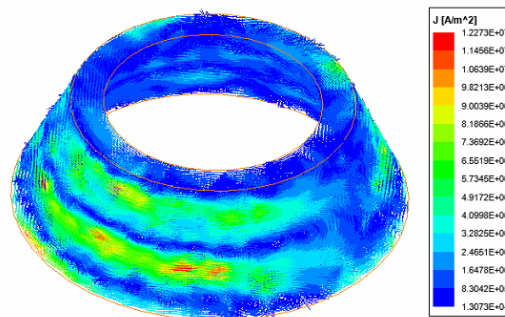
Dimensions of EDB:  $g = 1$  mm;  $d_x = 0.1$  mm;  $R_b = 30$  mm;  $H_{1-5} = 10$  mm;  $W_{1,5} = 1.5$  mm;  $W_{2,4} = 7$  mm;  $W_3 = 3$  mm;  $T_r = 5$  mm

Materials: rotor - copper, stator pole shoes - steel1008, magnets - NdFe30



**Figure 2:** Output performance of conical EDB for  $\beta = 65$ .

The current path is shown in Fig. 3. As can be seen, the main current path is right under the axial magnet. From the figure is seen where the current loop has its end. The direction of the restoring force is exactly at the spot where the maximal current is flowing.



**Figure 3:** Current loops in the rotor.



## 6. DISCUSSION

In the previous section, all obtained data are plotted and listed. The results support the theory that the induced currents and forces are speed-dependent. Also, the idea behind the conical arrangement is obvious. Raising  $\beta$  leads to balancing between radial and axial force generated in the EDB. For larger  $\beta$  radial force prevails and for smaller  $\beta$  axial force prevails.

In Fig. 2 can be seen that from around 20000 rpm the bearing spins stably. From that point, drag force is slowly decreasing but lift force strengthens. The negative angle  $\alpha$  indicates that EDB is unstable. It can be also noted that axial force is not varying for the whole range of speeds. It is caused by the non-changing axial part of the airgap. In this bearing, the force in the axial direction is generated even the eccentricity is zero. Similar results for radial EDB has been presented in [8]

This type of EDB is usually built as a stack. It means multiple stages are placed in parallel which leads to a bearing with better performance in a sense of higher stiffness and damping.

## 7. CONCLUSION

This paper deals with the conical electrodynamic bearing. FE analysis has been done for three different skewed rotors. The premise of balancing between radial and axial forces in the EDB has been proved. Also, it can be noted that stability is ensured above 20000 rpm. For stronger EDB can Halbach arrangement of magnets be used. It was analyzed beyond the coverage of this paper and the lift force for  $\beta = 65$  is almost doubled. EDB bearings are a promising device for high-speed application but because of low stiffness and damping and instability issues for lower frequencies. If they are adopted in the drive, the coupled AMB is usually used to stable the system at low speeds. Also, possible issues with the presented arrangement could be in sourcing such shaped magnets.

## ACKNOWLEDGMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports under institutional support and BUT specific research programme (project No. FEKT-S-20-6379).

## REFERENCES

- [1] G. Schweitzer and E. H. Maslen, *Magnetic bearings*. New York: Springer, 2009.
- [2] N. Amati, A. Tonoli, E. Zenerino, J. Girardello Detoni, and F. Impinna, "Design methodology of electrodynamic bearings," 01 2009.
- [3] T. A. Lembke, "Design and analysis of a novel low loss homopolar electrodynamic bearing," Stockholm, 2005.
- [4] J. Bárta, "The eddy current losses of high-speed solid rotor induction machine as function of the radial air-gap length," in *Proceedings of the 22nd Conference STUDENT EEICT 2016*. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2016, pp. 507–511. [Online]. Available: <http://hdl.handle.net/11012/83985>
- [5] V. Kluyskens and B. Dehez, "Analysis of the stability and the consumption of an electrodynamic bearing for different operation conditions," in *13th International Symposium on Magnetic Bearings*, 2012, pp. 5–18.
- [6] S. Xu and J. Fang, "A novel conical active magnetic bearing with claw structure," *IEEE Transactions on Magnetics*, vol. 50, pp. 1–8, 2014.
- [7] M. Hedlund, J. Abrahamsson, J. Pérez-Loya, J. Lundin, and H. Bernhoff, "Eddy currents in a passive magnetic axial thrust bearing for a flywheel energy storage system," *International Journal of Applied Electromagnetics and Mechanics*, vol. 54, pp. 1–16, 04 2017.
- [8] M. Lopes, G. Sotelo, E. Rodriguez, and R. Stephan, "An electrodynamic magnetic bearing," 09 2014.

# A Nodalization Study of Phebus FPT-1 Experiment in MELCOR 2.2

Tomas Cerny<sup>1</sup> and Michal Ptacek<sup>1</sup>

<sup>1</sup>Department of Electrical Power Engineering, Brno University of Technology, Technicka 12, 61600 Brno, Czech Republic

E-mail: [xcerny67@vut.cz](mailto:xcerny67@vut.cz), [ptacekm@vut.cz](mailto:ptacekm@vut.cz)

**Abstract**—Phebus, an integral experimental facility located in France, is scaled down version of a French 900 MWe pressurized water reactor (PWR) by a ratio of 1/5000. The goal of the facility is to study phenomena occurring during light water reactor severe accidents and validate models and computer codes used for calculation of severe accidents. The aim of this paper is to compare different modelling approaches in MELCOR 2.2 computer code and the influence on thermal hydraulic behavior and core degradation in comparison with experimental results.

**Keywords**—MELCOR, Phebus, FPT-1, ISP-46, nodalization, thermal-hydraulics

## 1. INTRODUCTION

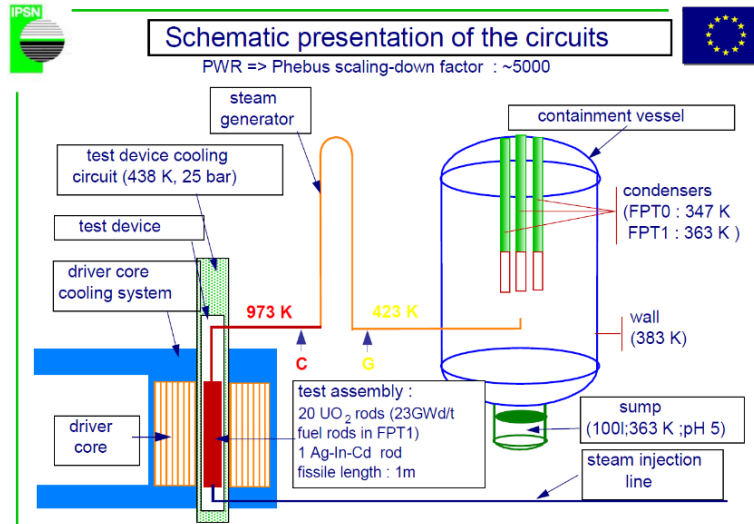
The Phebus FPT (Fission Product Tests) programme was carried out by IRSN, a French nuclear regulatory, between 1988 and 2010. Five integral experiments were conducted to study severe accident related phenomena with focus on core degradation, release, transport and deposition of fission products. Another objective of these experiments was to assess the capability of computer models and codes to simulate accident progression. Due to the broad phenomena focus, the Phebus tests are considered as one of the most essential experiments dedicated to severe accident research [1]. In this paper, the MELCOR code was selected to model these severe accident phenomena.

Based on [2], MELCOR is a fully integrated, engineering-level computer code that models the progression of severe accidents in light water reactor nuclear power plants (NPPs). MELCOR was developed at Sandia National Laboratories for U.S. NRC as a second-generation plant risk assessment tool. A broad spectrum of severe accident phenomena in both boiling and pressurized water reactors is treated in MELCOR. These include thermal-hydraulic response in the reactor coolant system, reactor cavity, containment building; core heat-up, degradation and relocation; core-concrete interaction; hydrogen production, transport and combustion; fission product release and transport behaviour. MELCOR applications include estimation of severe accident source terms and their sensitivities and uncertainties in a variety of applications. Each of the above mentioned process is handled by a specific package (CVH, FL, COR, etc.). The packages communicate with each other on a specific basis. The MELCOR code is designed as highly parametric code. Most of the models in each package and their parameters can be modified by sensitivity coefficients and sensitivity study is thus easily feasible.

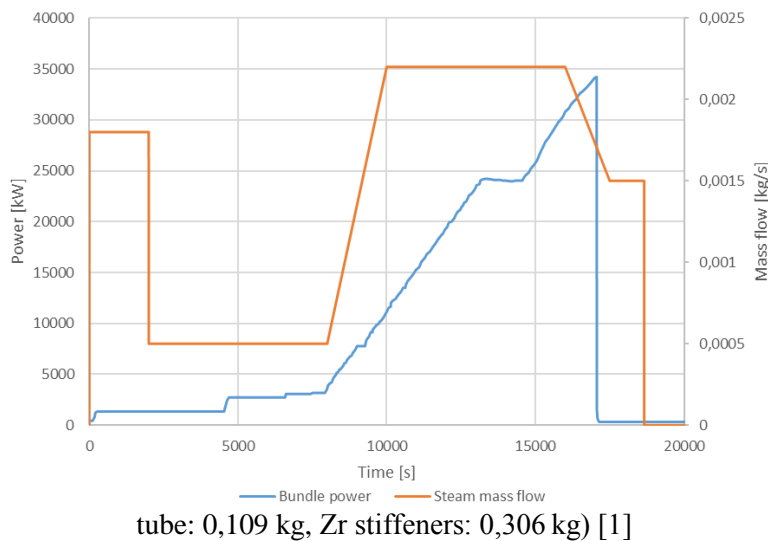
## 2. FACILITY AND EXPERIMENT DESCRIPTION

The Phebus facility, depicted in Fig. 1, is described using references [1,3]. The facility contains a driver core, which provides a fission power to a test fuel bundle, a cooling circuit and a containment vessel. The test bundle contains 20 fresh fuel rods laid on a zircaloy support plate and held in place by two zircaloy spacer grids. In the centre of the bundle, there is an absorber rod containing Ag-In-Cd. The test bundle is insulated from the driver core by zirconia shroud with inner ZrO<sub>2</sub> or ThO<sub>2</sub> layer, external ZrO<sub>2</sub> layer and an Inconel pressure tube with inner ZrO<sub>2</sub> coating. The test bundle is cooled by a mass flow of steam imposed at the core inlet. The experimental circuit made of Inconel-600 connects core exit with containment and simulates the thermal-hydraulic conditions of hot leg, steam generator and cold leg. The containment vessel collects the coolant, aerosols and gases transferred through the experimental circuit and simulates a PWR containment building.

The aim of the FPT-1 experiment was to simulate the conditions of Loss of Coolant Accident (LOCA) in PWRs. The experiment was conducted by gradually increasing bundle fission power up until about 36,5 kW. The bundle was cooled by pure steam with inlet pressure of 2 bar and mass flow varying from 0.5 to 2.2 g/s to provide strong oxidising conditions. After 17039 second after start of the experiment, the nuclear power was terminated and the containment vessel was isolated from the cooling circuit after 18660 seconds. The fission power and steam mass flow progression during test is depicted in Fig. 2.



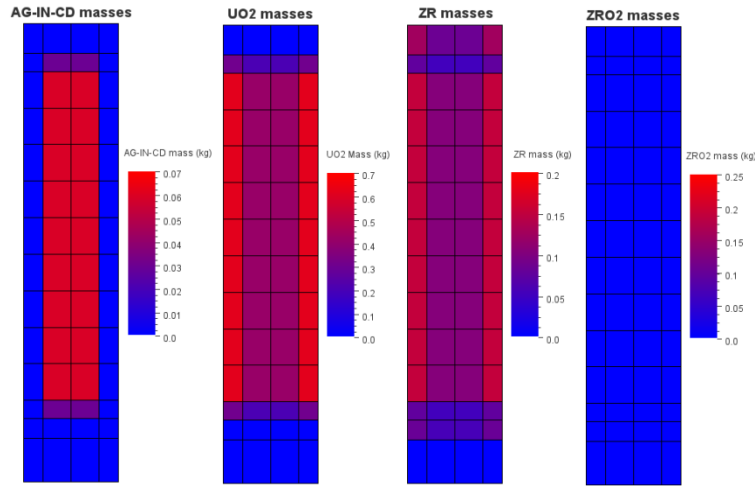
**Figure 1:** A schematic diagram of the Phebus FPT-1 experiment (FPT-1 fuel bundle mass inventory -  $\text{UO}_2$  fuel: 10,424 kg, Zr cladding: 2,932 kg, Ag-In-Cd control rod: 0,597 kg, Zr spacer grids: 0,156 kg, Stainless-steel control rod cladding: 0,138 kg, Zr control rod guide



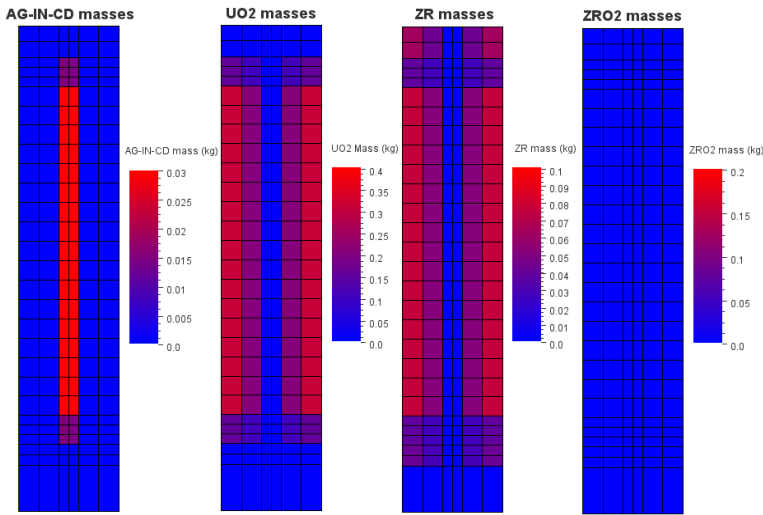
**Figure 2:** Steam mass flow rate and fission power of the fuel bundle [1]

### 3. PROPOSED NODALIZATION

The FPT-1 input deck was developed in MELCOR 2.2 using information and geometry from [1,4,5]. The modelling of core region uses COR package, which divides the core region into a mesh (nodes) with certain amount of masses of fuel, cladding, control rod poison, supporting and non-supporting structures. User also has to define the core geometry, cross-sectional areas for coolant flow, hydraulic diameters etc. to calculate core thermal response. The modelling approach to nodalization can influence core thermal response and degradation as well as calculation time of the input deck. For the purpose of this paper to demonstrate the impact of core nodalization modelling, the proposed medium and fine nodalization mesh of the fuel bundle is depicted in Fig. 3 and Fig 4, respectively.



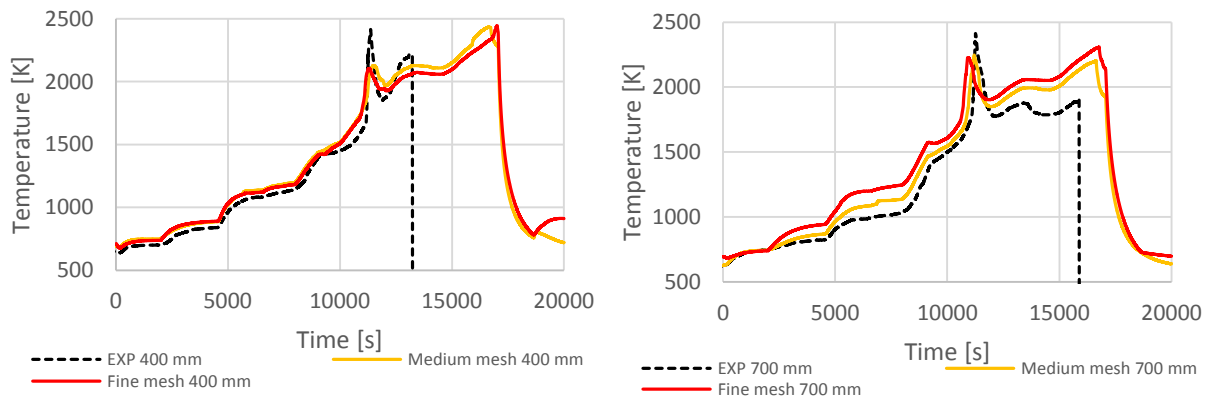
**Figure 3:** Medium mesh (14 axial and 2 radial levels) initial mass distribution



**Figure 4:** Fine mesh (28 axial and 3 radial levels) initial mass distribution

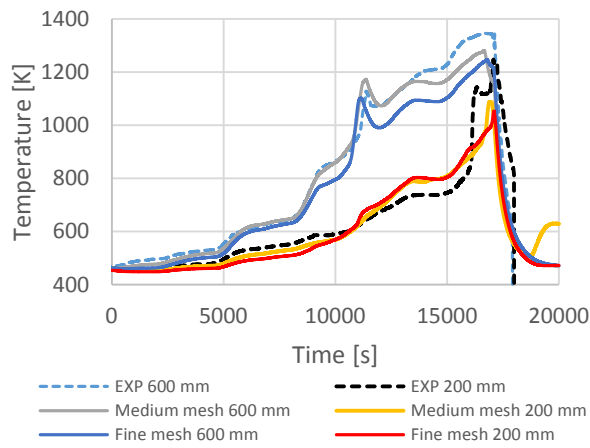
#### 4. COMPARISON OF RESULTS

To evaluate the influence of medium and fine mesh MELCOR nodalization on thermal-hydraulic behavior, five parameters were selected for comparison with experimental measurements from FPT-1 test and are presented in following Fig. 5, Fig. 6 and Fig. 7 including the evaluation of the total amount of generated hydrogen, respectively. Fig. 8 and Fig. 9 display the final state of fuel bundle damage.

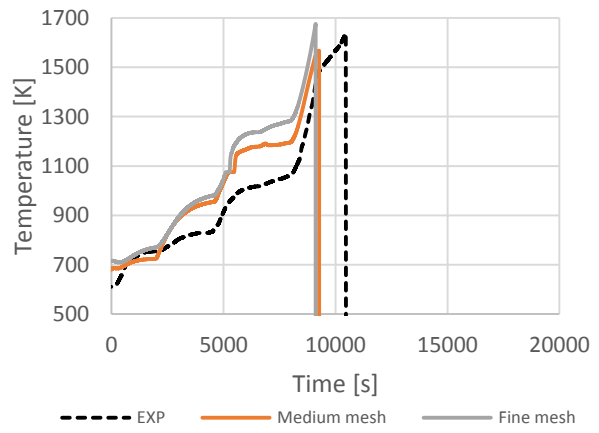


**Figure 5:** a) Temperature of outer fuel rods

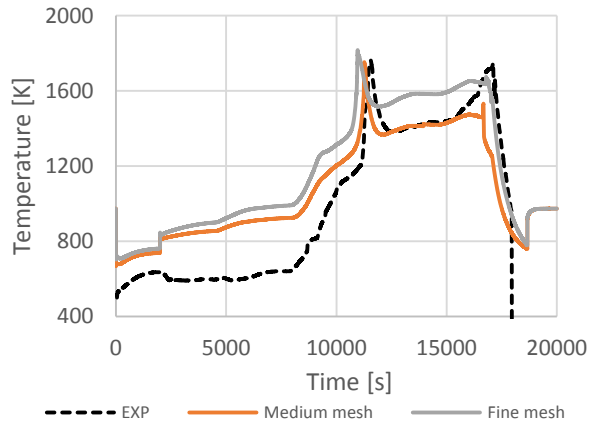
b) Cladding temperature of outer rods



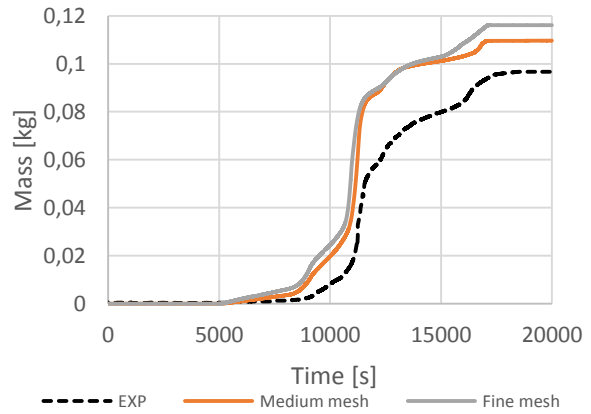
**Figure 6:** a) Temperatures inside core shroud



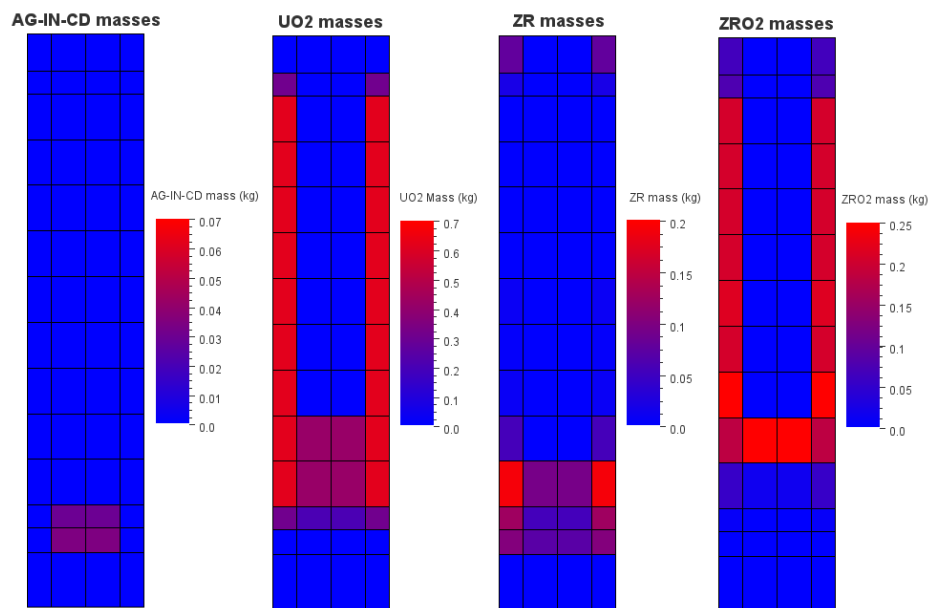
b) Temperatures of control rod at 700 mm



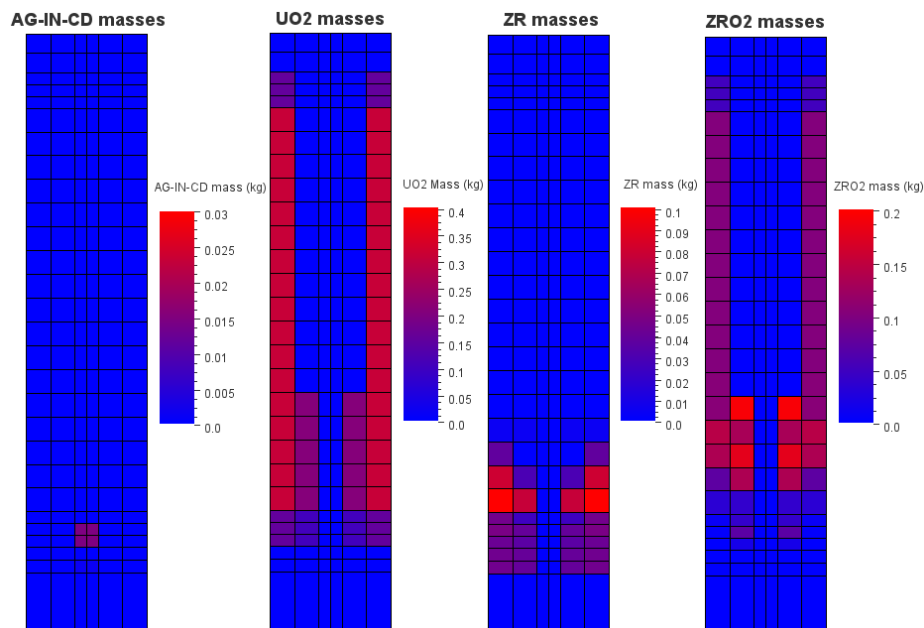
**Figure 7:** a) Temperature of coolant at core exit



b) Total amount of generated hydrogen



**Figure 8:** Medium mesh post-test mass distribution



**Figure 9:** Fine mesh post-test mass distribution

## 5. CONCLUSION

In this paper, a nodalization study of Phebus FPT-1 experiment in MELCOR 2.2 code has been performed. The study evaluates two nodalization meshes of the fuel bundle, one using 2 radial rings and 14 axial levels and the second one using 3 radial rings and 28 axial levels. An attempt to calculate a nodalization using 1 radial ring and 6 axial level was not successful since the code was crashing immediately causing an unknown error, therefore this nodalization was not taken into account. The masses of the fuel bundle (see Fig. 1 description) were accordingly distributed into mesh cells and the initial mass distribution at the start of the experiment is depicted in Fig. 3 and Fig. 4.

After MELCOR calculations of both nodalizations, the selected thermal-hydraulic parameters are compared with experimental measurements (dashed black lines) in Fig. 5, Fig. 6 and Fig. 7 a). The total amount of generated hydrogen is compared in Fig. 7 b), since it is a good indicator of a cladding oxidation process. From figures it can be conducted, that both nodalizations produce over-estimated results in comparison to experiment. However, the medium mesh nodalization provides more accurate results compared to experimental values for all of the evaluated parameters. Figures 8 and 9 show the mass distribution at the end of calculation. In both cases, the inner fuel rod ring was mostly destroyed, while the outer fuel rods remained intact. MELCOR calculations also revealed that almost all of the zirconium present in cladding underwent oxidation and transformed into  $ZrO_2$ . Similar core damage was found to be present after examination of the experimental fuel bundle using an X-ray image. The inner fuel rods were completely destroyed above the elevation of the first spacer grid located at 0,25 m from the bottom of the fuel bundle.

The MELCOR nodalization study revealed that the medium mesh provides a better overall calculation results in comparison with experiment than a fine mesh nodalization. The main advantage of the fine mesh is that it allows to model core damage more precisely. However, the code user has to bear in mind, that finer mesh strongly influences the calculation time. In case of medium mesh, the calculation took 9 minutes 32 seconds, while for fine mesh, the calculation took 28 minutes 9 seconds to complete (a 295 % increase). Since MELCOR is designed to calculate severe accidents in NPPs and the FPT-1 experiment has just one fuel bundle, the effect of nodalization on calculation time is therefore multiplied significantly for NPP calculation.

## ACKNOWLEDGEMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy



(CVVOZE). Authors gratefully acknowledge financial support from the Technology Agency of the Czech Republic (project No. TM02000039).

## REFERENCES

- [1] B. Clément and T. Haste, “Comparison Report on International Standard Problem ISP-46 (PHEBUS FPT1)”, NT SEMAR, July 2003.
- [2] L.L. Humphries, B.A. Beeny, F. Gelbard, D.L. Louie, J. Phillips, R.C. Schmidt and N.E. Bixler, “MELCOR Computer Code Manuals Vol. 1: Primer and User’s Guide, Version 2.2.19018 2021”, Sandia National Laboratories, Albuquerque (USA), January 2021.
- [3] P. March and B. Simondi-Teisseire, “Overview of the facility and experiments performed in Phébus FP”, *Annals of Nuclear Energy*, vol. 61, pp. 11-22, 2013.
- [4] P. Darnowski, M. Włostowski, M. Stępień, and G. Niewiński, “Study of the material release during PHÉBUS FPT-1 bundle phase with MELCOR 2.2”, *Annals of Nuclear Energy*, vol. 148, 2020.
- [5] P. Darnowski, P. Mazgaj, and M. Włostowski, “Uncertainty and Sensitivity Analysis of the In-Vessel Hydrogen Generation for Gen-III PWR and Phebus FPT-1 with MELCOR 2.2”, *Energies*, Vol. 14, No. 16, pp. 1-28, 2021.

# Plunging Brno into darkness: Measuring the contribution of public lighting to the total luminance of the night sky

F. Novák<sup>1</sup> and P. Baxant<sup>1</sup>

<sup>1</sup>Department of Electrical Power Engineering, Faculty of Electrical Engineering and Communication, Brno university of Technology, Czech Republic

E-mail: [xnovak1x@vutbr.cz](mailto:xnovak1x@vutbr.cz), [baxant@vut.cz](mailto:baxant@vut.cz)

**Abstract**— The article deals with the description of measurements of night sky luminance in urban agglomeration with emphasis on the determination of changes in the values of these luminances when public and architectural lighting is temporarily turned off. The measurements took place in Brno, the second largest city in Czech Republic, with the population around 382 thousand. The presented paper is one of the perspectives on the impact of human activity on the surrounding environment and an example of a possible solution to quantify the impact of the introduction of technologies that directly or indirectly affect the environment not only for humans, but also for almost all forms of life.

**Keywords**—ALAN, artificial light at night, luminance measurement, light pollution, public lighting, luminance analyzer, intrusive light, LumiDISP, astronomy.

## 1. INTRODUCTION

The main source of light pollution is in most cases public lighting, so for this experiment more than 43 thousand outdoor lamps were temporarily turned off and then a luminance measurement of the sky was carried out using a luminance analyzer.

Light pollution is the negative effect of artificial lighting in the outdoor environment. For a more impartial description, there is also the relatively recently introduced term “ALAN”, or artificial light at night, which aptly describes the nature of this phenomenon.

The issue of ALAN has many levels. The first level is economic since the light causing the disturbance must first be produced. The luminous flux emitted directly from the luminaires into the upper half-space (referred to as ULR) does not help us to illuminate the space in question, which is of course negative both from an economic and ecological point of view.

The second and more serious view of the problem is the environmental impact of ALAN. In synanthropic communities, most species of animals and plants are most likely affected by artificial light at night. Insects that are active at night are attracted to artificial light sources, which can lead to their depletion due to unnecessary energy wastage, or they become prey to predators for which they are easy targets in such a situation [1]. For example, fish are usually repelled by artificial light and have reduced melatonin levels [2], while amphibians have behavioral and juvenile developmental disorders [3]. ALAN also negatively affects mating behavior of some songbird species [4]. In plants, disturbing light can cause adverse growth changes [5].

Finally, ALAN also has negative effects on the human body. It fundamentally affects the production of melatonin in the pineal gland, which affects the quality of sleep and thus the regenerative processes in the body of humans and other mammals. It also increases levels of the stress hormone cortisol, which in combination with the above can cause seasonal depression and contribute to obesity and diabetes. Disturbing light at night also appears to influence the incidence of some cancers, particularly prostate and breast cancer, as reported, for example, by [6] and [7].

It is clear from the above information that the effect of ALAN is clearly negative in many respects. This raises the question of the contribution of public lighting systems to the artificial luminance of the night sky, and this was the main motivation for the experiment.

The luminance measurement was chosen because it appeared at the time to be the most objective

parameter for global impact assessment, not because it was the most important parameter.

In parallel with the described luminance measurements, aerial photography of Brno and its surroundings was also carried out, resulting in a photomap including all still functioning outdoor light sources [8].

## 2. PERFORMED MEASUREMENTS

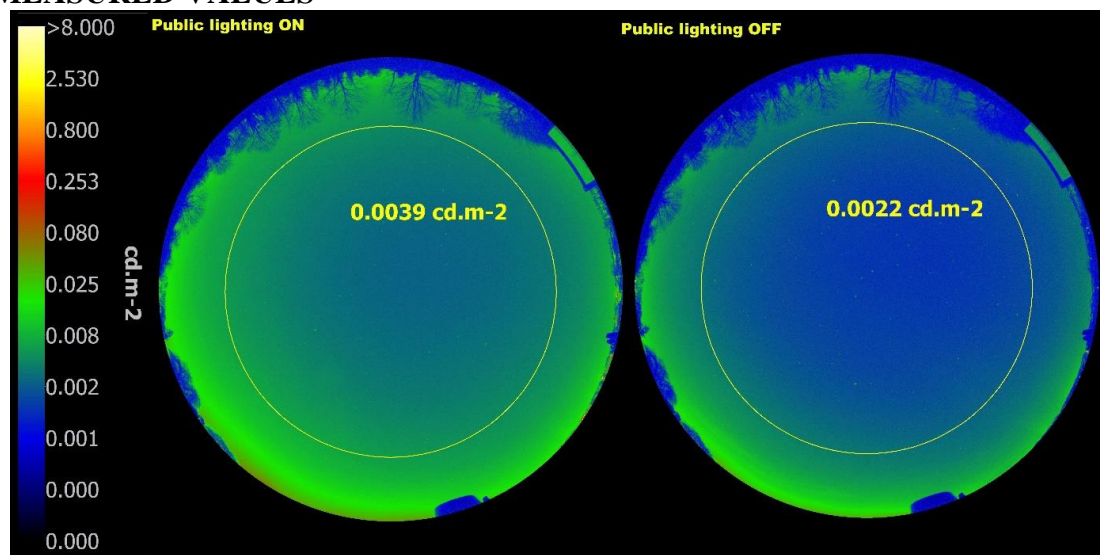
The idea to temporarily switch off the public lighting in Brno and measure the difference in night sky luminance was born at the end of 2020 and was initiated by assoc. prof. Baxant from the Faculty of Electrical Engineering and Communication of Brno University of Technology. In coordination with the Brno City Municipality, the director of TSB Ing. Rouček and the company Topgis s.r.o. (in charge of the aerial photography) the experiment was planned for the night of 10th April 2021. Due to the unfavorable weather forecast for this night, the event was in the last minute rescheduled to a night earlier.

The parking lot above the Kraví hora swimming pool (at GPS coordinates 49.2012858N, 16.5873644E) was chosen as the location for the measurements of the sky luminance, mainly because of its convenient location relatively in the center of the city, sufficient view mainly to the south and west and the absence of nearby light sources that would affect the measurements. The location is close to the Brno Observatory and Planetarium and the overall night lighting in this area is already relatively well designed without unnecessary sources of stray light such as advertisements.

The described measurements were performed using the LDA - LumiDISP luminance analyzer based on a Nikon D7500 APS-C SLR camera equipped with a Sigma 4.5 mm F2.8 EX DC HSM Circular Fisheye lens and a Sigma 135 mm F1.8 DG HSM lens.

The actual shutdown of the public lighting system was scheduled for 0:30 on April 10th, followed by a 3:30 turn-on. This time window was mainly influenced by the necessary interval for detailed aerial photography of the whole area. The climatic conditions were suitable for the measurements, with clear skies and an approximate temperature of 3 °C. Another advantage of the chosen date was that the Moon was in the waning crescent phase with only 2 % of the Moon surface illuminated and the measurement took place after the Moon has set, meaning it had no effect on the values measured.

## 3. MEASURED VALUES

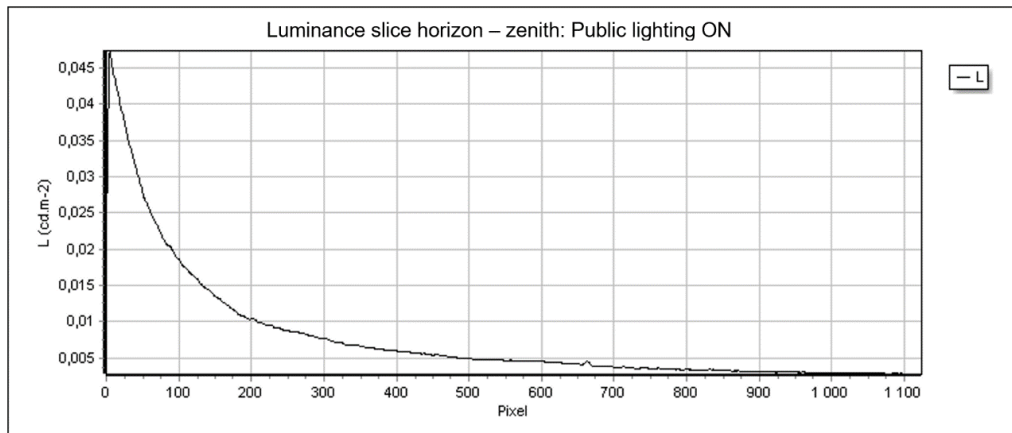


**Figure 1:** Views of the zenith and the average luminance values considering a detection angle of 120°.

These two images were taken with a panoramic fisheye lens and were created using HDR composition with exposure times from 30 s to 1/8000 s. The luminance maps themselves show a significant drop in luminance values of the whole sky (the color palette is of course set the same for both images).

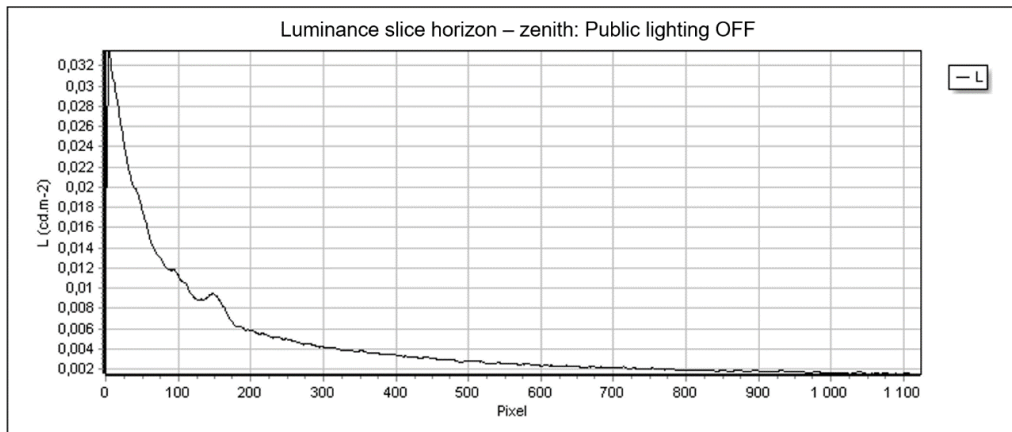
If we draw a circle around the zenith, which is defined by a viewing angle of 120° (or 60° from the zenith), and average all the luminance values contained in it, we get a value of 3.85 mcd.m-2 in the first case and 2.16 mcd.m-2 in the second case. This represents a 43.9% reduction in the percentage of sky veil luminance.

A slice was also made from the most intense illumination source in the lower left of the image to the zenith and the luminance was again measured on this slice. A circle with a radius of 6 pixels was used as the detector for all sections. This was done to reduce the number of random spikes and dips in the plot caused by CMOS sensor noise.



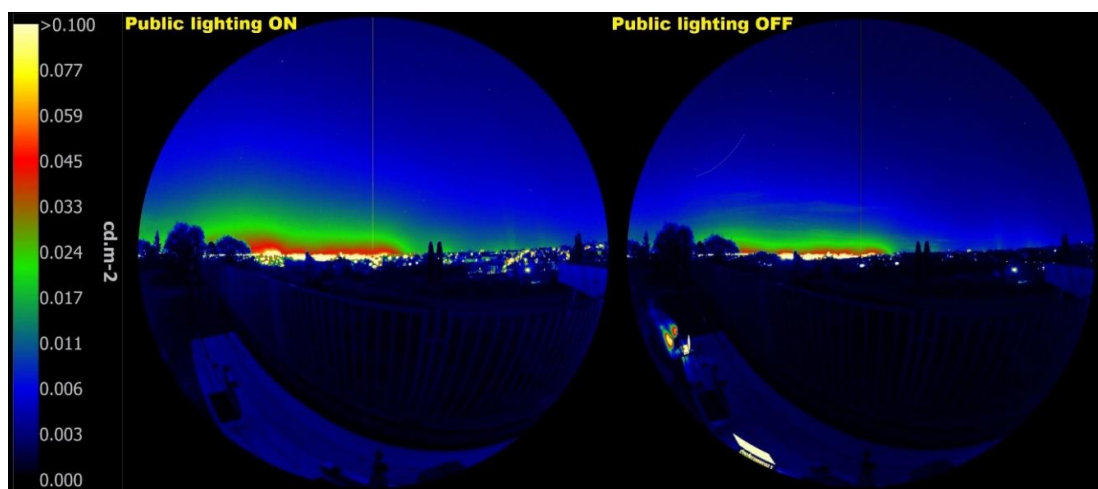
**Figure 2:** Luminance progression from horizon to zenith – public lighting turned on.

This was again repeated for the switched off state:



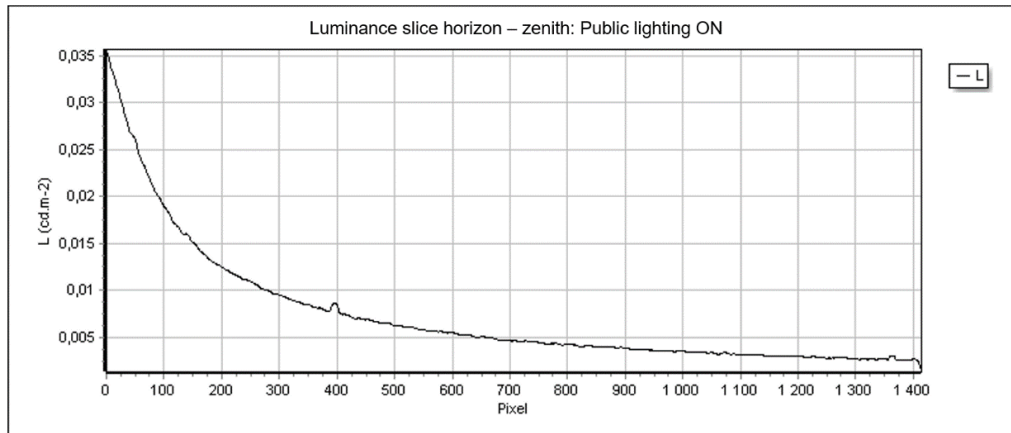
**Figure 3:** Luminance progression from horizon to zenith – public lighting turned off.

These graphs also show a significant decrease in both maximum and average sky luminance values - on average by 25 to 45 % (even though the curve looks almost identical, there is a different scale of the Y axis). Maximal measured value of luminance was 47.3 mcd.m-2 in the first case and 33.4 mcd.m-2 in the second, meaning a drop of 29.4 % was achieved in this case.



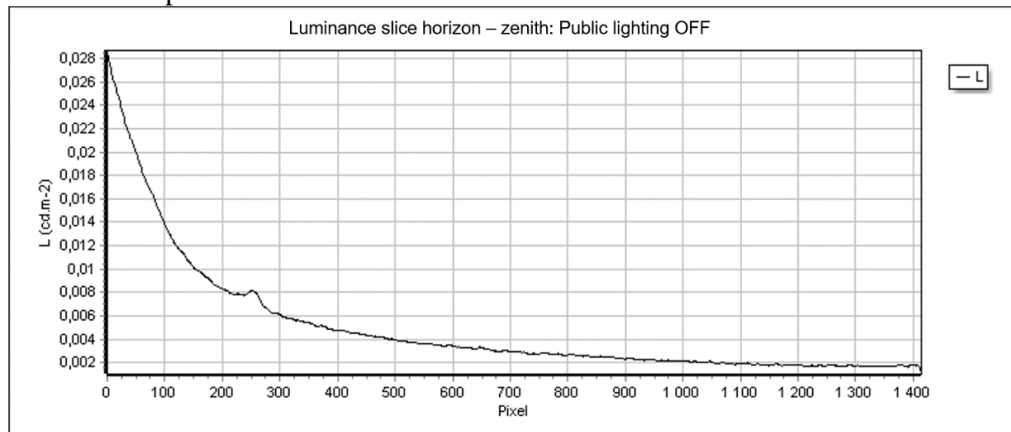
**Figure 4:** Luminance maps looking at the southern horizon.

Here again, the drop in sky luminance is immediately apparent across the entire image range, although many light sources on the horizon are no longer under control of the TSB company, and therefore these sources were not included in the experiment, and thus turned off. If these lighting arrays had been included, there would naturally have been an even more dramatic drop in luminance values. To investigate the sky luminance, a slice was again made through the center of the image.



**Figure 5:** Luminance progression from horizon to zenith – public lighting turned on.

The same was then repeated for the turned off state:



**Figure 6:** Luminance progression from horizon to zenith – public lighting turned off.

There is also a drop in measured luminance values, but mainly due to the larger number of lighting systems on the horizon that were not turned off, it is not as significant as in the first case - approximately 15 to 35% through the slice. Maximum luminance in this case dropped from 35,6 mcd.m-2 to 28,5 mcd.m-2, representing a decrease of 19,9 %.

#### 4. CONCLUSION

The experiment and subsequent measurements confirmed what was expected, i.e., that public lighting systems for roads, parks, monuments, and others contribute significantly to the luminance of the sky in and around the urban agglomeration. By switching off more than 43 000 public lighting points, a reduction to almost half of the original value of luminance was achieved (see Fig. 1 through 3).

The contribution of outdoor lighting systems to the artificial luminance of the night sky is influenced by the geometry of the luminaire and its output, its location and direction, but also the spectrum of the light emitted. Luminaires with a higher correlated color temperature contribute much more to the sky glow than luminaires with a lower CCT. Luginbuhl et al. in [9] report that a light source with  $T_c = 5100$  K contributes approximately 7 times more to the sky glow than an LPS lamp with the same luminous flux in a comparable luminaire. This implies the need in the future for outdoor lighting to use not only luminaires with appropriate geometry and direction of luminous flux, but also light sources with the lowest possible CCT, e.g., using LEDs of the PC amber type or even HPS lamps. This physical reason is a minor argument for using spectrums with longer wavelengths. A more compelling reason is the

sensitivity of living organisms to short wavelengths, especially in the blue region. This is understandable because all life forms evolved in an environment where the blue color of the scattered light of the sky provided, and still provides, information about the time of day and thus sets the biological clock of organisms. A compromise choice for outdoor applications are also the so-called biodynamic or adaptive light sources, which adapt the light flux and spectrum to evening, morning, and nighttime. An optimal complement to such a system would then also be the direct sensing of the presence of humans in the respective space, i.e., whether there is anyone to light for at all.

This experiment was also possible due to the unfavorable period in which the experiment took place, namely the fact that due to the COVID pandemic, a night curfew was in place from 9pm to 5am. The question arises as to whether it is necessary, especially in such a situation, to shine more than 43 000 light sources into streets with minimal traffic and almost no pedestrians - and these are the figures for Brno alone. The negative effects of intrusive lighting are unquestionable and go hand in hand with the energy consumption of such extensive lighting, regardless of its actual need.

## ACKNOWLEDGMENT

The authors would like to thank the Centre for Renewable Energy Research and Utilization (CVVOZE) in which this publication was written and for the financial support of the specific research project at the BUT (specific project no. FEKT-S-20-6449). Authors also gratefully acknowledge Brno City Municipality, Technical Networks Brno and Brno Observatory and Planetarium.

## REFERENCES

- [1] DALEY, Jason. *The Devastating Role of Light Pollution in the 'Insect Apocalypse'* [online]. Available from: <https://www.smithsonianmag.com/smart-news/light-pollution-contributes-insect-apocalypse-180973642/>
- [2] Anika Brüning, Franz Hölker, Steffen Franke, Torsten Preuer, Werner Kloas, Spotlight on fish: *Light pollution affects circadian rhythms of European perch but does not cause stress*, Science of The Total Environment, Volume 511, 2015, Pages 516-522, ISSN 00489697, Available from: <https://doi.org/10.1016/j.scitotenv.2014.12.094>
- [3] WISE, S. *Studying the ecological impacts of light pollution on wildlife: amphibians as models*. Journal of Thermal Biology. 2007, č. 6, 107-116.
- [4] DA SILVA, Arnaud, Mihai VALCU a Bart KEMPENAERS. Light pollution alters the phenology of dawn and dusk singing in common European songbirds. *Philosophical Transactions of the Royal Society B: Biological Sciences* [online]. 2015, **370**(1667). ISSN 0962-8436. Available from: <https://doi.org/10.1098/rstb.2014.0126>
- [5] Singhal, Rajesh & Kumar, Mahesh & Bose, Bandana. (2018). *Ecophysiological Responses of Artificial Night Light Pollution in Plants*. Russian Journal of Plant Physiology. 66. Available from: <https://doi.org/10.1134/S1021443719020134>
- [6] Zelinski EL, Deibel SH, McDonald RJ. *The trouble with circadian clock dysfunction: multiple deleterious effects on the brain and body*. Neurosci Biobehav Rev. 2014 Mar; 40:80-101. Available from: <https://doi.org/10.1016/j.neubiorev.2014.01.007>
- [7] Stevens, R. G., Brainard, G. C., Blask, D. E., Lockley, S. W., & Motta, M. E. (2014). *Breast cancer and circadian disruption from electric lighting in the modern world*. CA: a cancer journal for clinicians, 64(3), 207–218. Available from: <https://doi.org/10.3322/caac.21218>
- [8] Noc v Brně – mapa zhasnutého Brna [online]. Available from: <https://brno.cz/nocni-brno>
- [9] LUGINBUHL, Christian B., Paul A. BOLEY a Donald R. DAVIS. The impact of light source spectral power distribution on sky glow. Journal of Quantitative Spectroscopy and Radiative Transfer [online]. 2014, 139, 21-26. ISSN 00224073. Available from: <https://doi.org/10.1016/j.jqsrt.2013.12.004>



# Experimental investigation of sodium chloride neutron induced reactions for nuclear reactors cooled by molten salts

J. Burian<sup>1</sup>

<sup>1</sup>Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Electrical Power Engineering; Czech Republic

E-mail: [xburia22@vut.cz](mailto:xburia22@vut.cz)

**Abstract**—The article deals with the experiment within the Master's thesis written on Department of Electrical Power Engineering in Faculty of Electrical Engineering and Communication (FEEC) on Brno University of Technology (BUT). The behavior of NaCl in neutron field was studied via irradiation indium cylinders surrounded by the salt. As a result are charts of reaction rate for two nuclear reactions and three measuring channels.

**Keywords**—BUT, EEICT, FEEC, molten salt, molten salt reactor, neutron source, nuclear research

## 1. INTRODUCTION

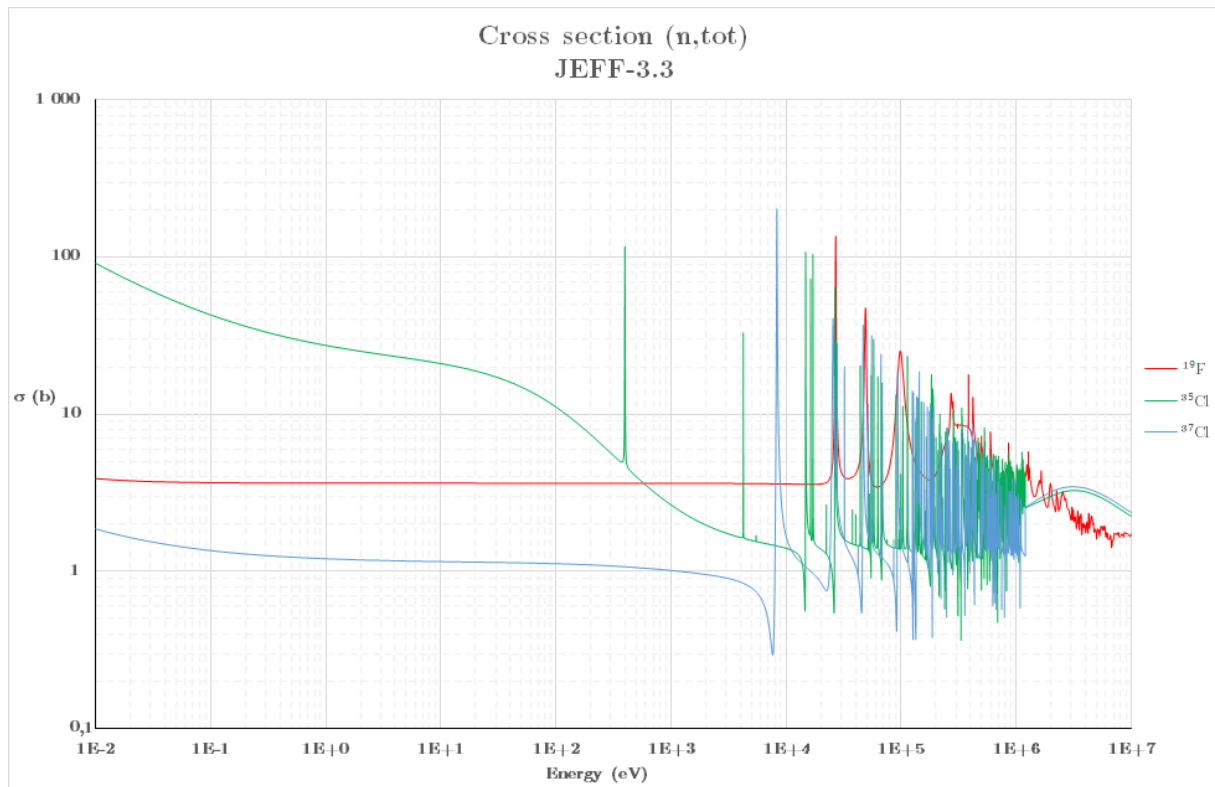
Molten salts have been used to cooling fission nuclear reactors with specific operating parameters due to their properties. Unlike the light water (for example), which can be heated to a maximum of 647.096 K (373.946 °C) at 22.064 MPa while maintaining a liquid state [1], saline solutions can be heated to a higher temperature in atmospheric pressure and they still stay liquid.

The reasons for using molten salts as a reactor coolant are their high boiling point (same samples are available in the tabs I and II), which allows generate steam with high parameters or run some chemical processes that needs high temperature to be processed. The thermal conductivity is higher in case of molten salts than in case of (light) water [2, 3]. Next advantage is absence pressure above atmospheric so the requirements for construction materials are low in terms of pressure resistance.

On the other hand using molten salt solutions as a reactor coolant brings a new technical problems that doesn't exist when the water is used for cooling the reactor. At first: corrosive behavior of (molten) salts place high demands on construction materials of reactor, pipes, pumps and seals which increases the cost of the structure [6]. Secondly, the relative high melting point (samples in tabs I and II) causes problems when shut down or maintenance work is needed. If the reactor is shutted-down it can't generate enough heat to keep the salt mixture in a liquid state. So there is a hazard that salt starts freezing, especially in the pipe-curves, which can clogged the pipelines, block the liquid flow and heat exchange and cause damage on the system. To avoid the problems, salt itself has to be removed from the system during shutting down for example by gradually diluting the saline solution with water. Other possible way is using the external heat sources like resistive heaters, resistive wires or gas burners (all around the pipes with molten salt inside). However this will bring new potential dangers in form of power loss (and therefore heat loss) and, in addition, will increase costs during outage due to the costs associated with energy consumption for heating the pipes with salt.

## 2. SALT EXAMPLES AND THEIR PROPERTIES

In the past, the saline mixture based on fluorine was used as a coolant for the nuclear reactors. The main goal was found the energy source for nuclear powered airplanes and molten salt reactors was chosen for that [7]. The reason why fluorine salts was used were their properties from the point of view of neutron physics. In the following graphs is evident that <sup>19</sup>F has lower effective of cross-section values in the field of thermal neutrons. Cross section is much higher in case of chlorine so thermal neutrons will be more absorbed than in fluorine. No matter that <sup>37</sup>Cl isotope has lower course of the cross section because in the nature is only 24.23 % of this isotope [9]. The effect of <sup>35</sup>Cl will be dominant. The cross section for neutron elastic scattering for chosen isotopes will be very similar as in chart 1.



**Figure 1:** Cross section for neutron reaction (total) [8].

On the following two tables (tab I and tab II) are properties of some salts based on chlorine or fluorine. The monitored properties are melting point, boiling point and density. If the boiling point is lower than melting point (for example  $\text{AlCl}_3$ ) the salt will sublime after reached the boiling point.

Although using chlorine salt coolant mixtures in thermal fission reactors is problematic and ineffective, they offers good solution for cooling fast breeder reactors or accelerator driven systems where problem with neutron capture is eliminated. In the first case is it due to fast neutrons and in the second one is enough number of neutrons generated by spallation reactions from spallation target. At the same time fluorine salts have better moderating properties which are needed for operating the thermal fission reactor unlike the reactors works with fast neutrons [4].

The coolant mixture wasn't only basic salt like examples in tabs I and II but the complex solutions were used. This step was chosen to improve or change properties of reactor coolant such as neutron moderation, thermal capacity, thermal conductivity. etc. The requirements were the lowest possible melting point, the highest possible boiling point, good moderation and thermodynamics properties, small corrosive effects and the lowest possible price. As an example could be  $\text{LiF-NaF-KF}$  (known as  $\text{FLiNaK}$ ),  $\text{LiF-BeF}_2$  ("FLiBe) or  $\text{NaF-ZrF}_4$  [4, 6].

**Table I:** Fluorine-based salts and their properties [4]

Compound	Melting point (°C)	Boiling point (°C)	Density (g/cm <sup>3</sup> )
$\text{AlF}_3$	1 290	1 280	2.88
$\text{BeF}_2$	554	1 169	1.99
$\text{CaF}_2$	1 418	2 500	3.18
KF	858	1 503	2.48
LiF	870	1 681	2.64
$\text{MgF}_2$	1 263	2 260	3.15
NaF	993	1 695	2.56

**Table II:** Chloride-based salts and their properties [4]

Compound	Melting point (°C)	Boiling point (°C)	Density (g/cm <sup>3</sup> )
AlCl <sub>3</sub>	192	180	2.48
BeCl <sub>2</sub>	399	482	1.89
CaCl <sub>2</sub>	722	1 935	2.15
KCl	770	1 420	1.98
LiCl	605	1 382	2.07
MgCl <sub>2</sub>	714	1 412	2.32
NaCl	801	1 465	2.16

### 3. EXPERIMENT

The goal of experiment was study the behavior of NaCl in the neutron field. The laboratory AmBe neutron source was used in the measuring set that was designed and compiled at the Department of Electrical Power Engineering. The assembly has the form of a box with four cylindrical holes in it, which goes through the whole breadth. The holes are market as channels – first channel is determined for neutron source and other ones are for samples for irradiation. First measuring channel (MCH1) is 115 mm away from the source channel (SCH) and is parallel to it. Second measuring channel (MCH2) is like MCH1, but 244 mm apart. The last – third measuring channel (MCH3) – is 430 mm away from the source and is perpendicular to the SCH.

The box was filled with pure NaCl and hermetically closed to avoid possible action with water (even with humidity). As a material for irradiation was chosen indium because of half-life of reaction yields and their easy measurable gamma-energy. Indium has two isotopes in its natural form: <sup>115</sup>In in 95.7% representation and <sup>113</sup>In (in 4.3 %) [9], therefore the isotope <sup>113</sup>In can be neglected (its effect will not be reflected in the results). Indium was in form of 12 cylinder with average weight of 31,91 g. The indium cylinders was stacked vertically, fixed to each other and put in one of three measuring channels and the irradiation process itself took two days (during weekend).

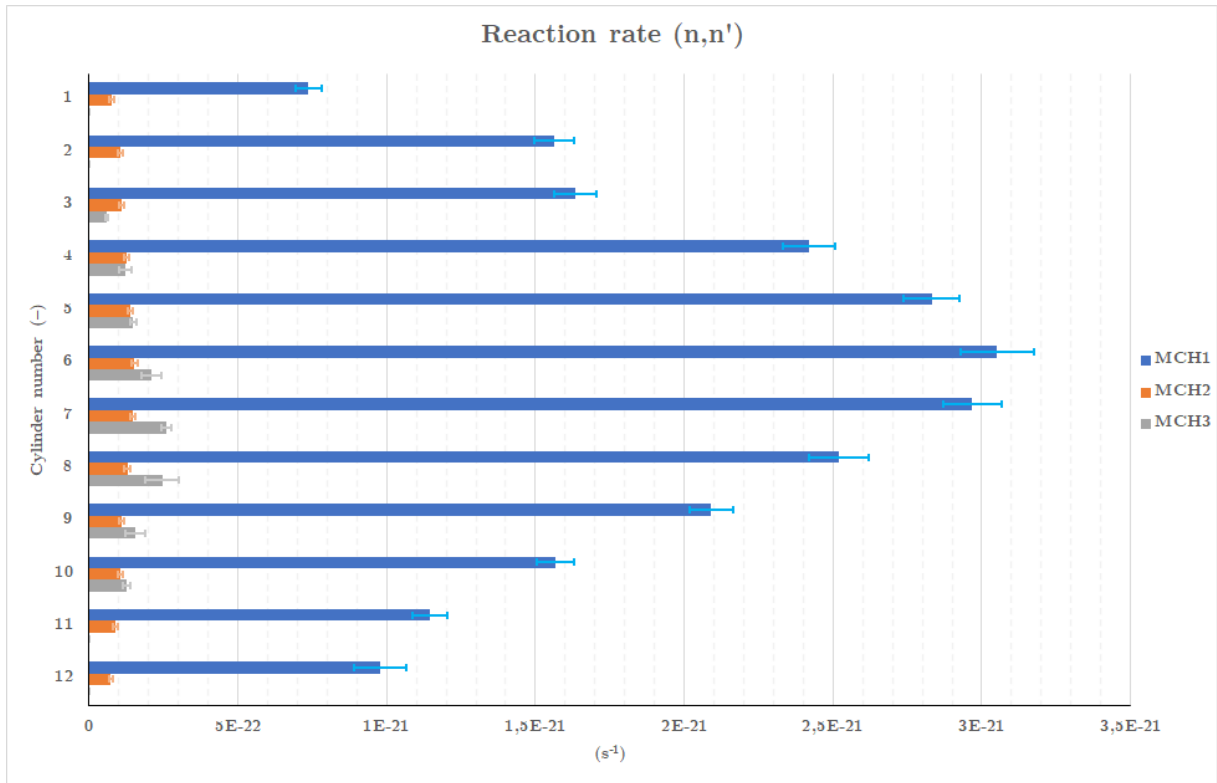
Irradiated samples were one by one moved into HPGe detector, where was their gamma spectrum analyzed. After measurement the obtained spectrums were processed in GAMWIN software and the areas of peaks from gamma-spectroscopy were written down with its uncertainty. After process were obtained values calculated in equation 1 to find the reaction rate in each position of the indium cylinders in the salt box.

$$R_R = \frac{M \cdot S(E_\gamma) \cdot \lambda \cdot \frac{t_{REAL}}{t_{LIVE}}}{N_A \cdot m \cdot \varepsilon_{FEP}(E_\gamma) \cdot I_\gamma(E_\gamma) \cdot (1 - e^{-\lambda \cdot t_{IRR}}) \cdot e^{-\lambda \cdot t_{COOL}} \cdot (1 - e^{-\lambda \cdot t_{REAL}})} \cdot \frac{1}{\eta_A} \quad (1)$$

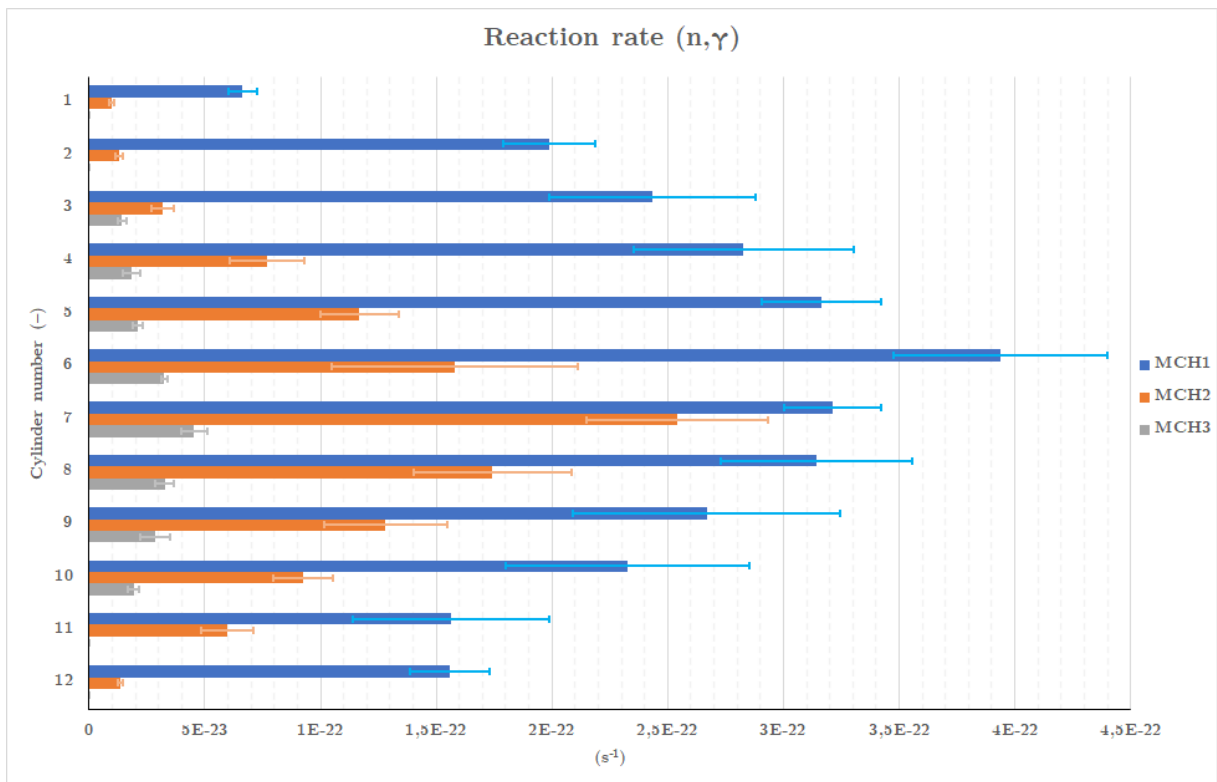
where  $M$  is molar weight of indium (g/mol),  $S(E_\gamma)$  is peak area (–),  $\lambda$  decay constant (s<sup>–1</sup>),  $t_{REAL}$  is real time of measurement (s),  $t_{LIVE}$  is time of measurement without dead-time (s),  $N_A$  represents Avogadro constant ( $N_A = 6,02214076 \cdot 10^{23} \text{ mol}^{-1}$ ),  $m$  is sample weight (g),  $\varepsilon_{FEP}(E_\gamma)$  is full absorption peak detection efficiency (–),  $I_\gamma(E_\gamma)$  is gamma-intensity (–),  $t_{IRR}$  time of irradiation (s),  $t_{COOL}$  is time between the end of irradiation and start of every measurement (s) and  $\frac{1}{\eta_A}$  is correction of self-shielding (–) [4].

Calculated values were processed into graphs (on fig. 2 and fig. 3) where is possible clearly view how was neutrons weakened by the salt. Reaction yields from (n,γ) reactions are less intensive than in case of (n,n') and in addition they have shorter half-life. This is the reason why the results of (n,γ) have higher fault.

From charts is also readable the asymmetry of distribution of the reaction rate. Values obtained from bottom cylinders (12, 11, 10, ...) have higher reaction rate than the upper ones. This is probably caused by neutron reflection from floor and bottom side of metal material from which is built the box with salt. The finally results can be compared with theoretical values obtained from ENDF libraries like American ENDF B–VIII.0, European JEFF–3.3, Chinese CENDL–3.2, Russian BROND–3.1 and others. This is also part of work that had been done but the results are not included in this article.



**Figure 2:** Results of reaction rate ( $n,n'$ ) from experiment [4].



**Figure 3:** Results of reaction rate ( $n,\gamma$ ) from experiment [4].

#### 4. CONCLUSION

Results of experiments provide evidence of the effect of salt on the neutron field. A total of 61 spectrums were collected: 18 from MCH1, 32 from MCH2 and 11 from MCH3. Peaks areas were processed according to Equation 1 and the results are visible on graphs 2 and 3. Unfortunately because of the COVID-19 pandemic, it wasn't possible to make more measurement and take more of spectrums. However, the current procedure and results will be used for further research. Experiment offer a many ways to continue: make more measurements to refine the results; exchange the salt from NaCl to other chloride based (for ex.  $\text{MgCl}_2$ ); irradiate another samples; etc.

#### ACKNOWLEDGMENT

Author is grateful to Mr. Dusan Kral, who was responsible for experimental setup and Mr. Karel Katovsky for his help, expert advice and comments.

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Author gratefully acknowledge financial support from the Ministry of Education, Youth and Sports of the Czech Republic under the Inter-Action project ADAR (LTAUSA18198).

#### REFERENCES

- [1] VELOSO, Marcelo A. *THERMODYNAMIC PROPERTIES OF WATER IN THE CRITICAL REGION*. INAC 2009: International Nuclear Atlantic Conference: Innovations in Nuclear Technology for a Sustainable Future. Rio de Janeiro, Brazil: Brazilian Association for Nuclear Energy, 2009, Pages 10. ISBN 9788599141038. [pdf](#)
- [2] REDDY, Ramana G. *Novel Molten Salts Thermal Energy Storage for Concentrating Solar Power Generation*. The University of Alabama (UA): United Technologies Research Center, Pages 27. [pdf](#)
- [3] TOULOUKIAN, Y.S., P.E. LILEY and S.C. SAXENA. *THERMOPHYSICAL PROPERTIES OF MATTER: Nonmetallic Liquids and Gases*. Volume 3. New York - Washington: IFI/PLENUM, 1970. SBN: 306-67020-3. [pdf](#)
- [4] BURIAN, Jiří. *EXPERIMENTAL AND CALCULATIONAL SALTS' PROPERTIES INVESTIGATION FOR MSR REACTORS FROM NUCLEAR DATA POINT-OF-VIEW*. Brno, 2021. Master's Thesis. BUT. Supervisor Doc. Ing. Karel Katovský, Ph.D. [pdf](#)
- [5] BEZDĚK, Z., J. HESOUN, Z. MAREŠ and Z. VALVODA. English–Czech and Czech–English disctionary on nuclear physics and nuclear technology. Prague: SNTL–Publishers of technical literature, 1985, Pages 344. ISBN L28-S1-VI-31/11928. 04-024-85.
- [6] W . D. Manly, J . W. Allen, W . H. Cook, J . H. DeVan, D . A . Douglas, H. Inouye, D . H. Jansen, P . Patriarea, T . K. Roche, G . M. Slaughter, A . Taboada, and G . Al. Tolson .*CONSTRUCTION MATERIALS FOR MOLTEN-SALT REACTORS*. Pages 595–625. [pdf](#)
- [7] ROSENTHAL, M.W. and R.B. KASTEN. *MOLTEN-SALT REACTORS—HISTORY, STATUS, AND POTENTIAL*. Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, 1969 [pdf](#)
- [8] N. Soppera, M. Bossant, E. Dupont, *JANIS 4: An Improved Version of the NEA Java-based Nuclear Data Information System*, Nuclear Data Sheets, Volume 120, June 2014, Pages 294-296. ENDF library: JEFF-3.3. DOI: [10.1016/j.nds.2014.07.071](https://doi.org/10.1016/j.nds.2014.07.071)
- [9] S.Y.F. Chu, L.P. Ekström and R.B. Firestone, *WWW Table of Radioactive Isotopes*, database version 2.0. 1999. <http://nucleardata.nuclear.lu.se/nucleardata/toi/>





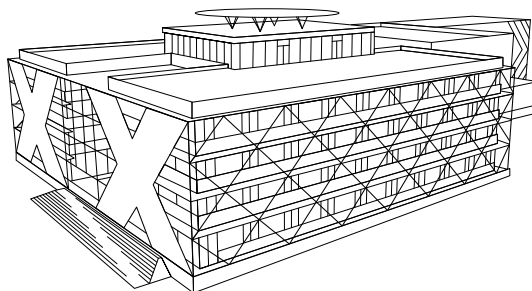
DĚLÁME CHYTRÉ TECHNOLOGIE  
A ŘÍDÍME DOPRAVU.

TADY I VE SVĚTĚ.

[www.cross-traffic.com](http://www.cross-traffic.com)

## KDO JSME?

Jsme to my, kdo z obyčejných měst dělá chytrá města. Staráme se o to, aby se dobře parkovalo, průjezd městem byl rychlejší a bez kolon, aby silnice a dálnice nebyly přetížené a aby bylo možné efektivně řídit dopravu. Je nás na to více než 160, máme vlastní tým IT vývojářů a spoustu dalších lidí, kteří se starají o hladký průběh našich projektů. Pracujeme v 55 zemích světa, ale sídlo máme ve Zlíně. Můžeš si vybrat technologie, které tě zajímají, a pracovat na projektech na té nejvyšší úrovni, které lze v našem oboru dosáhnout.



## KOHO HLEDÁME?

Technicky zdatné, jazykově vybavené zaměstnance zapálené pro elektrotechniku s chutí pracovat a rozvíjet se.

- PROGRAMÁTOŘY (SW & HW)
- SALES ENGINEERY
- ELEKTROINŽENÝRY
- ELEKTROTECHNIKY



VÁŽENÍ  
ZA JÍZDU



ŘÍZENÍ  
DOPRAVY



PARKOVACÍ  
SYSTÉMY



SILNIČNÍ  
METEOROLOGIE



DETEKCE  
PŘESTUPKŮ



DETEKCE  
DOPRAVY



5 TÝDNŮ  
DOVOLENÉ



13. PLAT



PRUŽNÁ  
PRACOVNÍ DOBA



NOTEBOOK, MOBIL  
A MOBILNÍ TARIF



PŘÍSPĚVEK  
NA STRAVOVÁNÍ



KURZY  
A VZDĚLÁVÁNÍ



PŘÍSPĚVEK NA  
SPORTOVNÍ AKTIVITY



DOSTANEŠ SE  
DO CELÉHO SVĚTA



## ZAÚJALI JSME TĚ?

Pošli životopis na [kariera@cross.cz](mailto:kariera@cross.cz).  
Potkáme se a pokud nám to bude spolu  
klapat, rádi tě vezmeme mezi sebe.



# An analysis of the birdcage resonators using magnetic field probes, for validation of optimal aspect ratio

K. Jurik<sup>1</sup> and P. Drexler<sup>1</sup>

<sup>1</sup>Brno University of Technology, FEEC, DTEEE, Czech Republic

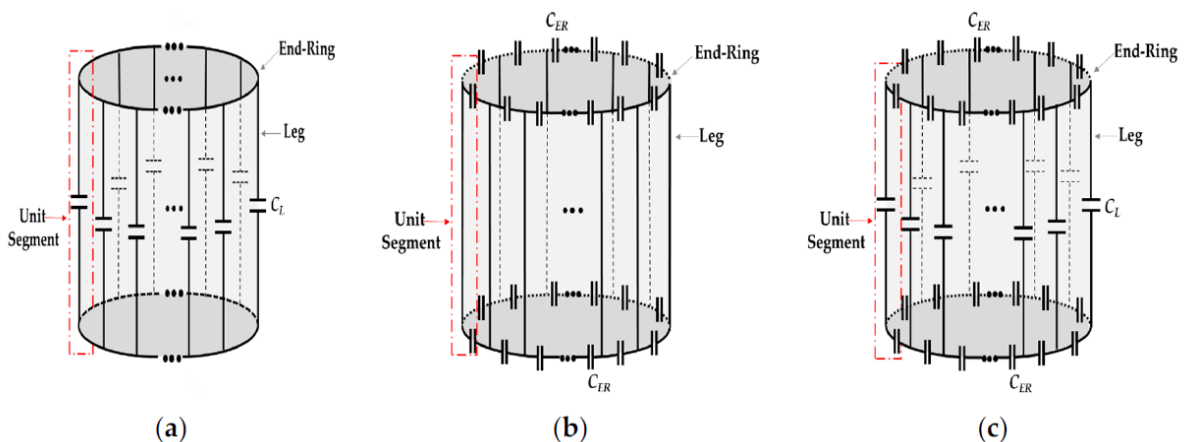
E-mail: karel.jurik@vut.cz, drexler@vut.cz

**Abstract**—Two probes were created in order to analyse the magnetic field of the birdcage resonators with different heights and aspect ratios. The first probe has a larger pickup loop at its end and was shielded to reduce the electric field coupling. The second probe with a smaller loop was not shielded at its end. The relative values of magnetic field intensity obtained using the two probes were similar, however, the first probe has better sensitivity. Both measurements confirm that there exists an optimal height and, therefore, an aspect ratio, of the resonators with respect to the generated magnetic field intensity.

**Keywords**—birdcage resonator, magnetic field probe, pick-up loop, current probe

## 1. INTRODUCTION

Birdcage resonators (sometimes referred to as birdcage coils) are usually used as the transmitting or receiving coils in magnetic resonance imaging (MRI) instruments. The resonators consist of even number of legs and two endrings at their ends. There are also capacitors, placed in the legs or in the endrings, which together with the inductance of the conductors creates the resonance [1]. If the capacitors are placed in the endrings, the resonator is high-pass type, if the capacitors are in the legs a low-pass type is created. There is also a combination of the two foregoing structures: a band-pass resonator (the three types are visualised in Figure 1). Each resonator has  $N/2$  resonant modes, where  $N$  is the number of legs. One of these modes (for high-pass the mode located at the highest frequency, for low-pass at the lowest frequency) possesses the cosine distribution of currents around its circumference, therefore creating a homogeneous magnetic field inside its volume. For high-pass type, two more resonant modes can be found (anti-rotational and co-rotational) which are associated with the resonance of the endrings themselves [2]. After the invention of the birdcage resonators, the description of their resonance modes relied mainly on analytical solution and lumped-element analysis [3]. With the increasing frequency of operation, the parasitic behaviour (such as radiation) became important, and the purely lumped-element approach was not accurate anymore. Then, the methods of numerical analysis started to be used more frequently for birdcage resonator diagnostics [4].

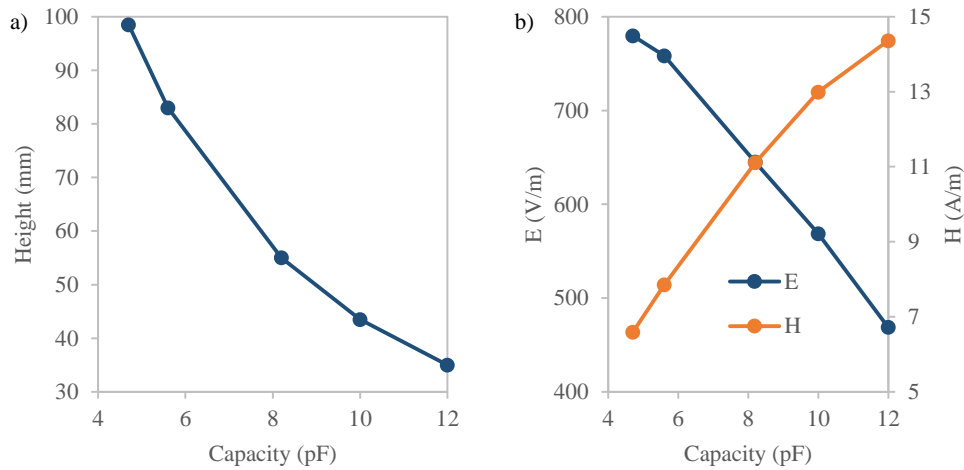


**Figure 1:** Different types of birdcage resonators - a) low-pass, b) high-pass, c) band-pass. Adapted from [5].

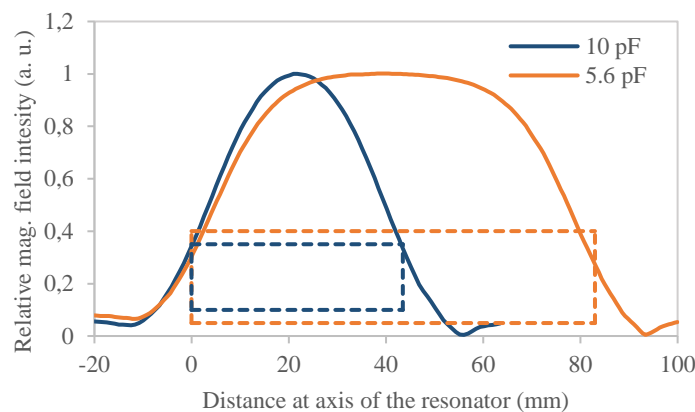
## 2. NUMERICAL ANALYSIS

Five high-pass birdcage resonators are analysed within this paper. The resonators differ in the capacity of inserted capacitors (Johanson Technology: SMD 0805, High-Q, 250 V), which ranges from 4.7 pF to 12 pF. The parameters of the examined resonators were simulated using Ansys HFSS. The diameter was fixed to 40 mm for all resonators and their height was adjusted to meet the frequency of 433 MHz of the fundamental resonance mode. The results of the analysis were already published and are described in detail in [6]. From the perspective of this paper, the most important is the information about the magnetic field in the middle of the resonator and the calculated resonator's height (Figure 2). The analysis results show clearly, that with increasing inserted capacity, therefore, decreasing height of the resonators, the magnetic field intensity in their centre is increasing.

However, the magnetic field is homogeneous only in the central part of the resonator. Closer to the endrings, the intensity of the field is decreasing. Therefore, the resonators with low aspect ratio (the resonators with the inserted capacity of 10 and 12 pF in this paper) are not suitable for MRI, since the volume of homogeneous magnetic field is restricted. For alternative application of the birdcage resonators (such as plasma processing [7]) these resonators might still be useful. The comparison of two examples with different volumes of homogeneous magnetic field is shown in Figure 3.



**Figure 2:** a) Height, b) Electric and magnetic field intensity in dependence on the inserted capacity. Adapted from [6].



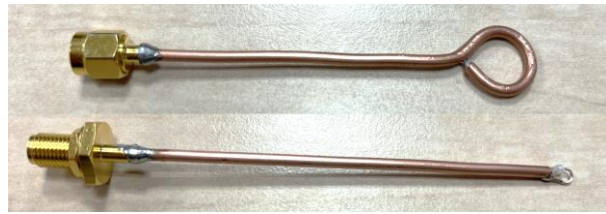
**Figure 3:** An example of the magnetic field distribution for the resonators with different heights (with 5.6 and 10 pF capacitors). The resonators are represented with dashed lines. Adapted from [6].

## 3. EXPERIMENTAL SETUP

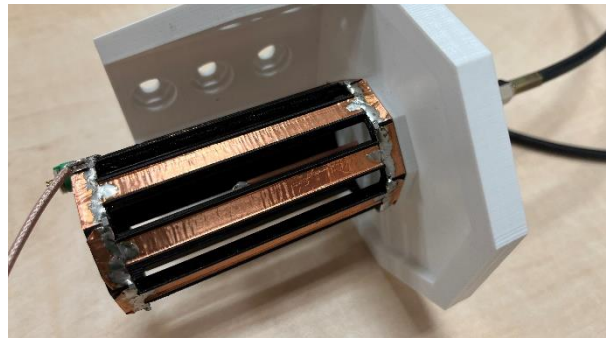
Two magnetic probes were created of semi-rigid coaxial cable (RG-405CU, QAXIAL). The first probe was created by bending a whole coaxial cable with respect to its minimum bending diameter, and the centre conductor was then soldered to the outer shield to create a closed loop with a diameter

approximately of 10 mm. The outer conductor shields the loop, which helps to reduce the response of the probe to the electric field created by the resonator. The second probe was fabricated with respect to higher spatial resolution. The outer shield was removed, and the loop was created by the central conductor only. The loop has diameter of approx. 3.5 mm. Both probes were completed with an SMA connector (Figure 4). The previously prepared birdcage resonators [6] were matched to  $50\ \Omega$  by a variable capacitor in series with the resonator, and together with the probes were placed to a 3D printed fixture, which ensures the proper centring of the probe inside the birdcage resonator (Figure 5).

The matched birdcage resonators and the probe were connected to vector network analyser (VNA) Rohde & Schwarz ZVL (0–6 GHz), and the magnetic field intensity was measured as the  $S_{21}$  parameter between the ports of the VNA. The  $S_{21}$  parameter was measured after fine-tuning of the adjustable capacitor (insertion of the probe caused a slight change in the impedance) and the maximal value was recorded. The measurement was done in the middle of the height of the resonator, and the plane of the probe loop was aligned perpendicular to the feeding point.



**Figure 4:** Magnetic field probes

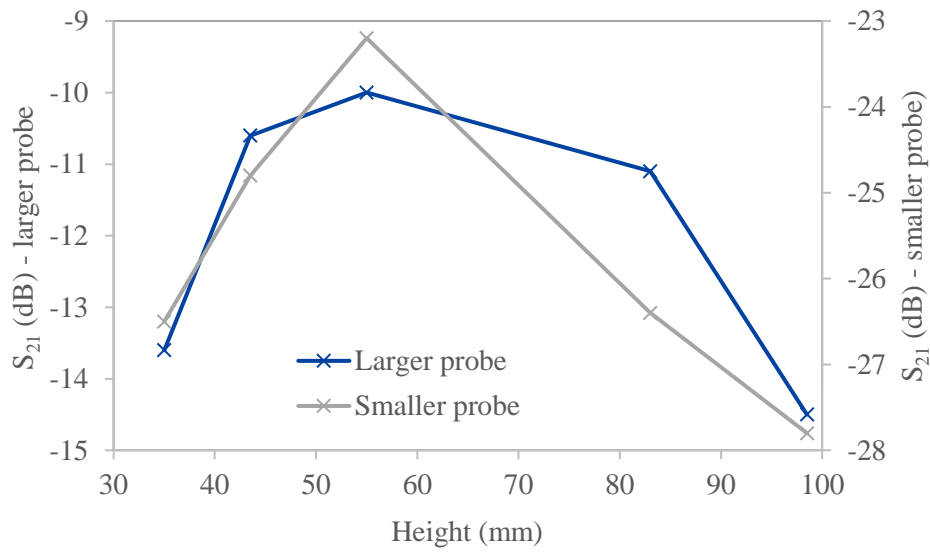


**Figure 5:** Measurement setup (the probe is located in the centre of the resonator)

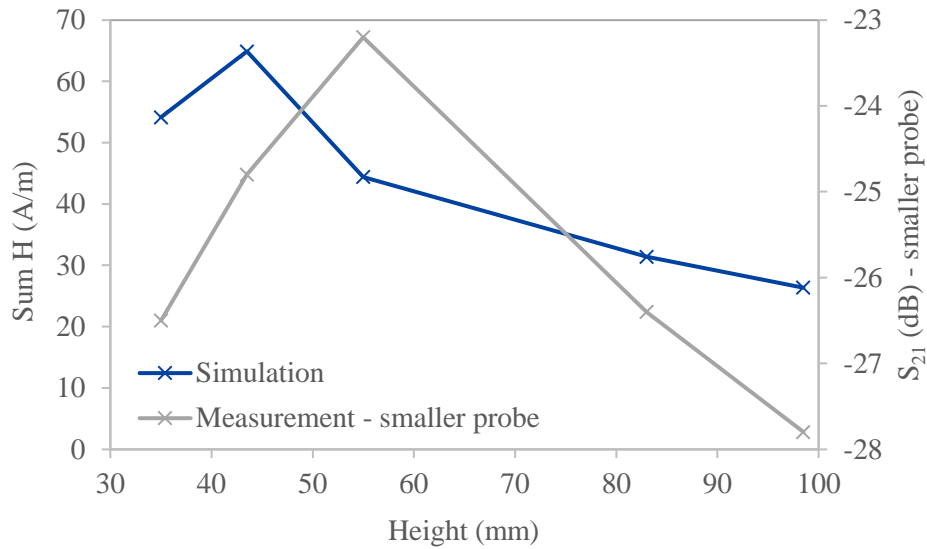
#### 4. RESULTS AND DISCUSSION

Figure 6 shows the dependence of the measured  $S_{21}$  parameters on the height of the resonators (see Figure 2 a)) for comparison with the capacity of the inserted capacitors. Both measurements show that the resonator with 8.2 pF inserted capacity possesses the most intense magnetic field in the middle of its height. The results are in accordance with the assumptions from the analytical calculation, which states that “the magnetic field amplitude at the centre is maximum when the length to diameter ratio is equal to  $\sqrt{2}$ ” [2]. For 40 mm of diameter, the optimal height is approx. 56.6 mm, which corresponds to the maximum in Figure below, which is located at 55 mm.

Besides the agreement of the analytical solution and the measured values, a discrepancy between the results of the simulation (Figure 2 b)) and measurement was found. According to the simulation, the magnetic field intensity should decrease with the increasing height of the resonators across the whole range of evaluated heights. However, the results of the simulation show only the maximum of magnetic field intensity in the centre of the resonator, while the probes are measuring the proximate surroundings as well. To diagnose this discrepancy, the results of the simulation were interpreted in a different way. The magnetic field intensity was cumulated (a sum was calculated over 201 points) in the range -0.5 mm to 0.5 mm around the centre of the resonator. The calculated values are compared to the  $S_{21}$  parameters measured with the smaller probe in Figure 7. There is still a discrepancy between the simulation and measurement, however the trends agree. The discrepancy is probably caused by non-ideal behaviour of the fabricated prototypes.



**Figure 6:** Measured  $S_{21}$  parameters for two magnetic field probes



**Figure 7:** Comparison of simulation and measurement using the smaller probe

## 5. CONCLUSION

The measurement with both probes confirms that there exists an optimal aspect ratio of the birdcage resonators, with respect to the created magnetic field intensity. The best results were achieved for the birdcage with a height of 55 mm. The probe with the smaller loop at its end exhibits lower sensitivity to the magnetic field, however, the measured intensity is steeper decreasing towards the non-optimal aspect ratios.

## ACKNOWLEDGMENT

This work has been supported by grant FEKT-K-22-7732 realised within the project Quality Internal Grants of BUT (KInG BUT), Reg. No. CZ.02.2.69 / 0.0 / 0.0 / 19\_073 / 0016948, which is financed from the OP RDE.

## REFERENCES

- [1] C. Hayes, W. Edelstein, J. Schenck, O. Mueller and M. Eash, "An efficient, highly homogeneous radiofrequency coil for whole-body NMR imaging at 1.5 T", *Journal of Magnetic Resonance* (1969), vol. 63, no. 3, pp. 622-628, 1985.
- [2] J. Mispelter, M. Lupu, and A. Briguet, "NMR Probeheads for Biophysical and Biomedical Experiments"
- [3] R. Pascone, et al., "Generalized electrical analysis of low-pass and high-pass birdcage resonators", *Magnetic Resonance Imaging*, vol. 9, no. 3, pp. 395-408, 1991.
- [4] N. Gurler and Y. Z. Ider, "Numerical methods and software tools for simulation, design, and resonant mode analysis of radio frequency birdcage coils used in MRI", *Concepts in Magnetic Resonance Part B: Magnetic Resonance Engineering*, vol. 45, no. 1, pp. 13-32, 2015.
- [5] S. F. Ahmad, Y. C. Kim, I. C. Choi, and H. D. Kim, "Recent Progress in Birdcage RF Coil Technology for MRI System", *Diagnostics*, vol. 10, no. 12, 2020.
- [6] K. Jurik, P. Drexler, D. Nespor, and L. Havlicek, "Parametric Optimization of a Birdcage Resonator for Low-pressure Plasma Excitation", pp. 455-461, Nov. 2021.
- [7] P. Guittienne, R. Jacquier, B. Pouradier Duteil, A. A. Howling, R. Agnello, and I. Furno, "Helicon wave plasma generated by a resonant birdcage antenna: magnetic field measurements and analysis in the RAID linear device", *Plasma Sources Science and Technology*, vol. 30, no. 7, Jul. 2021.

# Vanishing solutions of a second-order discrete non-linear equation of Emden-Fowler type.

J. Diblík<sup>1</sup> and E. Korobko<sup>1</sup>

<sup>1</sup>Brno University of Technology  
Faculty of Electrical Engineering and Communication  
Department of Mathematics  
Technická 3058/10  
616 00 Brno, Czech Republic

E-mail: [diblik@vut.cz](mailto:diblik@vut.cz), [222594@vut.cz](mailto:222594@vut.cz)

**Abstract**—The paper discusses a discrete equation of an Emden-Fowler type  $\Delta^2 v(k) = -k^3 (\Delta v(k))^3$  where  $v$  is a dependent variable,  $k$  is an integer-valued independent variable,  $\Delta v$  and  $\Delta^2 v$  are the first and second-order forward differences of  $v$ , respectively. The paper aims to prove the existence of a nontrivial and vanishing solution for  $k \rightarrow \infty$ . The equation is transformed into a system of two first-order difference equations, which makes it possible to apply previously known results when investigating the system.

**Keywords**—difference equation, Emden-Fowler type equation, asymptotic behaviour

## 1. INTRODUCTION

We consider the second-order difference equation

$$\Delta^2 v(k) = -k^3 (\Delta v(k))^3, \quad (1)$$

where the independent variable  $k \in \mathcal{Z}_{k_0}^\infty = \{k_0, k_0 + 1, \dots\}$ ,  $k_0$  is a positive integer, the dependent variable  $v: \mathcal{Z}_{k_0}^\infty \rightarrow \mathbb{R}$ ,  $\Delta v(k) = v(k+1) - v(k)$  and  $\Delta^2 v(k) = v(k+2) + 2v(k+1) - v(k)$  are the first and second-order forward differences. In the paper we prove the existence of a nontrivial solution  $v(k)$  such that  $\lim_{k \rightarrow \infty} v(k) = 0$ . It is known that second-order differential equations of Emden-Fowler type can have so called blow-up solutions (that is, solutions with a vertical asymptote). For example, Emden-Fowler differential equation  $y''(x) = y^3(x)$  has blow-up solutions  $y(x) = \pm \sqrt{2}/(x+K)$  where  $K$  is an arbitrary (but fixed) constant. By suitable transformation, this equation can be transferred to another one replacing the problem on the existence of blow-up solutions by one on the existence of vanishing solutions when  $t \rightarrow \infty$  for a new differential equation. Equation (1) is a discrete variant of such a differential equation. The paper aims to prove that there are solutions  $v = v(k)$  to equation (1) such that  $\lim_{k \rightarrow \infty} v(k) = 0$ . Note that differential equations of Emden-Fowler type are widely investigated because they serve as mathematical models of some phenomena appearing in cosmology, astrophysics and other sciences [1].

## 2. PRELIMINARIES

**Definition 1.** Let  $f, g: \mathcal{Z}_{k_0}^\infty \rightarrow \mathbb{R}$ . We denote  $f = O(g)$ , if there exists an index  $k_1 \geq k_0$  and a constant  $M > 0$  such that  $|f(k)| \leq M|g(k)|$ ,  $\forall k \in \mathcal{Z}_{k_1}^\infty$ .

For further investigations we need to refer to main results from [2] and [3]. The following preliminaries are taken from [2, 3]. Consider a system of discrete equations

$$\Delta Y(k) = F(k, Y(k)), \quad (2)$$

where  $F: \mathcal{Z}_{k_0}^\infty \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $Y = (Y_1, \dots, Y_n) \in \mathbb{R}^n$ . Throughout the paper assume that  $F$  is continuous with respect to  $Y$ . Define a set  $\Omega \subset \mathcal{Z}_{k_0}^\infty \times \mathbb{R}^n$  as

$$\Omega = \{(k, Y): k \in \mathcal{Z}_{k_0}^\infty, b_i(k) < Y_i < c_i(k), i = 1, \dots, n\}$$



where  $b_i, c_i: \mathcal{Z}_{k_0}^\infty \rightarrow \mathbb{R}$  are auxiliary functions such that  $b_i(k) < c_i(k), k \in \mathcal{Z}_{k_0}^\infty, i = 1, \dots, n$ . Let us define functions  $B_i, C_i: \mathcal{Z}_{k_0}^\infty \times \mathbb{R} \rightarrow \mathbb{R}, i = 1, \dots, n$  by formulas  $B_i(k, Y) := -Y_i + b_i(k), C_i(k, Y) := Y_i - c_i(k)$  and sets

$$\begin{aligned}\Omega_B^i &= \{(k, Y): k \in \mathcal{Z}_{k_0}^\infty, B_i(k, Y) = 0, B_j(k, Y) \leq 0, C_p(k, Y) \leq 0, \forall j, p = 1, \dots, n, j \neq i\}, \\ \Omega_C^i &= \{(k, Y): k \in \mathcal{Z}_{k_0}^\infty, C_i(k, Y) = 0, B_j(k, Y) \leq 0, C_p(k, Y) \leq 0, \forall j, p = 1, \dots, n, p \neq i\}\end{aligned}$$

for every  $i = 1, \dots, n$ .

**Definition 2.** The set  $\Omega$  is called the regular polyfacial set with respect to the discrete system (2) if

$$b_i(k+1) - b_i(k) < F_i(k, Y) < c_i(k+1) - b_i(k), \quad (3)$$

for every  $i = 1, \dots, n$  and every  $(k, Y) \in \Omega_B^i$  and if

$$b_i(k+1) - c_i(k) < F_i(k, Y) < c_i(k+1) - c_i(k), \quad (4)$$

for every  $i = 1, \dots, n$  and every  $(k, Y) \in \Omega_C^i$ .

Below we use the sets

$$\Omega(k) = \{(k, Y), Y \in \mathbb{R}^n, b_i(k) < Y_i < c_i(k), i = 1, \dots, n\},$$

$$\Omega_i(k) = \{(w), b_i(k) < w < c_i(k)\}, \quad i = 1, \dots, n.$$

We need the following minor modification of Theorem 4 in [2] where it is proved by Liapunov-like reasonings.

**Theorem 1.** [2, Theorem 4] Let  $\Omega$  be regular with respect to the discrete system (2) and let the function

$$G_i(w) := w + F_i(k, Y_1, \dots, Y_{i-1}, w, Y_{i+1}, \dots, Y_n)$$

be monotone on  $\overline{\Omega}_i(k)$  for every fixed  $k \in \mathcal{Z}_{k_0}^\infty$ , each fixed  $i \in \{1, \dots, n\}$ , and every fixed  $(Y_1, \dots, Y_{i-1}, Y_{i+1}, \dots, Y_n)$  such that  $(k, Y_1, \dots, Y_{i-1}, w, Y_{i+1}, \dots, Y_n) \in \Omega$ . Then, every initial problem  $Y(k_0) = Y^*$  with  $Y^* \in \Omega(k_0)$  defines the solution  $Y = Y^*(k)$  of the discrete system (2) satisfying the relation  $Y^*(k) \in \Omega(k)$  for every  $k \in \mathcal{Z}_{k_0}^\infty$ .

The following theorem is a slight modification of [3, Theorem 1] (see [4, Theorem 2] also) where it is proved by a topological method.

**Theorem 2.** Assume that the function  $F(k, Y)$  satisfies (2) and is continuous with respect to  $Y$ . Let inequality

$$F_i(k, Y) < b_i(k+1) - b_i(k) \quad (5)$$

hold for every  $i = 1, \dots, n$  and every  $(k, Y) \in \Omega_B^i$ . Let, moreover, inequality

$$F_i(k, Y) > c_i(k+1) - c_i(k) \quad (6)$$

hold for every  $i = 1, \dots, n$  and every  $(k, Y) \in \Omega_C^i$ . Then, there exists a solution  $Y = Y(k), k \in \mathcal{Z}_{k_0}^\infty$  of system (2) satisfying the inequalities

$$b_i(k) < Y_{i-1}(k) < c_i(k)$$

for every  $k \in \mathcal{Z}_{k_0}^\infty$  and  $i = 1, \dots, n$ .

### 3. AUXILIARY DISCRETE SYSTEM

Below we use the change of variables

$$v(k) = \pm \frac{\sqrt{2}}{k} (1 + Y_1(k)), \quad (7)$$

$$\Delta v(k) = \pm \Delta \left( \frac{\sqrt{2}}{k} \right) \cdot (1 + Y_2(k)), \quad (8)$$

$$\Delta^2 v(k) = \pm \Delta^2 \left( \frac{\sqrt{2}}{k} \right) \cdot (1 + Y_3(k)) \quad (9)$$

where  $Y_i(k)$ ,  $i = 1, 2, 3$  are new dependent variables. Formulas (7)–(9) transform (1) to the system of discrete equations

$$\begin{cases} \Delta Y_1(k) = \frac{1}{k} \cdot (Y_1(k) - Y_2(k)), \\ \Delta Y_2(k) = \left( \Delta \left( \frac{1}{k+1} \right) \right)^{-1} \Delta^2 \left( \frac{1}{k} \right) \cdot \left( -2k^3 \left( \Delta^2 \left( \frac{1}{k} \right) \right)^{-1} \cdot \left( \Delta \left( \frac{1}{k} \right) \right)^3 \cdot (1 + Y_2(k))^3 - 1 - Y_2(k) \right). \end{cases} \quad (10)$$

or, assuming  $|Y_2(k)| < 1$  and simplifying expression in the second equation, to

$$\begin{cases} \Delta Y_1(k) = \frac{1}{k} \cdot (Y_1(k) - Y_2(k)), \\ \Delta Y_2(k) = \left( -\frac{2}{k} + O\left(\frac{1}{k^3}\right) \right) \left( 2Y_2(k) + 3Y_2^2(k) + Y_2^3(k) + O\left(\frac{1}{k^2}\right) + O\left(\frac{Y_2(k)}{k^2}\right) \right). \end{cases} \quad (11)$$

#### 4. MAIN RESULTS

Consider the second equation in system (11) separately, that is, consider the equation

$$\Delta Y_2(k) = \left( -\frac{2}{k} + O\left(\frac{1}{k^3}\right) \right) \left( 2Y_2(k) + 3Y_2^2(k) + Y_2^3(k) + O\left(\frac{1}{k^2}\right) \right). \quad (12)$$

Let  $b_i(k) := -\varepsilon_i$ ,  $c_i(k) := \gamma_i$  where  $\varepsilon_i, \gamma_i$  are fixed positive numbers. Set

$$B_i(k, Y_1, Y_2) := -Y_i - \varepsilon_i, \quad C_i(k, Y_1, Y_2) := Y_i - \gamma_i, \quad i = 1, 2$$

Auxiliary sets  $\Omega_B^2, \Omega_C^2$  are reduced as follows

$$\Omega_B^2 = \{(k, Y_2) : k \in \mathcal{Z}_{k_0}^\infty, Y_2 = -\varepsilon_2\}, \quad \Omega_C^2 = \{(k, Y_2) : k \in \mathcal{Z}_{k_0}^\infty, Y_2 = \gamma_2\}.$$

We are going to apply Theorem 1 to the equation (12). This means that we need to show that (3) and (4) hold for  $i = 2$  where

$$F_2(k, Y_1, Y_2) = F_2(k, Y_2) = \left( -\frac{2}{k} + O\left(\frac{1}{k^3}\right) \right) \left( 2Y_2(k) + 3Y_2^2(k) + Y_2^3(k) + O\left(\frac{1}{k^2}\right) \right). \quad (13)$$

Inequality (3) is now modified to

$$0 = b_2(k+1) - b_2(k) < F_2(k, Y_2)|_{(k, Y_2) \in \Omega_B^2} < \gamma_2 + \varepsilon_2. \quad (14)$$

The function

$$F_2(k, Y_2)|_{(k, Y_2) \in \Omega_B^2} = F_2(k, -\varepsilon_2) = \left( -\frac{2}{k} + O\left(\frac{1}{k^3}\right) \right) \left( -2\varepsilon_2 + 3\varepsilon_2^2 - \varepsilon_2^3 + O\left(\frac{1}{k^2}\right) \right)$$

assumes positive values for all sufficiently large  $k$  if

$$2\varepsilon_2 - 3\varepsilon_2^2 + \varepsilon_2^3 = \varepsilon_2(\varepsilon_2 - 1)(\varepsilon_2 - 2) > 0$$

that is if  $\varepsilon_2 \in (0, 1) \cup (2, +\infty)$ . Because, in the derivation of system (11), we assumed  $|Y_2(k)| < 1$ , only values  $\varepsilon_2 \in (0, 1)$  can be used. The left inequality in (3) holds. The right inequality in (3) holds for all

sufficiently large  $k$  as well since the function  $F_2(k, \varepsilon_2)$  is vanishing.

Now, we show that (4) holds for  $i = 2$ . This inequality reduces to

$$-\varepsilon_2 - \gamma_2 = b_2(k+1) - c_2(k) < F_2(k, Y_2)|_{(k, Y_2) \in \Omega_C^2} < \gamma_2 - \gamma_2 = 0, \quad (15)$$

where

$$F_2(k, Y_2)|_{(k, Y_2) \in \Omega_C^2} = F_2(k, \gamma_2) = \left(-\frac{2}{k} + O\left(\frac{1}{k^3}\right)\right) \left(2\gamma_2 + 3\gamma_2^2 + \gamma_2^3 + O\left(\frac{1}{k^2}\right)\right).$$

The function  $F_2(k, \gamma_2)$  is negative for  $\gamma_2 \in (0, 1)$  and for all sufficiently large  $k$ . Therefore the right inequality in (15) holds. The left inequality in (15) holds too because the function  $F_2(k, \gamma_2)$  is vanishing as  $k \rightarrow \infty$ . Finally, we need to show that the function

$$G_2(w) := w + F_2(k, w) \quad (16)$$

is monotone on

$$\overline{\Omega}_2(k) = \{(w) : w \in \mathbb{R}, b_2(k) \leq w \leq c_2(k)\} = \{(w) : w \in \mathbb{R}, -\varepsilon_2 \leq w \leq \gamma_2\}$$

for every fixed  $k \in \mathcal{Z}_{k_0}^\infty$ . We will verify the monotony property by computing  $G'(w)$ . Since a direct computation of the derivative of the function  $F_2(k, w)$  expressed by (13) is not possible, we use its definition by the right-hand side of the second equation in system (10). We have

$$G_2(w) = w + F_2(k, w) = w + \left(\Delta\left(\frac{1}{k+1}\right)\right)^{-1} \Delta^2\left(\frac{1}{k}\right) \cdot \left(-2k^3 \left(\Delta^2\left(\frac{1}{k}\right)\right)^{-1} \cdot \left(\Delta\left(\frac{1}{k}\right)\right)^3 \cdot (1+w)^3 - 1 - w\right)$$

and, therefore,

$$G'_2(w) = 1 + \left(\Delta\left(\frac{1}{k+1}\right)\right)^{-1} \Delta^2\left(\frac{1}{k}\right) \cdot \left(-6k^3 \left(\Delta^2\left(\frac{1}{k}\right)\right)^{-1} \cdot \left(\Delta\left(\frac{1}{k}\right)\right)^3 \cdot (1+w)^2 - 1\right).$$

Because

$$\left(\Delta\left(\frac{1}{k+1}\right)\right)^{-1} \Delta^2\left(\frac{1}{k}\right) = -\frac{2}{k} + O\left(\frac{1}{k^3}\right) \quad \text{and} \quad -2k^3 \left(\Delta^2\left(\frac{1}{k}\right)\right)^{-1} \left(\Delta\left(\frac{1}{k}\right)\right)^3 = 1 + O\left(\frac{1}{k^2}\right),$$

we have

$$G'_2(w) = 1 + \left(-\frac{2}{k} + O\left(\frac{1}{k^3}\right)\right) \cdot \left(3 \left(1 + O\left(\frac{1}{k^2}\right)\right) \cdot (1+w^2) - 1\right)$$

and, for all sufficiently large  $k$ ,  $G'(w) > 0$ . The function  $G_2$  is monotone, Theorem 1 is applicable and, therefore, there exists a solution  $Y_2 = Y_2^*(k)$  to equation (12) satisfying inequality

$$-\varepsilon_2 < Y_2^*(k) < \gamma_2, \quad k \in \mathcal{Z}_{k_0}^\infty \quad (17)$$

where  $k_0$  is sufficiently large and positive numbers  $\varepsilon_2, \gamma_2, \varepsilon_2 < 1, \gamma_2 < 1$  are fixed. Note that this solution is not trivial. Now we use Theorem 2 to analyse the first equation of the system (11), that is, the equation

$$\Delta Y_1(k) = (Y_1(k) - Y_2(k)) / k. \quad (18)$$

We have proved that there exists a solution  $Y_2 = Y_2^*(k)$  of the second equation in the system (11) with asymptotic behaviour described by inequality (17). Let in (18) be such a solution assumed. Then

$$\Delta Y_1(k) = (Y_1(k) - Y_2^*(k)) / k.$$

Setting

$$F_1(k, Y_1, Y_2) = F_1(k, Y_1) = (Y_1(k) - Y_2^*(k)) / k$$

in Theorem 2, the auxiliary sets  $\Omega_B^1, \Omega_C^1$  are reduced to

$$\Omega_B^1 = \{(k, Y_2): k \in \mathcal{Z}_{k_0}^\infty, Y_1 = -\varepsilon_1\}, \quad \Omega_C^1 = \{(k, Y_2): k \in \mathcal{Z}_{k_0}^\infty, Y_1 = \gamma_1\}.$$

Then, for  $k \in \mathcal{Z}_{k_0}^\infty$ , inequality (5) becomes

$$F_1(k, Y_1)|_{(k, Y_1) \in \Omega_B^1} = F_1(k, -\varepsilon_1) = (-\varepsilon_1 - Y_2^*(k)) / k < b_1(k+1) - b_1(k) = 0. \quad (19)$$

Due to (17), we derive

$$(-\varepsilon_1 - Y_2^*(k)) / k < (-\varepsilon_1 + \varepsilon_2) / k$$

and inequality (19) will hold if  $\varepsilon_2 < \varepsilon_1$ . For  $k \in \mathcal{Z}_{k_0}^\infty$ , inequality (6) will be now

$$F_1(k, Y_1)|_{(k, Y_1) \in \Omega_C^1} = F_1(k, \gamma_1) = (\gamma_1 - Y_2^*(k)) / k > c_1(k+1) - c_1(k) = 0. \quad (20)$$

By (17), we derive

$$(\gamma_1 - Y_2^*(k)) / k > (\gamma_1 - \gamma_2) / k$$

and inequality (20) will hold if  $\gamma_1 > \gamma_2$ . Theorem 2 is applicable and, therefore, there exists a solution  $Y_1 = Y_1^*(k)$  to equation (19) satisfying inequality  $-\varepsilon_1 < Y_1^*(k) < \gamma_1, k \in \mathcal{Z}_{k_0}^\infty$  where  $k_0$  is sufficiently large and positive number and  $\varepsilon_2 < \varepsilon_1 < 1, \gamma_2 < \gamma_1 < 1$  are fixed. Note that this solution is not trivial because  $Y_2^*(k)$  is not trivial. Summarizing the above investigations, we conclude that the below theorem holds.

**Theorem 3.** *Let  $\varepsilon_i, \gamma_i, i = 1, 2$  be fixed positive numbers such that  $\varepsilon_2 < \varepsilon_1 < 1, \gamma_2 < \gamma_1 < 1$ . Then there exists a solution  $Y(k) = Y^*(k) = (Y_1^*(k), Y_2^*(k))$  to the system (11) such that*

$$-\varepsilon_i < Y_i^*(k) < \gamma_i, \quad i = 1, 2, \quad \forall k \in \mathcal{Z}_{k_0}^\infty \quad (21)$$

*provided that  $k_0$  is sufficiently large.*

## 5. CONCLUSION

By transformation (7), Theorem 3 implies that there exist two solutions  $v = v_\pm(k)$  of equation (1) such that

$$v_\pm(k) = \pm \sqrt{2} k^{-1} (1 + Y_1^*(k)), \quad \forall k \in \mathcal{Z}_{k_0}^\infty. \quad (22)$$

From (22) we have  $\lim_{k \rightarrow \infty} v_\pm(k) = 0$ . Moreover, from (21) we derive

$$\sqrt{2}(1 - \varepsilon_1) < kv_+(k) < \sqrt{2}(1 + \gamma_1), \quad -\sqrt{2}(1 + \gamma_1) < kv_-(k) < -\sqrt{2}(1 - \varepsilon_1).$$

## ACKNOWLEDGMENT

This research has been supported by the project of specific university research at Brno University of Technology, Faculty of Electrical Engineering and Communication, FEKT-S-20-6225.

## REFERENCES

- [1] R. Bellman, *Stability Theory in Differential Equations*, New York: Dover Publications, Inc., 2008.
- [2] J. Diblík, "Asymptotic behaviour of solutions of systems of discrete equations via Liapunov type technique", *Comput. Math. Appl.*, vol. 45, pp. 1041–1057, 2003, doi: [10.1016/S0898-1221\(03\)00082-8](https://doi.org/10.1016/S0898-1221(03)00082-8).
- [3] J. Diblík, "Discrete retract principle for systems of discrete equations", *Comput. Math. Appl.*, 42, pp. 515–528, 2001, doi: [10.1016/S0898-1221\(01\)00174-2](https://doi.org/10.1016/S0898-1221(01)00174-2).
- [4] J. Diblík, "Asymptotic behavior of solutions of discrete equations", *Funct. Differ. Equ.*, 11, (1–2), pp. 37–48, 2004.

# Testing of UWB Sensors For Use in Multi-robot Relative Localization Systems and Design of UWB Plugin for ROS/Gazebo Simulation Environment

P. Raichl<sup>1</sup>, P. Marcon<sup>1</sup>

<sup>1</sup>Faculty of Electrical Engineering and Communication, Brno University of Technology, 61600 Brno, Czech Republic

E-mail: [xraich02@vut.cz](mailto:xraich02@vut.cz), [marcon@vut.cz](mailto:marcon@vut.cz)

**Abstract**—This paper studies the possibility of using UWB sensors DWM1000 as a part of relative localization systems and a part of local avoidance systems in multi-robot systems, especially UAV swarms. The first part of the paper presents the behavior of UWB sensors DWM1000 in real, various conditions/environments. The second part of this paper describes ROS/Gazebo UWB plugin for relative localization. This plugin was created based on the first part of this paper and it is integrated with an open-source platform for UAVs from the Multi-robot Systems (MRS) group at the Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague. A Simple simulation of a multi-robot system with the designed, integrated plugin is shown.

**Keywords**—UWB, DWM1000, UAV Swarms, ROS, Gazebo, UWB plugin, UWB simulation, relative localization

## 1. INTRODUCTION

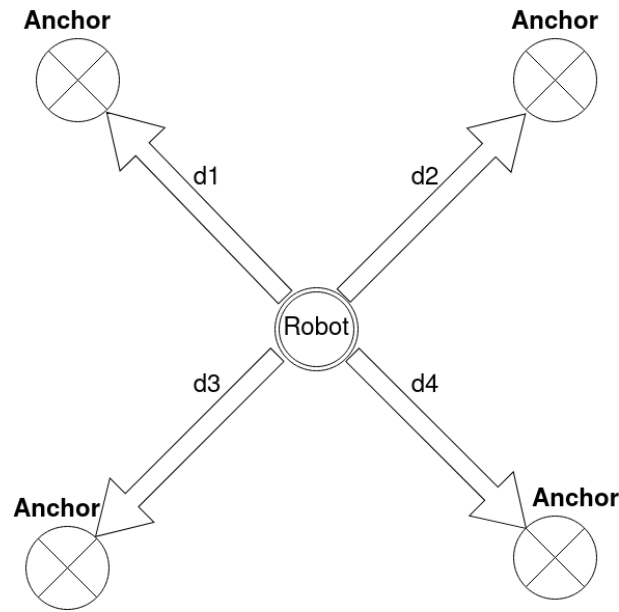
The coordination of multi-robot systems is a quite challenging issue and it is much more challenging when these robots are UAVs(Unmanned Aerial Vehicles), because of the different heights in which these robots can operate. To create the best movement action, the more global knowledge about the near environment every robot in the swarm has, the easier the problem is. The knowledge about the near environment, especially own position and velocity and position and velocity of coordination's partners, can be created by different types of sensors, cameras [1], infra-red emitters and receivers in specific layout [2] or UV LED markers [3].

In recent years, UWB technology has started to be used in indoor localization systems. The goal of this technology was to achieve high localization accuracy on distances higher than other technologies which can be used for localization systems such as WiFi, Bluetooth, and lower interference of signals. The Typical configuration of localization systems based on UWB modules is shown in fig. 1. This configuration consists of several anchors, i. e. UWB modules whose position is known and static, i.e. these modules are not moving, and 1 or more tags, UWB modules which are placed on moving robots and we are calculating their position. The minimal number of used UWB modules is 4(3 anchors, 1 tag), so the location of robots in the environment can be calculated by trilateration[4].

We intend to use UWB modules as a part of a relative localization system or as a part of a local avoidance system in multi-robot systems, especially UAV swarms. In this configuration, every robot measures its distance from others. As most of the papers are aimed at the study of UWB systems in an indoor environment, this paper brings knowledge of how UWB modules perform in indoor and outdoor conditions in an environment without obstacles between UWB modules.

Because building and testing of a relative positioning system for multi-robot systems is a quite difficult task where collisions between robots can occur, it is important to test every new feature of this system in simulation. Because we are using *ctu mrs platform*<sup>1</sup> [5] for our robots which is programmed in ROS and uses Gazebo environment for simulation, we need to design and program UWB plugin for this ecosystem because UWB plugin which could be used for simulation of relative positioning system doesn't exist. This plugin is also described in this paper.

<sup>1</sup>Open source platform for UAVs from Multi-robot Systems (MRS) group at Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague



**Figure 1:** Classic UWB localization configuration

## 2. MATERIALS AND METHODS

Nowadays, several manufacturers offer UWB modules, for example, NXP, Apple, CEVA, Qorvo. For our experiments, Qorvo DWM1000 was chosen. It is a module that integrates the antenna and everything needed for using DW1000IC. This module uses the SPI bus for communication with a superior microcontroller, where the code is running. As a superior microcontroller we used ESP8266 on the WeMos D1 development kit.

UWB sensors were studied in 3 different environments: Anechoic laboratory, Classic laboratory, Outdoor open space without obstacles. Conditions of tests are described below for each environment.

**Anechoic laboratory and Classic laboratory** This is an indoor environment and it is used primarily to verify the function of used UWB sensors and to test the accuracy of sensors on short distances, i.e. from a few cm to a few m. The accuracy was determined by comparing the measured distances with known distances. Between UWB sensors there were no obstacles. The difference between these two kinds of laboratories is that the Anechoic laboratory has special material on the walls which prevents signal reflections. On the contrary, in the classic laboratory, there are a lot of equipment and walls which can cause signal reflections.

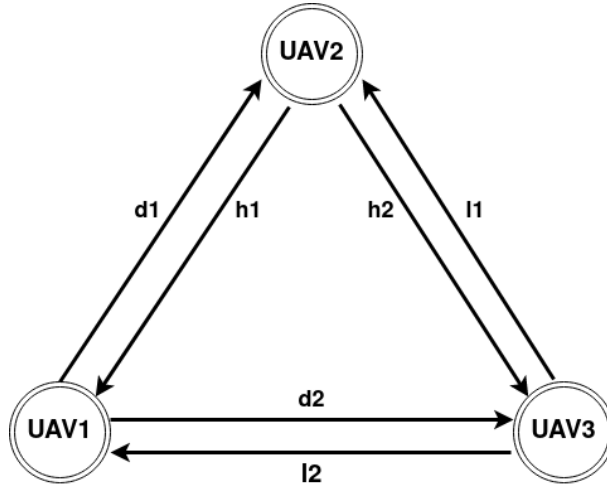
**Outdoor open space** This test was performed on the tennis court. The main purpose of this test was to find out if the accuracy from the laboratory can be achieved in outdoor space without obstacles. For this purpose, a tennis court was used.

**UWB for relative positioning in multi-robot systems** This approach is different from the typical configuration shown in fig. 1. In this approach every UWB module acts as a tag and anchor at the same time, i. e. every module calculates its distance from the others and this information can then be used for relative positioning in formation control or as a part of local collision avoidance system. This configuration is presented in fig. 2.

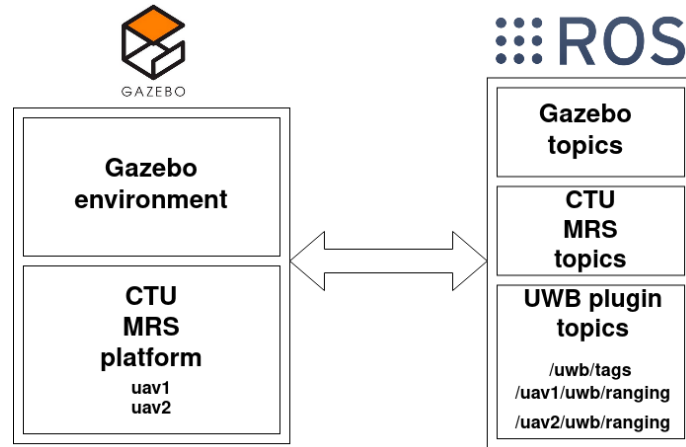
The information gained from measurements in different environments is used to design UWB Gazebo plugin. The plugin simulates the reception of UWB ranging measurements from different tags/anchors. The structure of the UWB plugin is shown in fig. 3. The plugin publishes into 2 topics: `/uwb/tags` - publishes its name. `/uav/uwb` - publishes distances to other UWB modules. The ranging message is similar to the output of UWB sensor DWM1000, see fig. 7. Every UWB module has to be paired with a robot by setting these parameters in *urdf file*<sup>2</sup>: model name, namespace..., see fig. 4.

<sup>2</sup>The Unified Robotic Description Format (URDF) is an XML file format used in ROS to describe all elements of a robot.





**Figure 2:** UWB Relative positioning approach



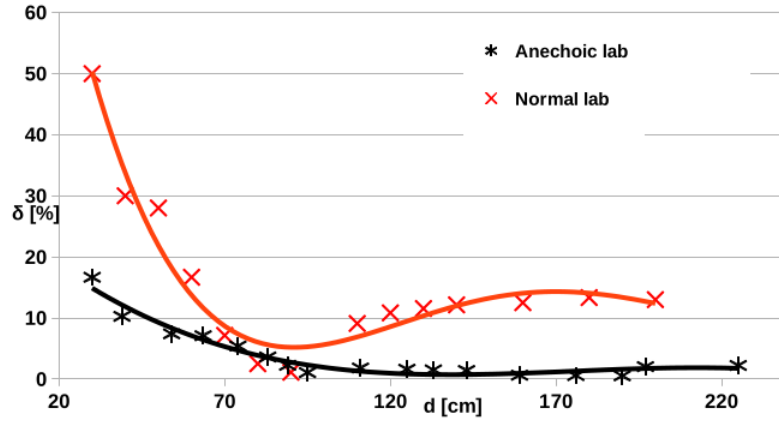
**Figure 3:** Structure of environment - topics

```
<gazebo>
  <plugin name='libuwb_plugin' filename='libuwb_plugin.so'>
    <update_rate>20</update_rate>
    <frequency>3.9936e9</frequency>
    <tag_link>right_tag</tag_link>
    <robot_namespace>${namespace}</robot_namespace>
  </plugin>
</gazebo>
```

**Figure 4:** Instalation of plugin - urdf file

### 3. RESULTS AND DISCUSSION

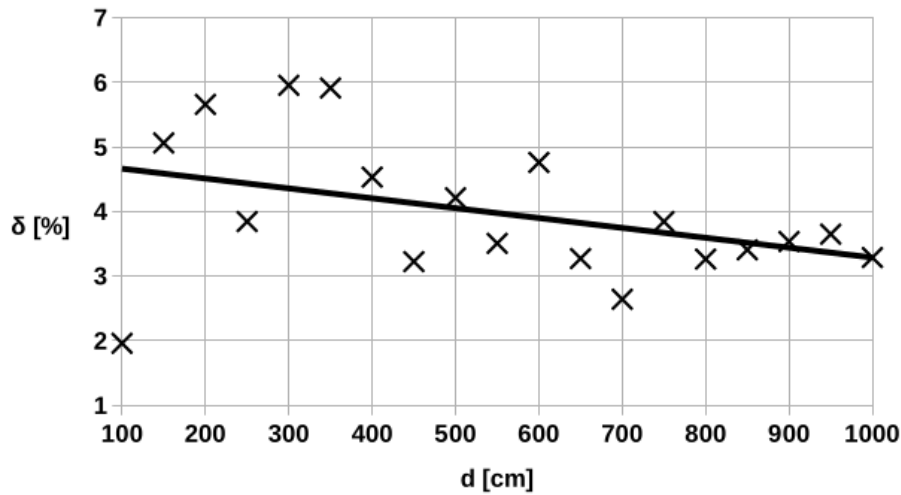
Figure 5 shows accuracy of UWB modules in Anechoic Lab and Classic lab. Accuracy in the graph is expressed as a relative error in the range of 30 to 220 cm. In the Anechoic laboratory, the accuracy agrees with datasheet information from the manufacturer, i. e. positioning accuracy is from 3 to 10 cm. In the Classic laboratory, the accuracy is worse, especially on short distances below 1m.



**Figure 5:** Measurement in the Anechoic Laboratory and the classic laboratory

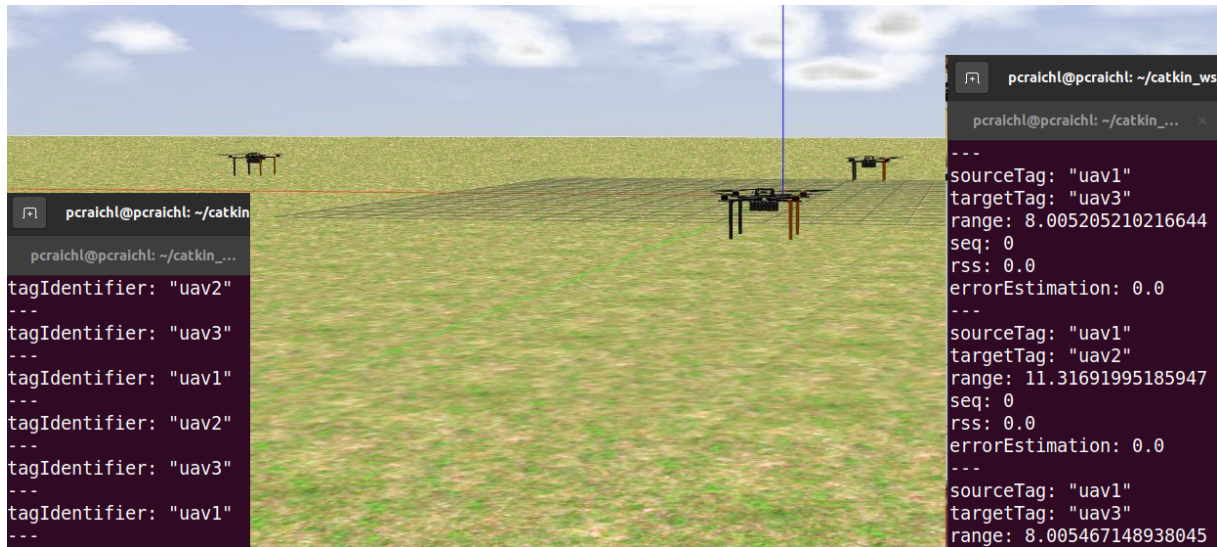
In figure 6 relative error of UWB modules from outdoor measurement in the range of 100 cm to 10 m is shown. The relative error is about 4% which is relatively low when compared to measurement in Classic lab. It is probably due to limited signal reflections in the environment without obstacles. This approves that UWB modules could be used for relative positioning of UAVs in swarms in an outdoor environment at least when flying above obstacles or after fusion with information from other sensors, especially cameras and lidars.

It is important to note that rotation of antenna has significant effects on the accuracy of UWB modules when even small obstacles are around it, which could make our approach for relative positioning in UAVs swarms problematic as UAVs can have different tilts and headings and other parts of UAVs can play a role of obstacles.



**Figure 6:** Measurement in the outdoor environment without obstacles

UWB plugin was integrated into ctu mrs architecture and simple simulation was done. In figure 7, there are 3 UAVs in simulation with integrated UWB plugin. Output from UWB topic was verified with ground true positions of robots. Verification was done for 3 different positions, see. tab. I. The positions match each other. The small difference between ground truth and UWB ranges is due to noise in UAV's movement and different timestamps of ground truth and UWB ranging.



**Figure 7:** 3 UAVs simulation - ctu mrs platform

**Table I:** Ground truth distances vs distances produced by UWB plugin for uav1 and uav2

Ground truth distance	UWB plugin distance	Difference
4.3056	4.4156	0.11
7.2564	7.1689	-0.0875
11.6259	11.4169	-0.209

#### 4. CONCLUSION

UWB modules DWM1000 from Decawave were tested in three different environments: Anechoic Laboratory, Classic Laboratory, and outdoor environment without obstacles. Accuracy of tested UWB modules represented by relative error is shown in fig. 5, 6. The results have shown that UWB modules could be used as a part of a relative positioning system in multi-robot systems, especially in the environment without obstacles or in heights above obstacles, where signal reflections are limited. ROS/Gazebo plugin of UWB modules was designed, integrated with ctu mrs platform, and tested in simple simulation with 3 UAVs, see fig. 7. In the future, the performance of UWB modules in an outdoor environment with obstacles could be tested.

#### ACKNOWLEDGMENT

This paper was funded from the general student development project at Brno University of Technology.

#### REFERENCES

- [1] M Saska, J Vakula and L Preucil. Swarms of Micro Aerial Vehicles Stabilized Under a Visual Relative Localization. In ICRA2014: Proceedings of 2014 IEEE International Conference on Robotics and Automation. 2014, 3570–3575.
- [2] J. Roberts and et al., "3-d relative positioning sensor for indoor flying robots." Autonomous Robots, vol. 33, no. 1-2, pp. 5-20, 2012
- [3] A. Ahmad et al., "Autonomous Aerial Swarming in GNSS-denied Environments with High Obstacle Density," 2021 IEEE International Conference on Robotics and Automation (ICRA), 2021, pp. 570-576, doi: 10.1109/ICRA48506.2021.9561284.
- [4] Liu, Liu, Bofeng Li, Ling Yang, and Tianxia Liu. 2020. "Real-Time Indoor Positioning Approach Using iBeacons and Smartphone Sensors" Applied Sciences 10, no. 6: 2003. doi:10.3390/app10062003
- [5] MRS UAV System - open source platform for UAV research. <https://ctu-mrs.github.io/>

# Optimization of electrostatic simulation by utilization of distribution equations with emphasis on design of capacitance sensors of inhomogeneity distance

V. Mancl<sup>1</sup>, P. Beneš<sup>1</sup>

<sup>1</sup>Brno University of Technology - FEEC, Czechia

E-mail: [xmancl00@vutbr.cz](mailto:xmancl00@vutbr.cz), [benes@vut.cz](mailto:benes@vut.cz),

**Abstract**—This paper deals with the problem of optimizing the simulation of electrostatic fields with a focus on electrode setup simulations for capacitive sensor design. Several simulations were performed in Ansys electronics desktop software in Maxwell 2D and 3D module. The simulations were carried out on a sensor model with parallel plates, between which the movable inhomogeneity was placed. In the simulation, the inhomogeneity was modeled as an solid object. The proposed optimization consisted in replacing the inhomogeneity as an object by the equation of permittivity distribution in the space between the plates of the capacitor. The obtained dependences of the capacitance on the position of the inhomogeneity of the two models were compared together. Also the time required to complete the simulation was discussed. In addition to evaluating the advantages and disadvantages of the two modelling approaches, the paper describes a method for simulating a continuous gradient distribution of inhomogeneity permittivity.

**Keywords**—Ansys Electronics Desktop, Dielectric distribution, Maxwell 2D, Maxwell 3D, Mesh operations, Parallel plate capacitor, Simulation

## 1. INTRODUCTION

Sensors based on capacitance measurement are an important part of industry and research. They are mostly used for indirect measurement of quantities. The quantities that can be measured this way are, for example, distance, humidity, pressure, flow or material thickness. [1]

The most important part of a capacitive sensor is its electrodes. The shape and material of the electrodes vary depending on the application and have a major influence on the key properties of the sensor. These properties must be known before manufacture. There are several methods to determine them. One of them is the analytical solution. Analytical solutions can only be used for simple cases where boundary conditions and the so-called "fringe effect" can be neglected. For more complex cases, there are often no exact analytical solutions or there are usually solutions that neglect many effects or are only empirical estimates.[2]

Another method is the production of a prototype and its subsequent measurement. This method is associated with higher costs and has very limited possibilities in terms of optimising the desired sensor properties. The last and most commonly used method is simulation based on finite element method. Simulations provide almost unlimited design and optimization possibilities. The biggest disadvantage is the computational complexity, which is also the most limiting factor.

This paper deals with simulations in Ansys Electronics Desktop using Maxwell 2D and 3D for electrostatic modelling. The simulations were created with the research objective in mind, which is a sensor that measures the distance of inhomogeneity from electrodes in a homogeneous dielectric. Simple models of the electrodes in a parallel capacitor setup were created. The models were used to investigate the possibility of speeding up more advanced simulations and thus reducing their computational complexity for a conventional desktop computer. To achieve the speed-up, a substitution of the measured object by the equation of permittivity distribution in the electrode space was proposed.

In addition to optimization, the possibility of simulating a continuous distribution of permittivity in the measured dielectric is also investigated.

## 2. BASE MODEL

The models for the simulations used in this work were based on the well-known model of a capacitor with parallel plate electrodes, which is described by the equation 1

$$C = \epsilon_0 \epsilon_r \frac{s}{d} \quad [F] \quad (1)$$

where  $\epsilon_0$  is the vacuum permittivity,  $\epsilon_r$  is the relative permittivity,  $s$  is the electrode area and  $d$  is the distance between the electrodes.

This equation is valid only if the boundary conditions are neglected or if the area of the electrodes is significantly larger than the distance between them. Without neglecting the boundary conditions, the so-called "fringe effect" occurs, which causes an inhomogeneous distribution of the electric field around the electrodes. The fringe effect with combination with a dielectric that does not fill the entire space between the electrodes, causes the capacitance of the capacitor to be dependent on the position of the dielectric. This dependence does not exist in the ideal case (see 1, analytical solution of the ideal model).

### 2.1. Using base model in simulations

According to the parallel capacitor model, two types of models were created, for which their individual variations were subsequently created. Both types contained two electrodes with dimensions of 100 x 100 x 1 mm and a gap of 80 mm. Intentionally, both types were created to be equivalent and to differ only in the form of inhomogeneity. The reason for this was to test methods of inhomogeneity replacement for the second type of simulation.

The first type of models was labeled "A" and contained an inhomogeneity between the electrodes of the capacitor in the form of a rectangular object that changed its position according to the parametric analysis. This object had predetermined dimensions of 100 x 100 x 5 mm and a relative permittivity of  $\epsilon_r = 5$ .

The second type, labeled "B", contained inhomogeneity in the form of a fill between the electrodes, which permittivity was defined as a global variable for which there is a possibility to write its value as a distribution function, which is defined as a function of the dependence of the permittivity on one of the three axes of the simulation. The Maxwell 2D and 3D environment provides only the basic mathematical functions. To create an equivalent model for both versions, a distribution function was derived using the signum (sgn) function. The following function was generated:

$$\epsilon(X) = \epsilon_{sgn1}(X) \cdot \epsilon_{sgn2}(X) \cdot \epsilon_{rh} + \epsilon_{ra} \quad [-] \quad (2)$$

$$\epsilon_{sgn1}(X) = (\text{sgn}(1000X + 0.5W_n - l) + 1) \cdot \text{sgn} \frac{1000X + 0.5W_n - l}{2} \quad [-] \quad (3)$$

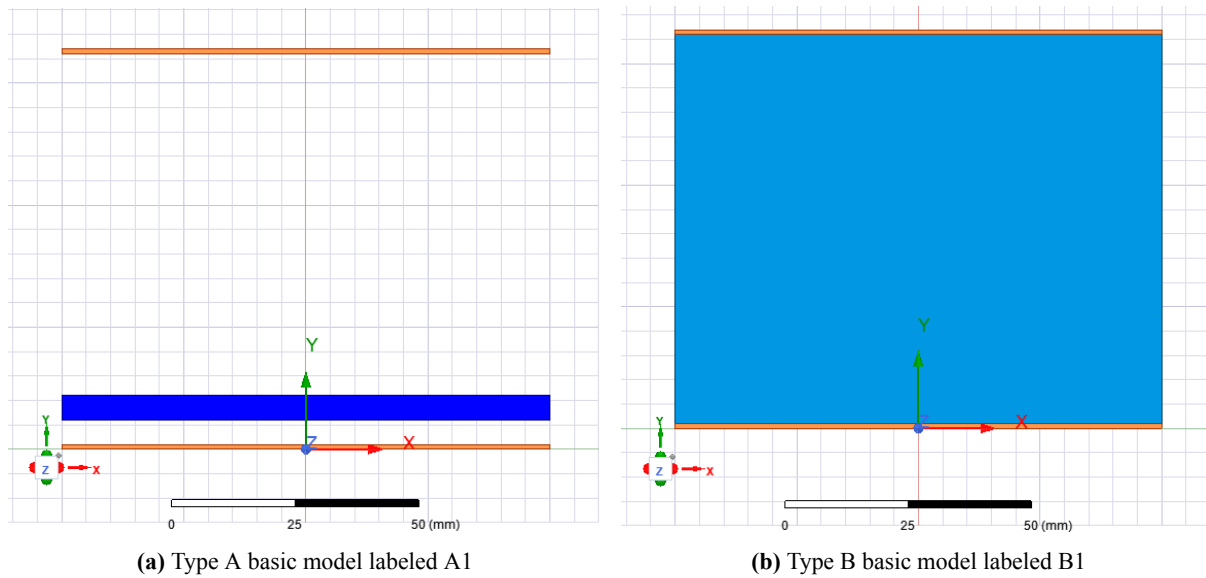
$$\epsilon_{sgn2}(X) = (\text{sgn}(-1000X + 0.5W_n + l) + 1) \cdot \text{sgn} \frac{-1000X + 0.5W_n + l}{2} \quad [-] \quad (4)$$

where  $\epsilon_{sgn1}$  and  $\epsilon_{sgn2}$  are functions that result in step function,  $\epsilon(X)$  is the final distribution equation,  $X$  are values of one of the main axis of simulation,  $W_n$  is the inhomogeneity width,  $l$  is position/distance of inhomogeneity,  $\epsilon_{rh}$  is relative permittivity without the relative permittivity of the area  $\epsilon_{ra}$ . If the area permittivity was set to  $\epsilon_{ra} = 2$  and the desired inhomogeneity permittivity was  $\epsilon_r = 5$ , then

$$\epsilon_{rh} = \epsilon_r - \epsilon_{ra} = 5 - 2 = 3 \quad [-] \quad (5)$$

Both types of models can be seen in the figure 1a and 1b. The main idea of speeding up/optimizing the simulation is based on the fact that if the position of the object changes in the type A simulation, it is necessary to re-generate the mesh. This is because the geometry of the model has changed. In type B, however, there is no movement of the object but only a change in the permittivity distribution. This avoids the need to re-generate the mesh. However, another problem arises here, where in the latter case it is not possible to use adaptive meshing and it is necessary to have a suitably generated universal

mesh beforehand. A universal mesh means a higher resolution and a higher complexity for the final computation, but decreases computing requirements for its generation.



**Figure 1:** Basic models of the parallel plate capacitor

### 3. VARIATION OF SIMULATIONS

Three variations were created for the basic model of the first type A and six variations for the second type B. An overview of the settings and differences of each variation is given in the table below see tab. I.

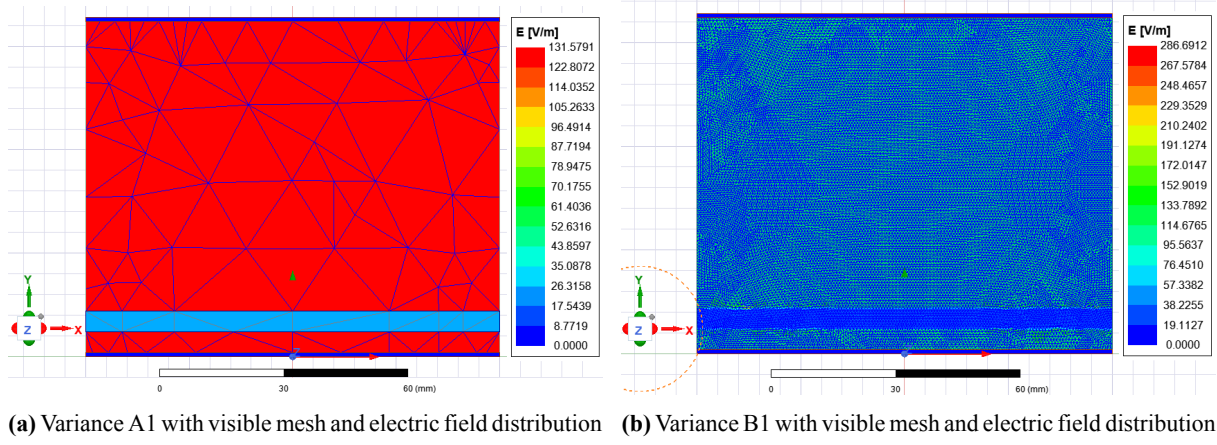
Description	Label	Accuracy	Refinement	Tetrahedron count	Time	One iteration
2D w/o fringe field	A1	0.1%	30%	210	0:00:32 s	0:00:02
2D w fringe field	A2	0.1%	30%	430	0:00:34	0:00:02
3D w fringe field	A3	0.1%	20%	281 000	1:36:7	0:05:59
2D w/o fringe field	B1	0.1%	30%	40 000	0:07:18	0:00:23
2D w fringe field	B2	0.1%	20%	163 000	0:24:55	0:01:29
3D w fringe field	B3	0.3%	30%	1 739 000	8:52:03	0:25:51
3D w fringe field	B4	0.2%	30%	200 000	0:11:02	0:00:35
3D w fringe field	B5	0.2%	30%	520 000	0:47:12 s	0:02:30
3D w fringe field	B6	0.2%	30%	734 000	1:05:14 s	0:03:52

**Table I:** Summary of properties of each simulation variation.

#### 3.1. Variation A1 and B1

These variations were created in the Maxwell 2D environment and are directly based on the base model and are equivalent to each other. There were no defined surroundings and thus no edge effects. The simulation was mainly used to check whether the equation 5 was correct and whether Ansys was working with the equations as expected. As a result, the dependence of the capacity on the position of the inhomogeneity was determined. As expected, there is no change and the capacity is constant. In figure 2a and 2b the distribution of the electrostatic field and mesh can be seen. The progression of capacitance over the inhomogeneity distance is shown in figure 3a. Due to the adaptive meshing the type B simulation has to "guess" where the inhomogeneity is by performing adaptive meshing. That results into the mesh with higher node count.





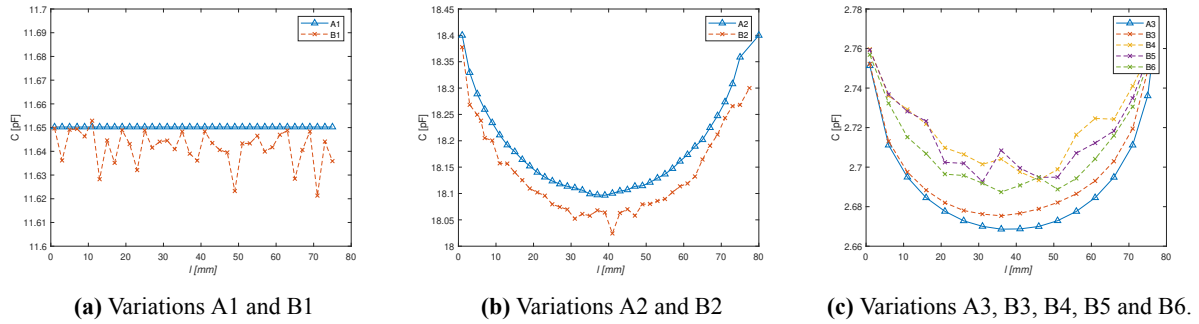
**Figure 2:** Variance A1 and B1

### 3.2. Variation A2 and B2

This is a variation of A1 and B1 where only the vacuum surroundings have been added. In this case, the capacitance depends on the position of the inhomogeneity. The resulting dependence of the capacitance on the inhomogeneity distance is shown in figure 3b

### 3.3. Variation A3, B3, B4, B5 and B6

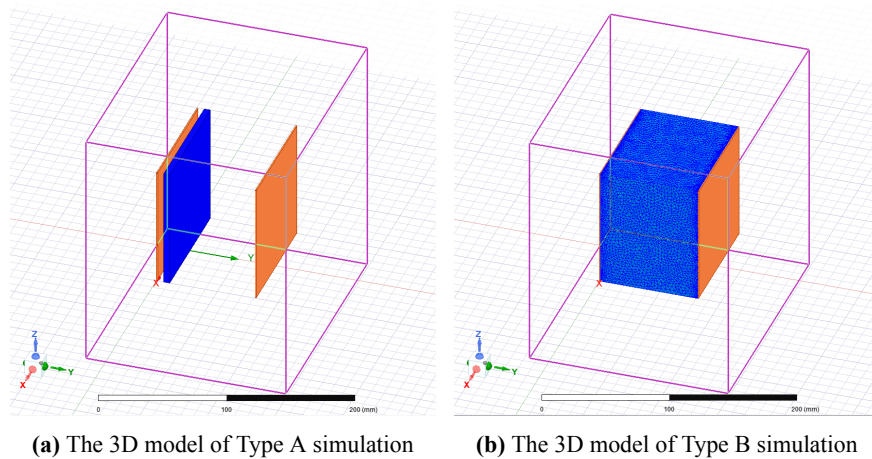
These variations were created by adding a third dimension to the 2D model. All variations contained a vacuum environment. Another modification was to round the edges to ensure proper convergence. Which also complicates mesh generation. Variation B3 was solved using the "copy equivalent mesh" and "continue adaptive passes" modes. [3] B4 to B5 used the solving method only with the first generated mesh. They differ from each other in the resolution of the mesh functions used. That is, the first generated mesh for B4 had a mesh resolution almost four times smaller than B6. The 3D models can be seen in figure 4a and 4b. The capacitance dependencies can be seen in figure 3c.



**Figure 3:** The capacitance dependency on inhomogeneity distance/position

### 3.4. Implications

The results show that the two types of models are equivalent. It can also be observed that in simple cases type B is slower and more computationally intensive. However, for more complicated models i.e. 3D and with rounded edges, it can be seen that with the right choice of mesh generation, sufficient results can be achieved. In doing so the mesh only needs to be generated once. The difference can be seen between A3 and B6 simulations. In the case of B6, the need to generate a new mesh for each iteration of the parametric analysis was eliminated, resulting in a speed-up of two minutes. It should be noted that the speeding up of the simulation was at the cost of a higher memory requirements, due to a mesh with having considerably more nodes, and at the same time a worse accuracy. However, it can be inferred that Type B simulations would be particularly appropriate when long-term simulations are to be performed on the model for different positions of different inhomogeneities. Furthermore, it can be inferred from the results that a high accuracy could also be achieved by proper initial settings for mesh generation.



**Figure 4:** Type A and type B models.

An important implication of the results is the ability to use the equations to define the distribution of permittivity in an object. In this way, a continuous distribution such as the Gaussian distribution can be defined. This distribution cannot be simulated by type A simulations.

#### 4. CONCLUSION

Simulations were performed on a simple model of a capacitive sensor with parallel plate electrodes. Two types of simulations were created. Type A with inhomogeneity defined as an object. Type B, where the inhomogeneity is defined by the equation of the dielectric distribution in the gap between the electrodes. Type A is suitable for simple and accurate models where it is easy to always re-generate the mesh. Type B is intended for situations where it is necessary to observe the effect of inhomogeneity that does not have precise boundaries and has a continuous distribution in an area. At the same time, type B is suitable for simulations on which analysis will be carried out over a long period of time. The disadvantage of type B is its memory requirements for simple models and also its greater complexity in mesh generation. For more complex models, this disadvantage is not as substantial. It must be added that further research will be needed to reach a clear conclusion in the matter of the comparison of type A and type B simulations.

The most important conclusion is that the distribution of permittivity in a dielectric can be defined by a function. This has been verified using a B-type simulation with the distribution function 5, which results were in accordance with the equivalent A-type simulation model. This approach will be used in future simulations of the dependence of the capacitance of capacitive sensors on the concentration of materials in space.

#### ACKNOWLEDGMENT

This work was supported by the project FEKT-S-20-6205 Industry 4.0 in automation and cybernetics, Internal Grant Agency of Brno University of Technology.

#### REFERENCES

- [1] L. K. Baxter. "Capacitive Sensors. Design and Applications". New York: IEEE Press, 1997. ISBN 0-7803-1130-2.
- [2] R. P. Feynman, R. B. Leighton, M. Sands. "The Feynman lectures on physics. Volume II, Mainly electromagnetism and matter". New millennium edition. New York: Basic Books, 2010.
- [3] Ansys, Inc. "ANSYS Electromagnetics Suite 2021 R1 HELP" 2022. Pittsburgh, Pensylvanie, USA.

# Multiscale Modelling of Li-Ion Battery Volume Expansion

M. Mačák<sup>1</sup>, P. Vyroubal<sup>1</sup>

<sup>1</sup>Department of Electrical and Electronic Technology, Brno University of Technology, Brno 616 00, Czech Republic

E-mail: [martin.macak@vut.cz](mailto:martin.macak@vut.cz), [vyroubal@vut.cz](mailto:vyroubal@vut.cz)

**Abstract**—Mechanical characteristics of Li-ion batteries play an important part in the design of electric devices or energy storage systems. Proper mechanical design can increase the safety and the longevity of devices. Apart from external forces, Li-ion batteries also experience internal volume changes due to inherent physical processes, especially volume changes caused by intercalation and deintercalation of lithium ions into electrode materials. Currently, these changes are mainly studied experimentally using complex methods, while numerical simulations of these processes are surprisingly rare. This work presents different numerical approaches for studying internal volume changes in Li-ion batteries. The results show that it is possible to study volume changes at different length scales, but it is necessary to apply a suitable simplification of the model to obtain macroscopic deformation.

**Keywords**—Li-ion battery, numerical simulation, volume expansion, intercalation, electrochemistry

## 1. INTRODUCTION

In recent years, mechanical behavior of Li-Ion batteries became an interesting area due to their importance in the design and longevity of energy storage systems. The structural integrity and damage resistance are important factors in the transportation industry and especially for emerging electrical vehicles. One specific property of Li-ion battery is its sensitivity to external and internal mechanical loads. These loads can lead to internal short circuits, which could result in thermal runaway and rapid dissipation of energy [1, 2].

Furthermore, volume changes are directly related to inherent electrochemical processes. These changes can be caused by lithium migration, in which electrode materials change their volume due to lithium intercalation and deintercalation. This process seems to be the most significant one as graphite can expand up to 13%, while common cathode materials up to 5 % [3]. Additionally, effects such as gas formation due to side reactions or lithium plating can lead to further volume changes. Both reversible and irreversible volume changes increase with the degradation of the battery cell. As a result, accurate prediction of volume changes is necessary to ensure safety and longevity of batteries designed for long term use [4]. Experimental investigation of these changes is often difficult, as the maximum deformation is usually in tenths of millimeters. To estimate the internal deformation of the layers, it is necessary to use techniques such as micro-CT or destructive methods such as disassembling the battery [1]. The disadvantage of experimental methods is that they require a lot of time and usually they can be carried out only on a smaller scale.

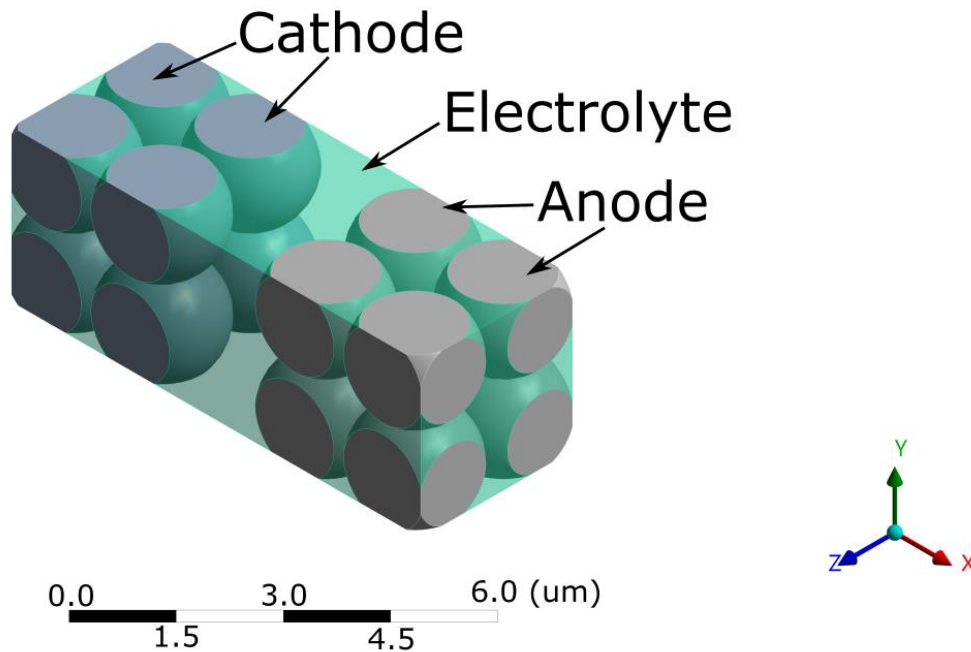
On the other hand, numerical simulations can provide a suitable alternative to the experiments. Currently, most research on internal volume changes is carried out on the microscopic scale [5-7]. While this approach is the most accurate, as it directly describes the electrochemical nature of the process, its complexity makes it nearly impossible to apply it on the macroscopic scale. Currently, most of the macroscopic models focus on the effects of external forces, while the internal volume changes remain suppressed [8].

The presented work shows a possibility of numerical modelling of internal volume changes in Li-ion batteries caused by elastic volume change caused by intercalation and deintercalation at different length (particles, layers, and whole battery) scales. The simulations show that a simplification of the model based on the geometry is necessary to obtain precise results in appropriate time.

## 2. NUMERICAL MODEL

As it is possible to study volume changes of Li-ion batteries from different perspective, it is necessary to use appropriate models for each case based on the geometry scale. The simulations presented in this work focused on a description of volume changes at different scales (particles, electrode layers and the whole battery). Electrochemical and mechanical properties of an 18650 NCA battery were used for simulations [2, 3, 9]. The volume changes were caused only by lithiation and delithiation as it seemed to be the most important mechanism. For all simulations, it was assumed that the anode volume was increased by 10%, while the cathode volume was decreased by 5% [1]. This process represented a discharge of the battery. The structural simulations were carried out in Ansys Mechanical software, while the electrochemical simulation was carried out in Ansys Fluent.

Microscopic simulation used a numerical model based on the models presented in [5] and [6], which describe in detail the transport and intercalation of lithium ions into the cathode/anode host materials. The transport of lithium inside the particles is only governed by diffusion, while the transport of lithium ions in the electrolyte is derived from a concentrated solution theory, which considers both solute-solvent and solute-solute interaction. The charge transfer chemistry at the solid-electrolyte interface is described through a Butler-Volmer equation. In this model, the volume expansion can be directly related to lithium concentration in a particle, as the initial and maximum concentrations are known parameters [6]. The simplified geometry is shown in Fig. 1. The radius of anode and cathode particles was set to 1  $\mu\text{m}$ .

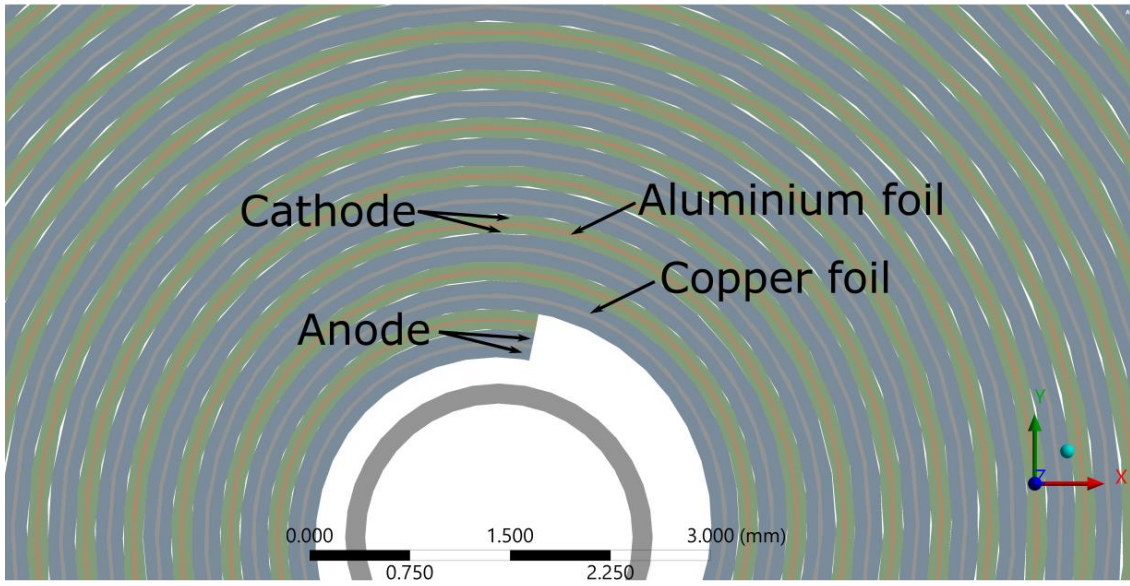


**Figure 1:** Simplified microstructure of a Li-Ion battery. Each particle has a radius of 1  $\mu\text{m}$

Simulations which focused on large scale volume changes were described by a simplified approach. Even a commonly used physics-based Newman's P2D model [10] does not readily holds information about lithium concentration in the electrodes. In these models the anode, the cathode and the electrolyte are usually considered as single joint zone, so it would not be possible to define individual volume changes for different materials. For this reason, it was not necessary to resolve whole electrochemical model. Instead, only estimated volume changes were set for the anode (+10%) and the cathode (-5%) materials. The simulation was carried out in Ansys Mechanical software, as the problem became purely mechanical. The thickness of layers was defined as: anode – 80  $\mu\text{m}$ , aluminium foil – 30  $\mu\text{m}$ , cathode – 55  $\mu\text{m}$  and copper foil – 30  $\mu\text{m}$ . The thickness of steel casing was set to 250  $\mu\text{m}$ . The height of the jelly roll was set to 60 mm. Mechanical properties of materials were obtained from [2]. The geometry for the whole battery simulation consisted of a steel casing with a width of 18 mm and a height of 65 mm. Inside the casing, there was a 60 mm tall domain, which represented the jelly roll structure. A similar approach to the method shown in [11] was used to obtain average material properties of the simplified



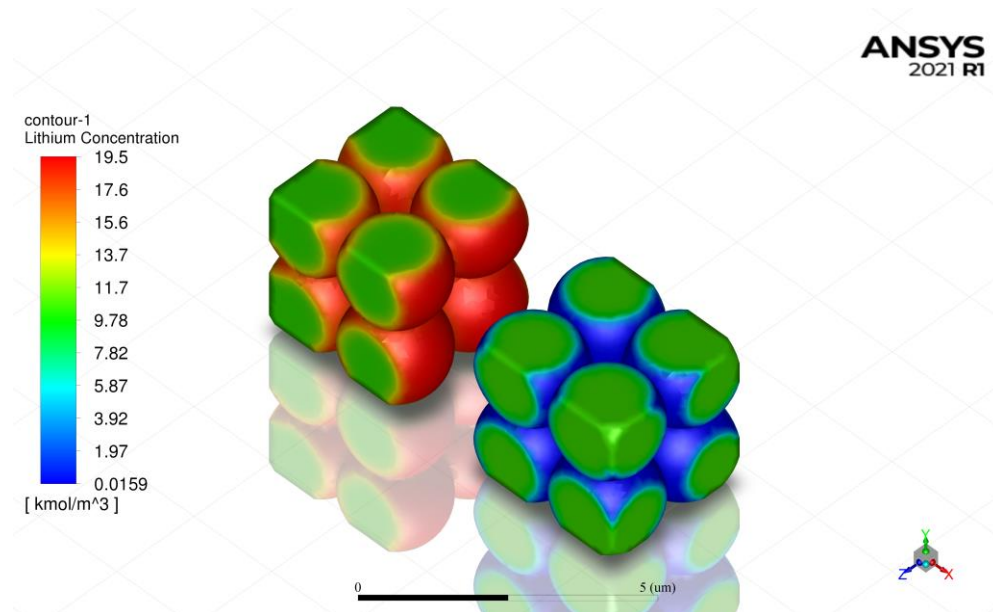
continuum domain.



**Figure 2:** Detail of the jelly roll layer structure with visible individual layers.

### 3. RESULTS

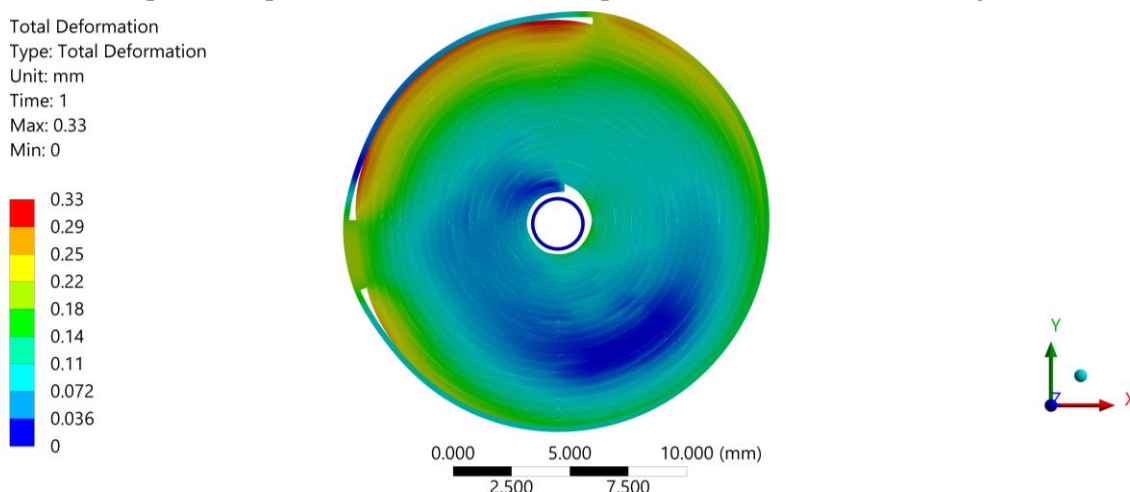
Results from microscopic simulation show that it is possible to track lithium concentration inside the electrode particles. As lithium ions migrate from the anode, they are intercalated into the cathode structure, which is shown by the increased lithium concentration near the particle surface. Lithium ions inside the particles are then transported by a diffusion mechanism. In this case the volume changes are not very significant. This is caused by the fact that the actual active layer of the electrode consists of many particles, and the overall volume change is the sum of individual contributions. In order to increase the particle volume by 10%, its radius has to be only increased approximately 1.032 times. This means that the original radius of 1  $\mu\text{m}$  would be 1.032  $\mu\text{m}$  at the end of the simulation, so the change is not very noticeable.



**Figure 3:** Lithium concentration inside microscopic structure of Li-Ion battery.

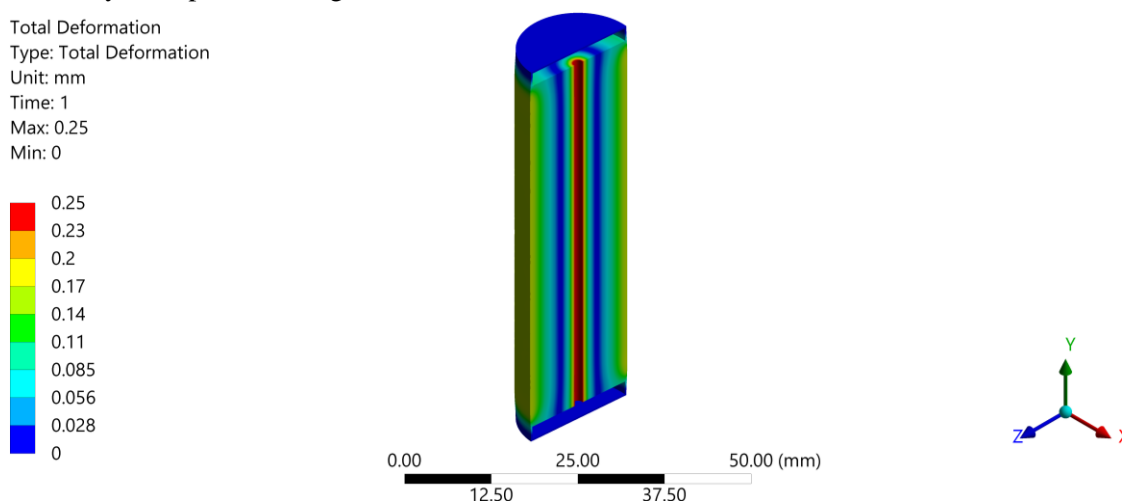
Fig. 4 shows the macroscopic deformation of the jelly roll structure. In this case, the volume change was set to individual cathode and anode layers. The maximum deformation was visible at the end of the cathode layer (top of the geometry) and the top left area, which was empty. The average deformation of

the steel casing was approximately 0.18 mm. Generally, the inner layers experienced more strain as they had more free space to expand. This effect was also reported in the work of Willenberg et al. [1].



**Figure 4:** Deformation of a jelly roll structure with resolved individual layers.

Even the resolution of individual layers may be too detailed for the study of multiple batteries in battery packs. Fig. 5 shows the result of volume changes in the whole battery. Even though, the jelly roll structure was considered as a continuum with averaged values, the results generally agree with the previous simulation with resolved layers. The average outward deformation of the steel casing was approximately 0.17 mm. The main difference, in contrast to the previous simulation, is that in this case, the inner layers experienced significant deformation.



**Figure 5:** Deformed geometry of an 18650 NCA battery.

#### 4. CONCLUSION

Presented simulations show that it is possible to study volume changes of Li-ion batteries at different length scales. However, it is necessary to choose an appropriate model for the desired application. While the simulation at the microscopic scale should be the most precise, the volume changes can be studied only at this scale. To study macroscopic effects by this method, it would require resolving complex and large geometries, which would be extremely computationally expensive. On the other hand, simplified methods show good accuracy while being relatively quick. The resolution of individual layers is an effective method to study processes inside individual batteries. To study volume changes of batteries in battery packs, defining of individual layers may become too complicated, so a connection of layers into a single zone may seem to be the best option. Generally, these simulations can help with identifying critical areas in the design of batteries or even battery packs, which might lead to a mechanical failure. These models can be also extended to include the effects of thermal expansion, Joule heating, inelastic volume changes or gasification, which would increase their accuracy.



## ACKNOWLEDGMENT

This work was supported by the BUT specific research program (project No. FEKT-S-20-6206).

## REFERENCES

- [1] L. K. Willenberg, P. Dechent, G. Fuchs, D. U. Sauer, and E. Figgemeier, “High-Precision Monitoring of Volume Change of Commercial Lithium-Ion Batteries by Using Strain Gauges”, *Sustainability*, vol. 12, no. 2, 2020, doi 10.3390/su12020557.
- [2] C. Zhang, J. Xu, L. Cao, Z. Wu, and S. Santhanagopalan, “Constitutive behavior and progressive mechanical failure of electrodes in lithium-ion batteries”, *Journal of Power Sources*, vol. 357, pp. 126-137, 2017, doi 10.1016/j.jpowsour.2017.04.103.
- [3] S. Schweidler, L. de Biasi, A. Schiele, P. Hartmann, T. Brezesinski, and J. Janek, “Volume Changes of Graphite Anodes Revisited: A Combined Operando X-ray Diffraction and In Situ Pressure Analysis Study”, *The Journal of Physical Chemistry C*, vol. 122, no. 16, pp. 8829-8835, Apr. 2018, doi 10.1021/acs.jpcc.8b01873.
- [4] C. R. Birkel, M. R. Roberts, E. McTurk, P. G. Bruce, and D. A. Howey, “Degradation diagnostics for lithium ion cells”, *Journal of Power Sources*, vol. 341, pp. 373-386, 2017, doi 10.1016/j.jpowsour.2016.12.011.
- [5] A. M. Colclasure and R. J. Kee, “Thermodynamically consistent modeling of elementary electrochemistry in lithium-ion batteries”, *Electrochimica Acta*, vol. 55, no. 28, pp. 8960-8973, 2010, doi 10.1016/j.electacta.2010.08.018.
- [6] G. M. Goldin, A. M. Colclasure, A. H. Wiedemann, and R. J. Kee, “Three-dimensional particle-resolved models of Li-ion batteries to assist the evaluation of empirical parameters in one-dimensional models”, *Electrochimica Acta*, vol. 64, pp. 118-129, 2012, doi 10.1016/j.electacta.2011.12.119.
- [7] R. Xu, Y. Yang, F. Yin, P. Liu, P. Cloetens, Y. Liu, F. Lin, and K. Zhao, “Heterogeneous damage in Li-ion batteries: Experimental analysis and theoretical modeling”, *Journal of the Mechanics and Physics of Solids*, vol. 129, pp. 160-183, 2019, doi 10.1016/j.jmps.2019.05.003.
- [8] L. Tang, J. Zhang, P. Cheng, and X. Hu, “Homogenized modeling methodology for 18650 lithium-ion battery module under large deformation”, *PLOS ONE*, vol. 12, no. 7, Jul. 2017, doi 10.1371/journal.pone.0181882.
- [9] C. M. Julien and A. Mauger, “NCA, NCM811, and the Route to Ni-Richer Lithium-Ion Batteries”, *Energies*, vol. 13, no. 23, 2020, doi 10.3390/en13236363.
- [10] K. Uddin, S. Perera, W. Widanage, L. Somerville, and J. Marco, “Characterising Lithium-Ion Battery Degradation through the Identification and Tracking of Electrochemical Battery Model Parameters”, *Batteries*, vol. 2, no. 2, 2016, doi 10.3390/batteries202001.3
- [11] P. Vyroubal and T. Kazda, “Equivalent circuit model parameters extraction for lithium ion batteries using electrochemical impedance spectroscopy”, *Journal of Energy Storage*, vol. 15, pp. 23-31, 2018, doi 10.1016/j.est.2017.10.019.

# Complex systems resilience qualification

P. Dvorský<sup>1</sup>, O. Baštán<sup>2</sup>, and P. Fiedler<sup>3</sup>

<sup>1,2,3</sup>Brno University of Technology  
Faculty of Electrical Engineering and Communication  
Department of Control and Instrumentation  
Czech Republic

E-mail: [191395@vut.cz](mailto:191395@vut.cz), [156224@vut.cz](mailto:156224@vut.cz), [fiedlerp@vut.cz](mailto:fiedlerp@vut.cz)

**Abstract**—This paper focuses on resilience qualification and propose a new method of measuring resilience within the systems. Within the paper is discussed the concept of resilience. Further, there is a defined general complex system with its main system functions. Based on that, the paper proposes a methodology and a model by which could be resilience of systems evaluated. An example of an application of a proposed resilient metric to a specific system is also discussed here.

**Keywords**— Resilience metrics, Main system functions, Criterion space, Criterion oninon model

## 1. INTRODUCTION

Within every industrial system, there is a need for knowledge of the actual state and overall healthiness of the system. However, for the complex systems, it is very difficult to enumerate and analyze their potential hazards and failures and also analyze them within different areas like safety, security, etc.

Therefore, there is increasing demand for an overall solution that could cover up all these areas and make analyzing complex systems more effortless. For complex systems, there needs to be introduced other metric that can incorporate all these attributes within itself and give some evaluation. One such approach to this problem of evaluating the properties of complex systems and their states could be resilience.

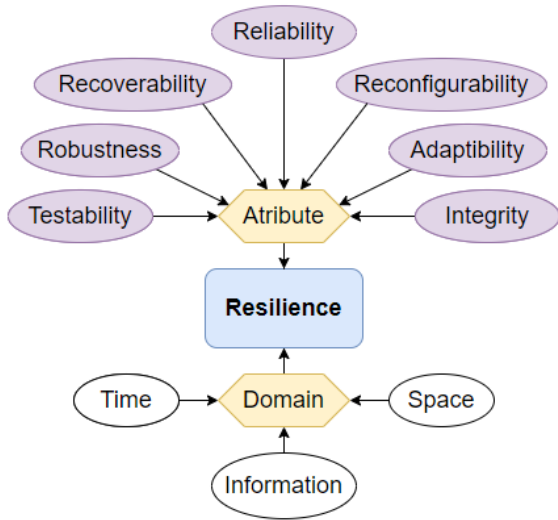
Several definitions of resilience have been offered in past years. Many are similar, though many overlaps with a number of already existing concepts such as robustness, fault-tolerance, flexibility, survivability, and agility, among others.[1] Within this paper, we will look at resilience same as in [2], thus on the property of a system or device that can be divided into different domains and attributes - figure 1. It is necessary to add that resilience is an attribute of a system, not an add-on to a system. Therefore, a design principles which enhance resilience may be design principles which achieve other desirable characteristics of systems.[3]

The definition of resilience by Prengerzer [4] is as follows “measure of a system’s ability to absorb continuous and unpredictable change and still maintain its vital functions.” Furthermore, Allenby and Fink [5] defined resilience as the “capability of a system to maintain its functions and structure in the face of internal and external change and to degrade gracefully when it must.” Based on [4], [5] and [2], we can say that a resilient system is a system that is flexible enough, further resilient system is able to respond to unpredictable adverse and harmful situations and conditions, adapt to them and return to its original state of operation after their disappearance. It is a fundamental feature for correct operation of a system in various conditions, but due to inconsistency in terminology and generality of usage, it is difficult to quantify and qualify the resilience of systems.

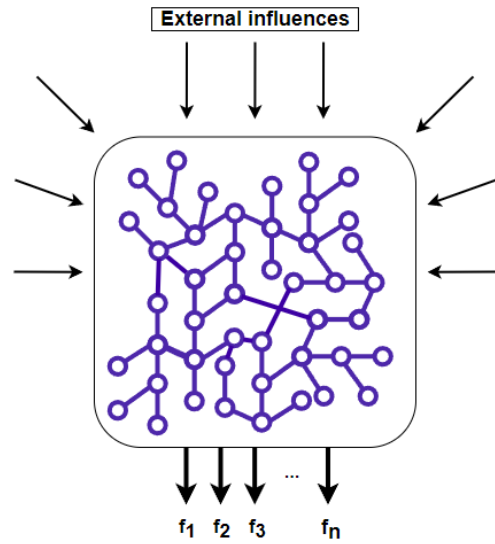
Despite that, there are also lots of approaches how to qualify resilience described in [1]. In this paper, we propose our approach to how could be the resilience of such systems qualified and measured, so it could be seen, if any pattern and approaches for an increase of resilience really work.

## 2. GENERAL COMPLEX SYSTEM DEFINITION

Let’s consider a generally complex system whose internal structure is unknown to us. The system can be generally linear or nonlinear and its internal structure can be arbitrarily complex, the system can consist of any number of subsystems of any structure, hierarchy and their interconnections (see figure 2). The



**Figure 1:** Elements of resilience



**Figure 2:** General complex system

system is primarily designed to perform  $f_1$  to  $f_n$  output functions Based on these output functions and the qualities they provide, we could be able to evaluate whether the system is functional or not, eventually, if the output functions are provided with sufficient or insufficient quality.

### 3. MAIN SYSTEM FUNCTION DEFINITION AND LIMITS OF MAIN FUNCTION DEFINITION

The main function of a system is any function of the system that is performed by the system and for which the system was primarily designed. There can be 1 to  $n$  system functions that are performed by the system (see figure 2). A number of functions and their form must be always enumerated and specified for a particular system.

Let us also consider that within the evaluation of these main functions we will be based on three domains described in [2], namely in space, time and information domains ( $S$  = space/structure,  $T$  = time,  $I$  = Information).

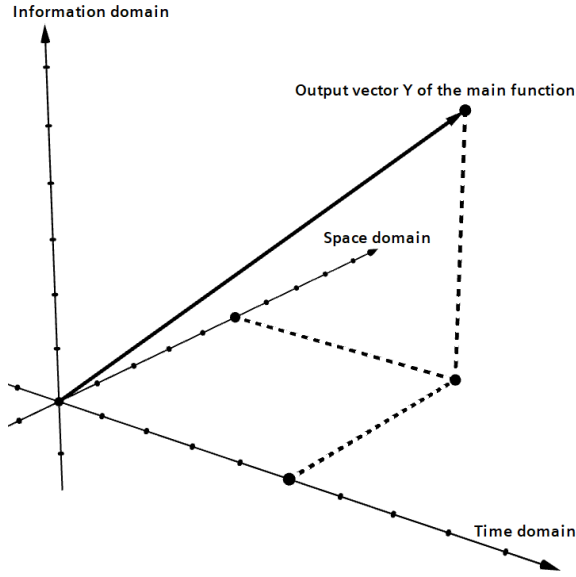
For simplicity, let's assume that the system has only one main function  $f$  to perform and such a function has only one output  $Y$ . If we consider that the output of the function  $Y$  can be divided into elementary components within individual domains, we are able to express the output  $Y$  as a vector in the coordinates  $S, T, I$  as follows  $Y = Y_s, Y_t, Y_i$  (see figure 3).

Furthermore, the main function  $f$  depends on the inputs  $U$  ( $U_s, U_t, U_i$ ), the unknown influencing disturbances  $E$  ( $E_s, E_t, E_i$ ) and the states of the system  $X$  ( $X_s, X_t, X_i$ ) (see figure 4). In general, the main function  $Y = f(U, X, E)$ , can be decomposed into individual components of particular domains as follows:

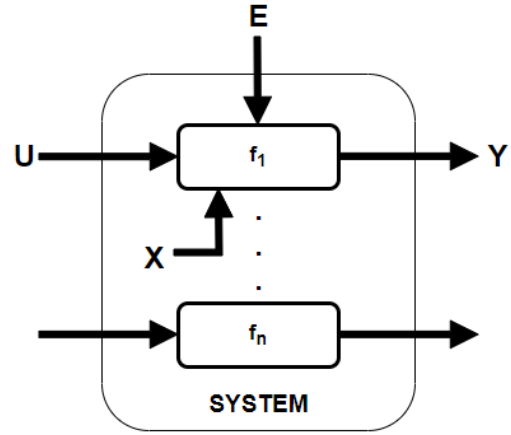
$$(Y_s, Y_t, Y_i) = f(U_s, U_t, U_i, X_s, X_t, X_i, E_s, E_t, E_i) \quad (1)$$

### 4. DEFINITION OF CRITERION SPACE AND EVALUATION OF OUTPUT FUNCTION QUALITY

If we accept the fact that the output  $Y$  is decomposable into the three components mentioned above and it is therefore generally a three-dimensional vector. We can further assume that there will exist some space in this respect where the output vector  $Y$  will be located during the correct performance of the main functions of the system.



**Figure 3:** Main function output vector  $Y$



**Figure 4:** Main functions model

Therefore, if we find the limits in the space where the outputs of the main function are acceptable to us, i.e., the space where the main function acquires acceptable values for its purpose and system correctly fulfils the main functions, we get so-called criterion space. If the output  $Y$  is located outside this criterion space the system does not properly perform its main function and if the output of the system is inside the criterion space, the system correctly fulfils its main function. The above mentioned is graphically expressed in figure 5.

If we will continue in previous considerations about the system where  $U$  are ideal inputs to a function from the system,  $X$  are ideal states of this system and  $E$  represents all errors, faults and disturbances that affect the system itself, we can evaluate the degree of resilience of the system according to how large the disturbance  $E$  has to be in order to be the output vector  $Y$  taken out of the defined criterion space. In other words, we try to find out how large and in which domains disturbance  $E$  will lead the output vector  $Y$  out of the defined criterion space defined by us when we have ideal inputs  $U$  and states  $X$ .

This can be also formulated as follows: We are trying to find how large disturbance  $E$  and in which domains are still tolerable by the system so that the output of the main function is still within our criterion space.

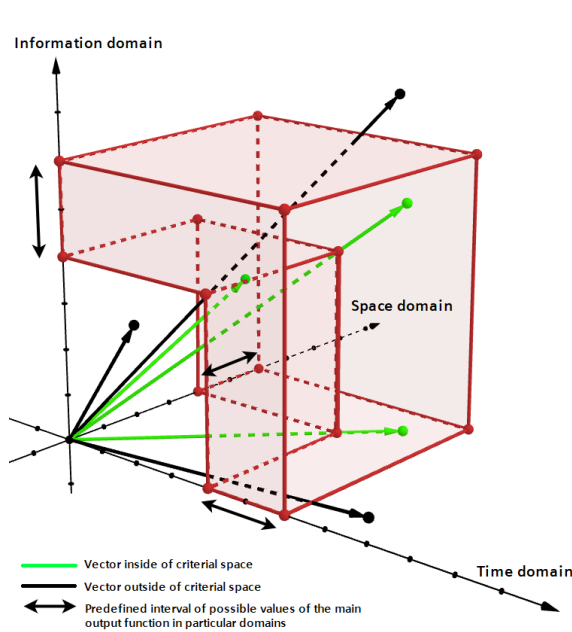
The shape of the criterion space can be defined either by intervals of values in individual domains, which the output  $Y$  of the main function can take for proper operation (see figure 5). In the case of more complex systems, we can define the required shape of the criterion space according to the needs of the application instead of the above-mentioned basic intervals.

The criterion space in the figure 5 is defined as follows:

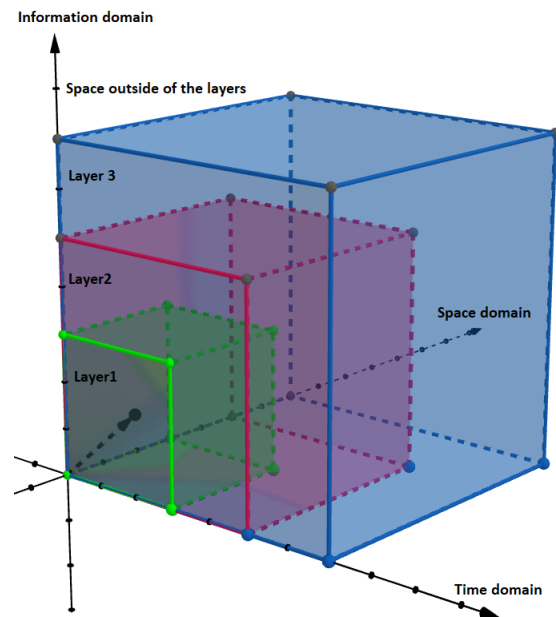
$$Y_{min} < Y < Y_{max} \iff \begin{matrix} Y_{T_{max}} < Y_T < Y_{T_{max}} \\ Y_{S_{max}} < Y_S < Y_{S_{max}} \\ Y_{I_{max}} < Y_I < Y_{I_{max}} \end{matrix} \quad (2)$$

In the above-mentioned treatise is captured only the fact that the output  $Y$  of the main function of the system is or is not located within the criterion space, i.e., the main function performs correctly or incorrectly.

This approach, therefore, looks at the problem only from a "black and white perspective" - System works or System doesn't work. However, this approach does not fit very often with real-world scenarios where the reality is not binary, but it is far more diverse and has several shades. It means that the system still can perform the main function somehow correctly while disturbances and adverse conditions take place, but with lower quality.



**Figure 5:** Criterion space with output vector  $Y$



**Figure 6:** Criterion onion model

The criterion space could therefore be further stratified into several layers - the onion layer model (see figure 6), which corresponds to the output quality of the main function  $Y$ . An example is in the following figure 6. The origin of the model represents the best possible operation mode of the system. If the vector  $Y$  is located within the smallest layer "Layer1" the output quality of the main function is, for example, 100%, if the output  $Y$  is in "Layer2", then the output quality is 80% until the vector is outside all layers, which may indicate that the main function has a quality of 0%, i.e., the main function is performed incorrectly or not at all.

From this point of view, it is possible to evaluate how the methods and design patterns used for resilience increase really affected the behaviour of the output of the main function and thus the resulting resilience of the system itself.

## 5. EXAMPLE OF PROPOSED RESILIENCE METRICS USAGE

Let's imagine a construction site where an excavator and a truck with a tipping trailer for material transportation are located. For the application, the tipping trailer needs to be filled with a certain material. The main function of the system which is considered here is to fill the tipping trailer correctly, i.e., the output  $Y$  of this main function is to fill the tipping trailer of a certain amount of material of a given type in a given time.

If we consider this particular example, in the space/structure domain, we require to be the tipping trailer filled with material of a certain weight/volume with a maximum defined deviation, i.e. an interval of possible values is given to guarantee the correct functionality of the main function in the space domain (see figure 5). For example, this criterion takes into account the fact that if we supply too much material to the tipping trailer, we could overload it. On the other hand, if we supply too little material, it could lead to a problem or error further in the following systems, or to the non-optimal function of the system or following systems - systems inefficient handling. The criterion in the information domain can be viewed in such a way that the loaded material must meet, for example, a minimum percentage purity in some respects. Thus in this case, for example, the loaded material has to meet at least 95% purity of sand without various admixtures from the structure site or it has to have a certain grain roughness (E.g. sand of a certain fraction), etc. The time-domain criterion can be set up in a way that the tipping trailer needs to be filled with a material in certain time slot, which cannot be exceeded.

Such a system can be affected by various types of disturbance in the form of machine failures, environmental effects, etc. However, if we guarantee that the system will perform its main function despite these disturbances, i.e. output  $Y$  will be within the defined criterion space, we could say about the system, that

the system at the given monitored time interval was resilient to the disturbances that were influencing the system during that interval.

## 6. DISCUSSION

The concept of resilience is very complex and hard to grasp at the moment. The above proposed metric for system resilience measurement is so far an idea and a draft, which does not have a fully expressed mathematical foundation within this paper.

Despite that, the proposed onion model could be also applicable to one of the key features of resilience, namely graceful degradation. Graceful degradation is a term in terms of resilience, where the system goes through several phases when its individual modules with the same or similar functionality fail until complete system failure occurs.[2] The onion model would be appropriate to measure the progress of graceful degradation.

The proposed model is also valid for attributes flexibility and recoverability. Based on that criterion space we could see if the disturbances and affecting adverse conditions will bring the output  $Y$  outside of the criterion space. It would also be possible to find out whether the output  $Y$  will return in the criterion space after the adverse conditions have subsided and in what quality - whether the system is capable of a complete or only partial recovery of its functionality.

## 7. CONCLUSION

In this paper, the concept of resilience was introduced. Also, there was briefly explained, why the concept is needed and why it is valid in the context of a complex system and their evaluation.

The aim of the paper was to bring a draft of a new method of measuring resilience in systems. Within the paper was defined the form of the general system and its main functions. Further, there were listed domains in which these functions are examined and analyzed. Based on these definitions, the concept of resilience measurement was introduced which is based on evaluation of the outputs of the main functions, their qualities and their presence in the criterion space. Last but not least, a general case of a system to which the resilience measurement methodology was applied was presented.

The future direction of this concept is its further extension, to provide a detailed mathematical expression of the problem and to apply this approach to real systems.

## ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-20-6205—“Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0” financially supported by the Internal science fund of Brno University of Technology.

## REFERENCES

- [1] S. Hosseini, K. Barker, and J. E. Ramirez-Marquez, “A review of definitions and measures of system resilience”, *Reliability Engineering & System Safety*, vol. 145, pp. 47-61, 2016, doi: [10.1016/j.ress.2015.08.006](https://doi.org/10.1016/j.ress.2015.08.006).
- [2] V. Castano and I. Schagayev, *Resilient computer system design*. Switzerland: Springer, 2015.
- [3] S. Jackson and T. L. J. Ferris, “Resilience principles for engineered systems”, *Systems Engineering*, vol. 16, no. 2, pp. 152-164, 2013, doi: [10.1002/sys.21228](https://doi.org/10.1002/sys.21228).
- [4] A. Prezenger, “Systems resilience: a new analytical framework for nuclear nonproliferation”, Albuquerque, NM: Sandia National Laboratories, p. 27, 2011.
- [5] B. Allenby and J. Fink, “Toward Inherently Secure and Resilient Societies”, *Science*, vol. 309, no. 5737, pp. 1034-1036, Aug. 2005, doi: [10.1126/science.1111534](https://doi.org/10.1126/science.1111534).



# Study of the effect of tissue conductivity on electroporation

M. Hemzal<sup>1</sup>, V. Aubrecht<sup>1</sup>

<sup>1</sup>Brno University of Technology, FEEC, Department of Power Electrical and Electronic engineering, Czechia

E-mail: [Martin.Hemzal@vut.cz](mailto:Martin.Hemzal@vut.cz), [aubrecht@vut.cz](mailto:aubrecht@vut.cz),

**Abstract**—This paper deals with two different dynamic models of electrical conductivity and a static model for electroporation. Those models were used to calculate electric field distribution in the tissue, the current and the volume of the tissue that will be subjected to electroporation. The results are then compared to the experimental model. Improved model can be used to better predict ablation zones.

**Keywords**—biliary stent, electrical conductivity, electroporation, FEM, occlusion

## 1. INTRODUCTION

Cholangiocarcinoma is a type of cancer that affects the bile duct and is often treated with the insertion of a metallic stent. The patency rate of the metallic biliary stent is reported between 14 and 321 days. Hence, the occlusion of the stent can cause dramatic problems and needs to be dealt with. The occlusion is often treated with mechanical cleaning, radiofrequency ablation, or stent-in-stent placement. However, none have shown satisfactory long-term results [1].

Electroporation is a novel less invasive technique that utilizes pulsed electric fields to induce a transmembrane voltage to either increase permeability of the cell by creating nanopores in the membrane, reversible electroporation (300 V/cm), or to disrupt the membrane to a point that the cell dies, irreversible electroporation (700 V/cm). The success of the treatment is highly dependent on the subjection of the undesirable tissue with sufficient electric field. To accurately predict the ablation zone, numerical model must incorporate parameters such as electrode shape and spacing, applied voltage, or number of pulses. Moreover, due to the nature of the electroporation, electrical conductivity depends on the magnitude of the electric field, because of the creation of pathways with higher electrical conductivity during permeabilization.

This paper focuses on the effect of dynamic electrical conductivity on the distribution of electric field inside the tissue, volume of electroporated tissue and the total electric current. Two models of dynamic electrical conductivity were chosen and the numerical results were compared with the experimental results.

## 2. EXPERIMENTS

Experiments by [1] were performed both in ex vivo and in vivo settings on porcine livers. For the ex vivo setting, tissue bands of liver (3 mm thick) were inserted into the lumen of the stent to model the occlusion. Both active electrodes were in contact with the inserted liver tissue. For the in vivo setting, the lower edge of the liver segment S3 was cut into 3 mm thick pieces. The pieces were then inserted into the lumen of the stent with the follow-up of the catheter. Both active electrodes were in contact with the stent. The protocol of the irreversible electroporation procedure was 100 microsecond pulses of rectangular shape with a frequency of 1 Hz with four different values of voltages 300 V, 650 V, 1000 V and, 1300 V [1].

## 3. DYNAMIC ELECTRICAL CONDUCTIVITY MODELS

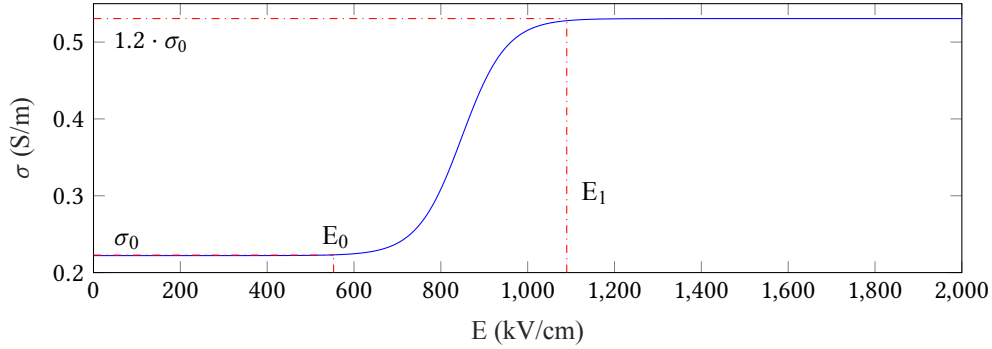
It is a well-known fact that the electrical conductivity varies non-linearly with the strength of the electric field applied to the tissue. The relationship between the electrical conductivity and the electric field is usually described with the help of sigmoidal functions. Two models were selected to calculate the distribution of the electric field and current in the occluded biliary stent scenario.

The first model [2] used in this paper was derived from the measurements done on normal and tumorous human liver tissue. One hundred square-wave pulses with a width of 100  $\mu$ s were delivered to the tissue.

Electrical conductivity was calculated at the end of the first pulse based on the known geometry of the mold that the tissue was inserted into. The initial electrical conductivity was calculated using a low direct current pulse. The values were fitted to the sigmoidal function [2], see Equation 1

$$\sigma(E) = \sigma_0 + \frac{\sigma_1 - \sigma_0}{1 + D e^{\frac{E-A}{B}}}, \quad (1)$$

where the fitting parameters were  $A = 747$ ,  $B = 51.30$ ,  $D = 20.54$ ,  $\sigma_1 = 125.9 \text{ S/m}$ . Graphical presentation of the function is in the Figure 1. The base electrical conductivity of the liver is  $0.22 \text{ S/m}$  and the sigmoidal function plateaus at  $0.53 \text{ S/m}$ . The two turning points  $E_0$  and  $E_1$  at which the electroporation occurs are  $553 \text{ V/cm}$  and  $1090 \text{ V/cm}$  [2].



**Figure 1:** Dynamic electrical conductivity of Model 1 [2].

The second model [3] was derived from experiments on rabbit livers. It contains the dependence of the electrical conductivity on the magnitude of the electrical field and the number of pulses applied. The electrical conductivity was calculated at the end of each pulse at different magnitudes of the electric fields and then it was fitted to the Heaviside function 2 [3]

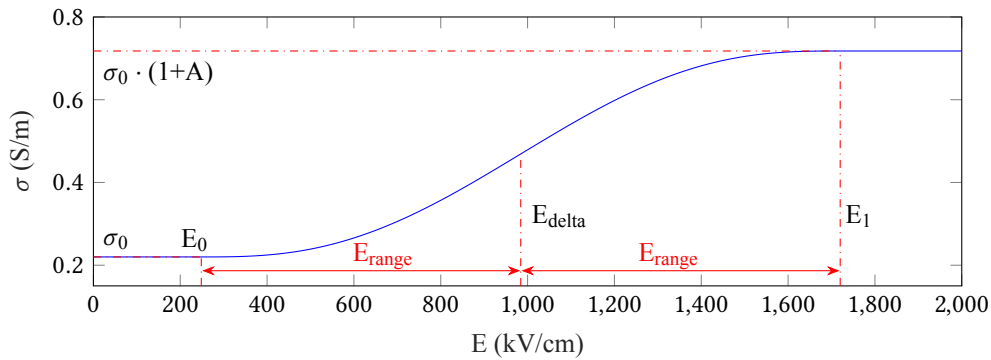
$$\sigma(E, N) = \sigma_0 \cdot (1 + A \cdot \text{flc2hs}(E - E_{del}, E_{range})), \quad (2)$$

where  $\sigma_0$  represents the electrical conductivity before the train of pulses.  $A$  is the increase factor after the treatment. The first turning point  $E_0$  can be calculated by the equation 4. Second turning point  $E_1$  is determined from the measurements. The electrical conductivity saturates at  $0.71 \text{ S/m}$  after 60 pulses [3]

$$A = 2.3919 \cdot (1 - (1 + N \cdot 0.0018^{-1})^{-2.802}), \quad (3)$$

$$E_0 = 8.0861 \cdot 10^3 N^{-1} + 2.9056 \cdot 10^4 e^{-0.0027N}. \quad (4)$$

The parameters  $E_{del}$  and  $E_{range}$  can be determined from Figure 2



**Figure 2:** Dynamic electrical conductivity of Model 2 [3].

#### 4. NUMERICAL MODEL

To compare the experimental results with the numerical simulations, a model of occluded biliary stent was created using COMSOL Multiphysics (v.6, Stockholm, Sweden). It consists of a block of cylindrical liver tissue with the diameter of 35 mm with bile duct inserted into the block with diameter of 10 mm. The bile duct wall is 1 mm thick. A band of liver tissue 3 mm thick was inserted into the lumen of the bile duct to model the obstruction. Two gold plates 10x2 mm with the thickness of 0.15 mm were modeled as electrodes. The stent was modelled as a mesh of nitinol wires with a stent mesh size two by two millimeters. The height of the model is of 100 mm. Half symmetry was used to decrease the computational time.

Dirichlet boundary condition was set on the surface of one electrode with a potential of 300-1300 V the other electrode surface was set to 0 V

$$\Gamma_{e0} = 300 - 1300 \text{ V}, \quad (5)$$

$$\Gamma_{e1} = 0 \text{ V}. \quad (6)$$

On the outer surface of the model, Neumann boundary condition  $\Gamma_n$  was set

$$\Gamma_n = -\frac{\partial \varphi}{\partial n} = 0. \quad (7)$$

To solve the electric potential, Laplace equation 8 is used. The distribution of the electric field was obtained using the gradient of the electric potential .

$$\nabla \cdot (\sigma(E, N) \nabla \varphi) = 0, \quad (8)$$

$$E = -\nabla \varphi. \quad (9)$$

Here the  $\sigma(E, N)$  represents the electrical conductivity of the liver tissue as a function of the electric field, N is the number of applied pulses and  $\varphi$  is an electric potential. Total current delivered to the tissue was calculated by integrating the normal current density over the surface of the electrode

$$I = \iint_S j \cdot dS. \quad (10)$$

Three parametric simulations were performed to evaluate the effect of the electrical conductivity on the distribution of the electric field and the current inside the tissue. One with the static electrical conductivity of  $\sigma_0 = 0.22 \text{ S/m}$ , another two with the dynamic models described in Section 3, all with four different voltages of 300 V, 650 V, 1000 V and 1300 V respectively. The distribution of the electric field was evaluated on a cross-sectional area of the model in the middle of the electrodes.

#### 5. RESULTS

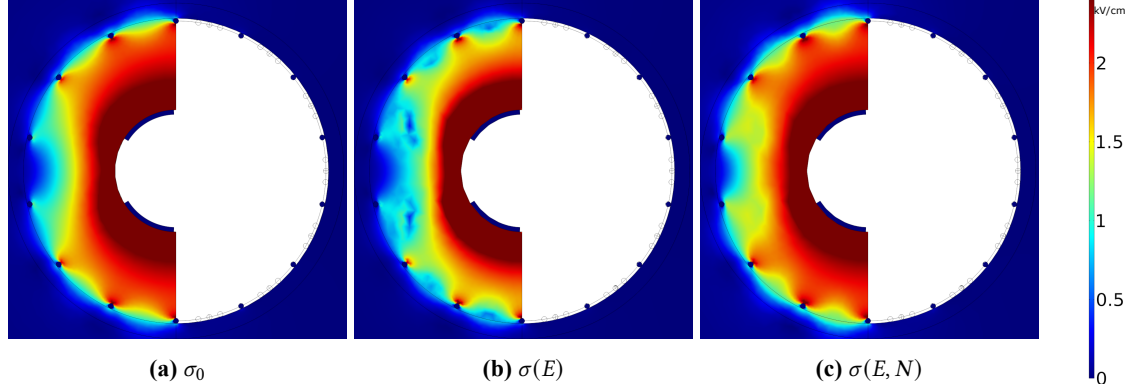
The numerical results were compared with the experimentally obtained current values in Table I. The currents calculated by the model with static electrical conductivity are much lower than those obtained by the experiments. The first model predicts higher currents, but the best fit to the experimental data is performed by the second model, which incorporates the number of pulses applied. The distribution of

**Table I:** Computed and measured currents

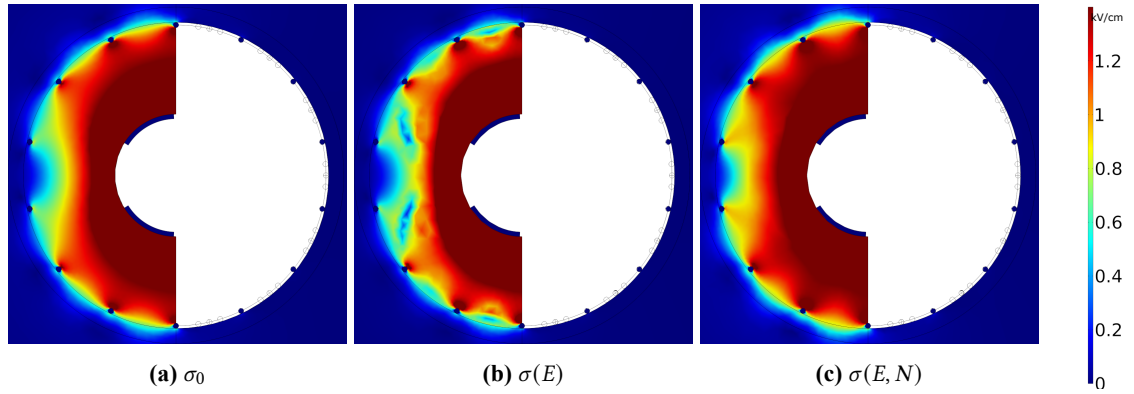
Model	300V	650 V	1000 V	1300 V
experiments	1.1 A	3.1 A	6 A	-
$\sigma$	0.63 A	1.36 A	2.1 A	2.7 A
$\sigma(E)$	0.81 A	2.82 A	4.73 A	6.2 A
$\sigma(E, N)$	0.92 A	3.49 A	6.2 A	8.3 A

the electric field in the tissue in all models is restricted to the inserted tissue that models the obstruction. This is caused by the stent that behaves as Faraday cage like device and protects the healthy tissue from

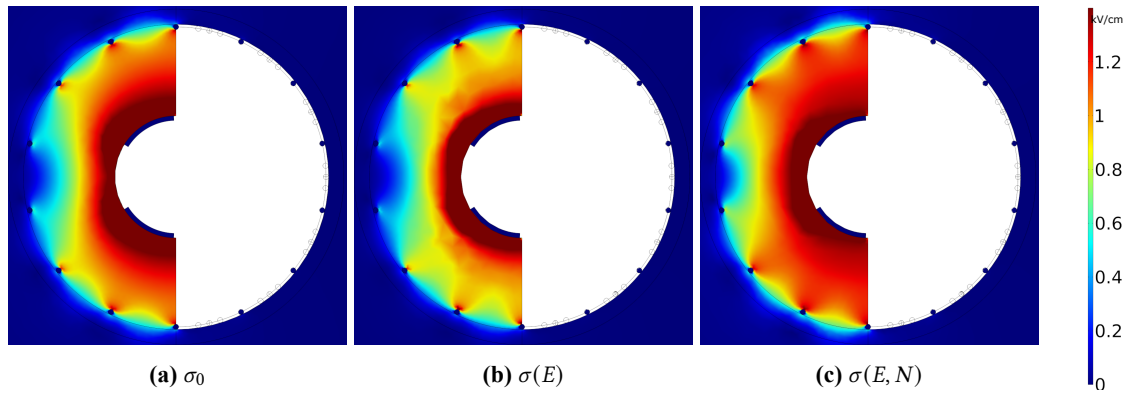
electroporation. The models of  $\sigma_0$ ,  $\sigma(E)$  and  $\sigma(E, N)$  allows for direct comparison of the electrical conductivity effect on the electric field distribution (EFD). For the 1300 V the EFD differs slightly. Higher intensities of the electric field for the dynamic models are around electrodes and wires of the stent, Figure 3. For the lower voltages 1000 V and 650 V, Figures 4 and 5 the change in the EFD is more noticeable with higher intensities of the electric field around wires of the stent. Also the EFD is more concentrated between the electrodes for the  $\sigma(E)$  and  $\sigma(E, N)$  unlike the model  $\sigma_0$  where the EFD is distributed more evenly around electrodes.



**Figure 3:** Electric field distribution at 1300 V



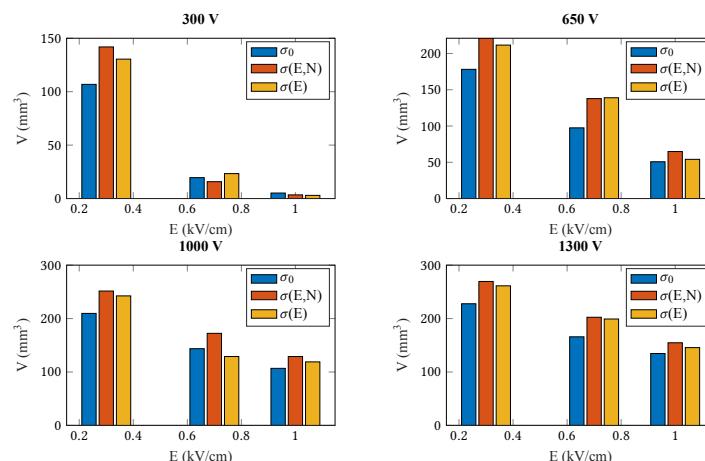
**Figure 4:** Electric field distribution at 1000 V



**Figure 5:** Electric field distribution at 650 V

This leads to higher ablation volumes in models with the dynamic electrical conductivities. Ablation volumes were calculated for all models, see Figure 6. The thresholds were chosen to represent reversible

electroporation (300 V/cm) and irreversible electroporation (700 V/cm). The value of 1000 V/cm represents more conservative value for irreversible electroporation.



**Figure 6:** Ablation volumes

## 6. CONCLUSION

Due to the fact that it is hard to measure or visualize the electroporation effect during the surgical procedure, it is important to have an accurate model to visualize the ablation volume and the distribution of temperature in the tissue ahead of time. The dynamic electrical conductivity is an important parameter in the numerical model that influences both the electric field distribution and the temperature. To yield a successful irreversible electroporation procedure it is important to subject tissue to a sufficient electric field, typically 700 V/cm. On the other hand, higher intensity of the electric field does not contribute to the electroporation effect and leads to higher temperatures during the procedure which can ultimately lead to the thermal damage of the tissue. The best prediction of the total average current delivered to the tissue is presented by the model  $\sigma(E, N)$  with 10 % error rate.

The electrical conductivity is influenced by other factors such as the temperature, length of the pulse and frequency of the pulses. In the future we want to incorporate those parameters in the numerical model to better predict the parameters that are needed for successful accomplishment of the electroporation procedure.

## ACKNOWLEDGMENT

This research work has been carried out in the Centre for Research and Utilization of Renewable Energy (CVVOZE). Authors gratefully acknowledge financial support from the Ministry of Education, Youth and Sports under institutional support and BUT specific research programme (project No. FEKT-S-20-60379). This work was also supported by the Ministry of Health of the Czech Republic, grant no. NU21-08-00561.

## REFERENCES

- [1] T. Rohan, T. Andrasina, T. Juza, P. Matkulcik, D Cervinka, I. Svobodova, V. Novotna, V. Bernard, V. Valek and S. Goldberg, "Experimental model of occluded biliary metal stent recanalization using irreversible electroporation via a tubular catheter", *INTERNATIONAL JOURNAL OF HYPERTHERMIA*, vol. 38, no. 1, pp. 393-401, 2021, [doi.org/10.1080/02656736.2021.1875061](https://doi.org/10.1080/02656736.2021.1875061).
- [2] N.B. White, S. Bhonsle, R.C.G. Martin, R.V. Davalos, "Electrical Characterization of Human Biological Tissue for Irreversible Electroporation Treatments, *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 4170-4173, 2018, [doi: 10.1109/EMBC.2018.8513341](https://doi.org/10.1109/EMBC.2018.8513341).
- [3] Z. Yajun, Z. Shuang, N.B. White, L.Hongmei, Y. Chenguo, R.V. Davalos, "Development of a Multi-Pulse Conductivity Model for Liver Tissue Treated With Pulsed Electric Fields", *Frontiers in Bioengineering and Biotechnology*, vol. 8, 2020, [10.3389/fbioe.2020.00396](https://doi.org/10.3389/fbioe.2020.00396).

# History of EVs and measuring battery under external force

M.Šedina<sup>1</sup> and T.Kazda<sup>1</sup>

<sup>1</sup>Department of Electrical and Electronic Technology, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technická 10, 616 00 Brno, Czech Republic

E-mail: [xsedin00@vut.cz](mailto:xsedin00@vut.cz), [kazda@vut.cz](mailto:kazda@vut.cz).

**Abstract**— This article describes the modern history of electromobility with the prediction of its market growth and the drivers of change. There are mentioned Li-ion battery's anode materials with their basic parameters and one of the conversion materials, silicon, which represent one of the possible materials for enhancing the capacity of the anode. Usage of the Si opens a few problems which are unnoticeable with actual anode materials, but there are some possible ways how to solve them. One of their solutions could be external force applied on the battery surface, which should reduce inflation of the battery and improve internal conduction.

**Keywords**— Electromobility, Si anode, Li-ion battery, External force

## 1. INTRODUCTION

Every day, we use devices powered by batteries, whether they are smartphones, wearables, laptops, tablets, or devices that we need for our work. There are more than 7 billion people in the world, and everyone needs transport. One of the clean transporting trends is electromobility. This trend opens a new challenge of researching batteries. The research is mostly focused on longer lifetime, higher capacity, or lighter batteries.

## 2. MODERN ELECTROMOBILITY

The history of electromobility is older than we can think. Firsts prototypes of electric vehicles (EVs) were made in the first half of the 19<sup>th</sup> century and the 10s of the 20<sup>th</sup> century fought face to face with internal combustion vehicles (ICV), but Ford's mass production and longer driving range wins the fight and production of EVs was suspended in the time of WW2. Then during the rest of the 20<sup>th</sup> century, there were few EVs, their production was typically driven by an oil crisis.

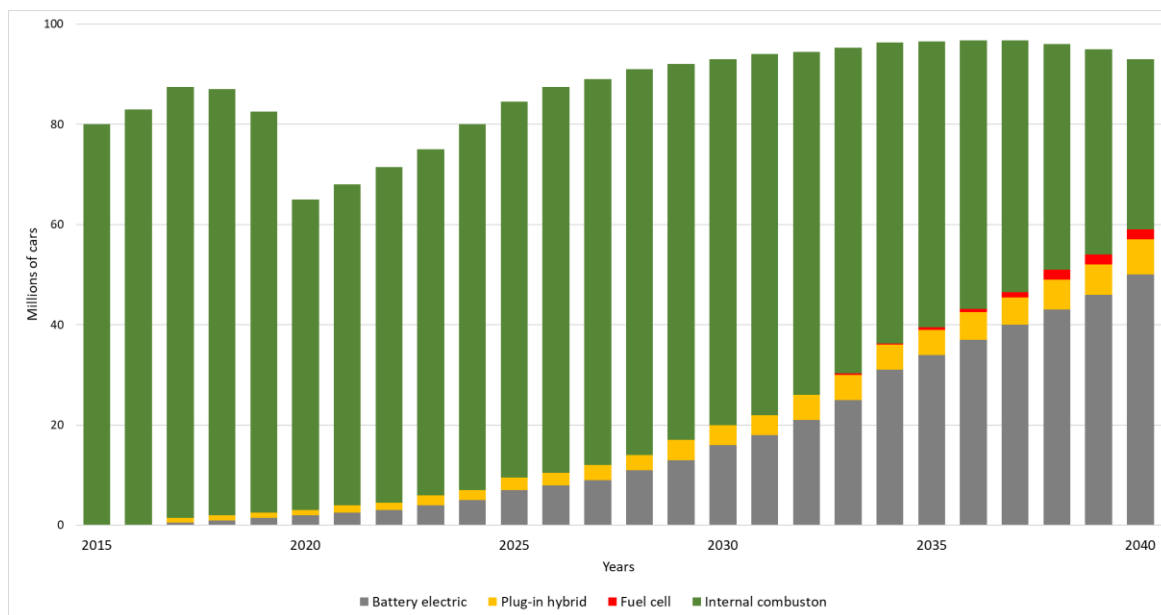
A new restart of the powertrain electrification in the automotive sector started Toyota with their model Prius Hybrid in 1997. Prius was the symbol of ecological travelling, and his sales were successful, the highest sales had on the west coast of the USA.

In 2008 Elon Musk with his company Tesla wanted to test the market with their fully electric car Tesla Roadster. Tesla Roadster is based on the Lotus Elise and is the first serial EV with Li-ion batteries. Thanks to them, Tesla Roadster was able to reach a range of more than 300 km. Tesla sold almost 2500 cars and showed that EVs have a future [1][2].

In the 10s of the 21<sup>st</sup> century, almost every producer starts to experiment with EVs, some with prototypes, some with production cars. As a representative prototype should be mentioned Škoda Octavia Green eLine, which is based on second-generation Skoda Octavia and served for Press presentation [3]. One of the production cars, which should be mentioned is the BMW i3 which was on the market from 2013 and will ends this year. The special thing about this car is, that in the early sales years had a version called Rex (Range extender), which means that the car had a small IC engine that



work as a generator for an extended range of the car. This conception did not work very well, so BMW withdrew it from the market [4].



**Figure 1:**Global annual passenger vehicle sales by drivetrain[5]

Figure 1 clearly shows that electric vehicles will have a place in the future on the market. The assumption is that by 2040 at least half of the cars will be equipped with alternative drives (fuel cell, battery). It can be assumed that these conditions may still change, as some carmakers already mention that they plan to stop selling internal combustion vehicles around the year 2030, alone the EU is considering banning new ICVs from 2035. It is possible to assume that all operations of ICVs will be banned in the EU, as it is written in the Green Deal [5][6].

### 3. ANODE

Nowadays generation of Li-ion batteries uses two types of anode material. The first and the most common material is graphite in configuration  $\text{LiC}_6$  theoretical capacity  $372 \text{ mAhg}^{-1}$  and potential 50-250 mV vs Li/Li<sup>+</sup>. This material can be used with all cathode materials[7]. Another type is made from ceramic  $\text{Li}_4\text{Ti}_5\text{O}_{12}$  (LTO), this type of anode is used for  $\text{LiMn}_2\text{O}_4$  (LMO) and  $\text{LiNi}_{1-x-y}\text{Mn}_x\text{Co}_y\text{O}_2$  (NMC) cathode materials. LTO has a working potential of less than 1.55 V vs Li/Li<sup>+</sup> with a theoretical capacity of  $175 \text{ mAhg}^{-1}$  [8].

One of the ways to improve Li-ion batteries is solved by increasing the capacity of the anode. The most promising conversion material, which can lead to the increased capacity of the anode is silicon. This material has a more possible used modification, but mostly tested is  $\text{Li}_{15}\text{Si}_4$  with theoretical capacity  $3579 \text{ mAhg}^{-1}$  with potential lower than 0.5 V vs Li/Li<sup>+</sup>[9]. The advantages of silicon are in its low ecological impact and large amount of it in Earth's crust. On the other hand, there are a few problems with the usage of Si on the anode. The biggest problem is large volume change during cycling (up to 300 %), which leads to damage of SEI and its re-grow and the battery inflation. Another thing connected with volume changes is the cracking and disintegration of Si particles. The result of this is a rapid decrease of capacity, which makes usage of Si on anode difficult[10].

#### 4. MEASURING BATTERY UNDER EXTERNAL FORCE

My research is focused on testing batteries under constant external force, which I moderate on battery by the holder with metal clamps and spring, as batteries we use pouch cells which are used by Škoda Auto in their EVs (see Figure 2). Forces on the springs used for the experiment were 50, 100, and 200 N on each spring. The springs have a maximum ( $F_m$ ) force of 641.8 N and a spring constant ( $R$ ) of 13.22 N/mm. For the setting moderated force, we use equation (1),  $F$  is moderated force and  $s$  is spring stroke.

$$F = R * s \quad (1)$$

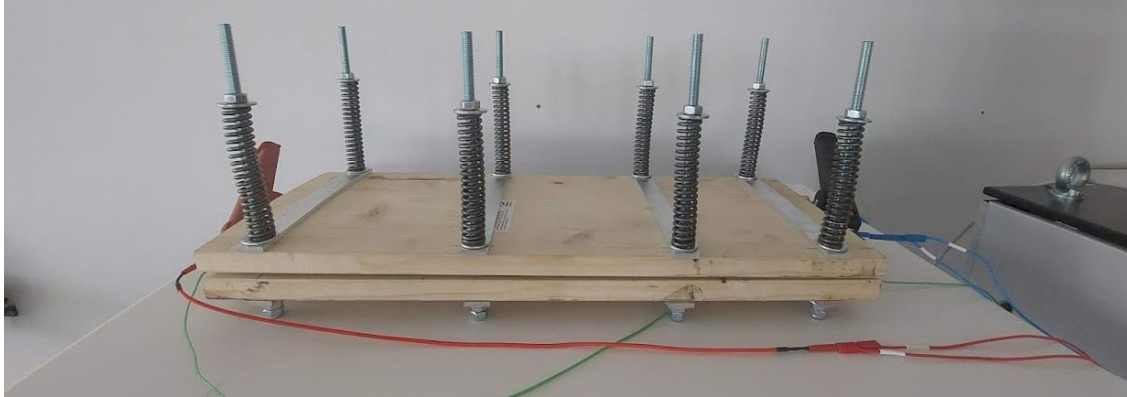
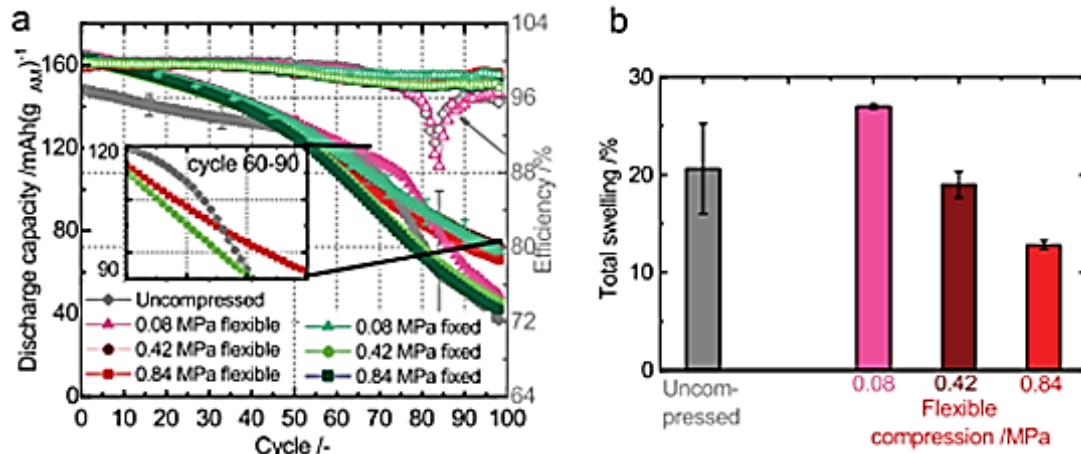


Figure 2: Battery during the measurement

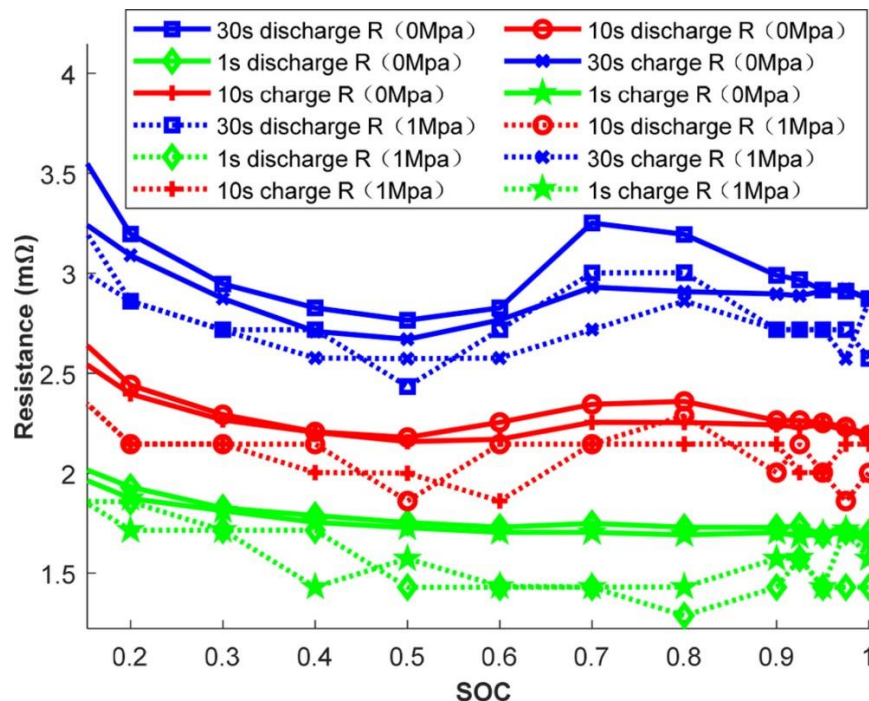
At this moment I am using four pouch cells, one is used for uncompressed measurement and one for each force. The battery is compressed during the cycling and measuring electrochemical impedance spectroscopy and then is uncompressed and run the parametric test.

There are several studies focused on this research. The study of Müller et al. was focused on testing different levels of pressure and style of compression affecting the battery. In their experiment, they used uncompressed cells, flexible compressed (with springs), and fixed compression. Figure 3a shows the results of long-term cycling. Compression had a positive effect on the stability of the capacity drop and increase starting capacity. Best results were achieved with flexible compression, which shows almost two times higher remaining capacity than the uncompressed cell. Figure 3b then shows results of an inflation measurement, where pressure 0.84 MPa wins with low diversification of the result and total swelling only 13 % [11].



**Figure 3:** Effect of compression on battery with Si anode [11]

Another study made by Zhou et al. compares the function of NMC pouch cells under pressure (1 MPa) and without pressure (0 MPa). New cells show that before the experiment the pressure did not affect the capacity, the result was verified by OCV. Their result shows that external pressure can effectively reduce the internal resistance of the battery (in Figure 4). The study also showed tests on the aged battery. The capacity was raised by 1.6 % just by using the external pressure on the battery. The pressure didn't affect OCV on the aged battery. After the measurements, the batteries were left for two weeks and after the decompression of the compressed battery, they measured an irreversible increase of capacity[12].



**Figure 4:** Comparison of internal resistance before and after pressure [12]

## 5. CONCLUSION

The article is describing the modern history of electromobility and the prediction of the market share of EVs in the future with their drivers.

In the article are described current anode materials with their parameters and new possible material for anode with its advantages and disadvantages. There is described one of the possible ways to reduce the problem with the usage of Si and how external pressure can positively affect the battery.

There is also described my mechanism for generating the external force and the methodology of my measurement.

## ACKNOWLEDGMENT

This work was supported by the specific graduate research of the Brno University of Technology No. FEET-S-20-6206. This work was supported by Škoda auto.

## REFERENCES

- [1] “Worth the Watt: A Brief History of the Electric Car, 1830 to Present”, *Car and Driver*, 02018. [Online]. Available: <https://www.caranddriver.com/features/g15378765/worth-the-watt-a-brief-history-of-the-electric-car-1830-to-present/>. [Accessed: 14-Feb.-2022].
- [2] “The History of the Electric Car”, *ENERGY.GOV*, 02014. [Online]. Available: <https://www.energy.gov/articles/history-electric-car>. [Accessed: 07-Mar.-2022].
- [3] “Skoda - Octavia Green E Line Concept”, *ALL CAR INDEX*. .
- [4] “BMW i3 REX – Mám rád elektromobily!”, *Auto.cz*, 02016. .
- [5] *Intelligence in Mobile Battery Applications (D5.1 Desk Research & Data Analysis IMBA – Release 1)*. 2019.
- [6] “Zelená dohoda pro Evropu”, *Evropská komise*, 02019. [Online]. Available: [https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal\\_cs](https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_cs). [Accessed: 14-Feb.-2022].
- [7] S. Schweidler, L. de Biasi, A. Schiele, P. Hartmann, T. Brezesinski, and J. Janek, “Volume Changes of Graphite Anodes Revisited: A Combined Operando X-ray Diffraction and In Situ Pressure Analysis Study”, *The Journal of Physical Chemistry C*, vol. 122, no. 16, pp. 8829-8835, Apr. 2018.
- [8] T. -F. Yi, S. -Y. Yang, and Y. Xie, “Recent advances of  $\text{Li}_4\text{Ti}_5\text{O}_{12}$  as a promising next generation anode material for high power lithium-ion batteries”, *Journal of Materials Chemistry A*, vol. 3, no. 11, pp. 5750-5777, 2015.
- [9] Y. H. Xu, G. P. Yin, and P. J. Zuo, “Geometric and electronic studies of  $\text{Li}_{15}\text{Si}_4$  for silicon anode”, *Electrochimica Acta*, vol. 54, no. 2, pp. 341-345, 2008.
- [10] F. Luo et al., “Review—Nano-Silicon/Carbon Composite Anode Materials Towards Practical Application for Next Generation Li-Ion Batteries”, *Journal of The Electrochemical Society*, vol. 162, no. 14, pp. A2509-A2528, Oct. 2015.
- [11] V. Müller et al., “Effects of Mechanical Compression on the Aging and the Expansion Behavior of Si/C-Composite/NMC811 in Different Lithium-Ion Battery Cell Formats”, *Journal of The Electrochemical Society*, vol. 166, no. 15, pp. A3796-A3805, Nov. 2019.
- [12] L. Zhou et al., “A study of external surface pressure effects on the properties for lithium-ion pouch cells”, *International Journal of Energy Research*, vol. 44, no. 8, pp. 6778-6791, Jun. 2020.

# SEM Micromanipulator as Current Collector for Battery In-situ Characterization

O. Klvač<sup>1</sup>, T. Kazda<sup>1</sup>, Y. Fam<sup>2</sup>, and L. Novák<sup>2</sup>

<sup>1</sup> Department of Electrical and Electronic Technology, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technická 10, 616 00 Brno, Czech Republic

<sup>2</sup> Thermo Fisher Scientific Brno, Vlastimila Pecha 12, 627 00, Brno, Czech Republic

E-mail: [xklvac02@vut.cz](mailto:xklvac02@vut.cz), [kazda@vut.cz](mailto:kazda@vut.cz), [yakub.fam@thermofisher.cz](mailto:yakub.fam@thermofisher.cz), [libor.novak@thermofisher.cz](mailto:libor.novak@thermofisher.cz),

**Abstract**—The aim of the work is to develop an in-situ method for electrode microstructure investigation by scanning electron microscopy (SEM). In the first part, the preparation of a model battery using FIB-SEM system is described, which consists of lithium metal, ionic liquid, and lithium titanate oxide (LTO) attached on a micromanipulation needle by platinum deposition. The needle is used for handling and at the same time as a current collector, which facilitates subsequent analysis. Open-circuit voltage (OCV), cyclic voltammetry (CV), and galvanostatic cycling measurements (GCPL) was done on the cell assembled (and imaged) in SEM. After seven complete cycles, a cross-section of LTO electrode using FIB was prepared and the surface solid electrolyte interface (SEI) layer was observed. Challenges connected with battery preparation are described and possible solutions for future testing are proposed. The work is evolved in cooperation with Thermo Fisher Scientific Brno.

**Keywords**— SEM, FIB, in-situ, Li-ion, LTO

## 1. INTRODUCTION

Lithium-ion batteries (LIBs) are secondary energy sources with high energy density, low self-discharge, and long life. They are widely used in portable electronics, electric vehicles, and energy storage systems [1], [2]. In recent years, the demand for these types of batteries has increased mainly due to the rapidly evolving electromobility and the need to store energy from renewable sources [3]. For these reasons, Li-ion cells are one of the fastest growing markets with an annual rate of about 13.8% and considerable attention is paid to research [4].

Typically used electrochemical methods (GCPL, CV, PEIS, ...) can determine the number of important cell parameters, such as capacity, internal resistance, plateau position, and more. The cause of changes in these parameters can be estimated on a theoretical level only. However, better knowledge is essential for research, performance improvement, and the application of new materials [5].

Larger-scale changes at the cell design level can be visualized non-destructively using CT methods [6]. However, for processes at the microstructure level, the resolution is not sufficient. For this, the battery is usually discharged after a defined number of cycles, disassembled, and the electrode surface or cross-section is scanned [7], [8]. During this, the cell is destroyed and a direct correlation with electrochemical data is missing. Direct characterization can be performed using XRD in special in-situ cells, on the other hand, the observed changes are spatial averages only [9].

In-situ electron microscopy allows both – direct characterization at the microstructure level of individual grains. As the system operates in a vacuum, it is needed to prevent electrolyte evaporation. It is possible to use environmental chambers, or replace typical electrolytes with ionic liquids or solid state electrolytes. It is also necessary to use low electron beam currents for SEM imaging to avoid experiment distortion by local charging or lithium plating. Several in-situ battery cycling approaches for SEM [10], [11], [12] and TEM [13], [14] have been published.

This work is focused on the design of an experimental setup of a simple half-cell consisting of lithium metal, ionic liquid, and the examined sample using an SEM-FIB system. The sample is attached to a micromanipulation needle, which also serves as a current collector during electrochemical measurements. This allows using any electroactive material, quick disassembling the half-cell, and observing grain-level changes induced by cycling.



## 2. EXPERIMENTAL

For the in-situ experiment, Scios 2 DualBeam SEM (Thermo Fisher Scientific) with gallium FIB was used. The experimental scheme is shown in Figure 1.

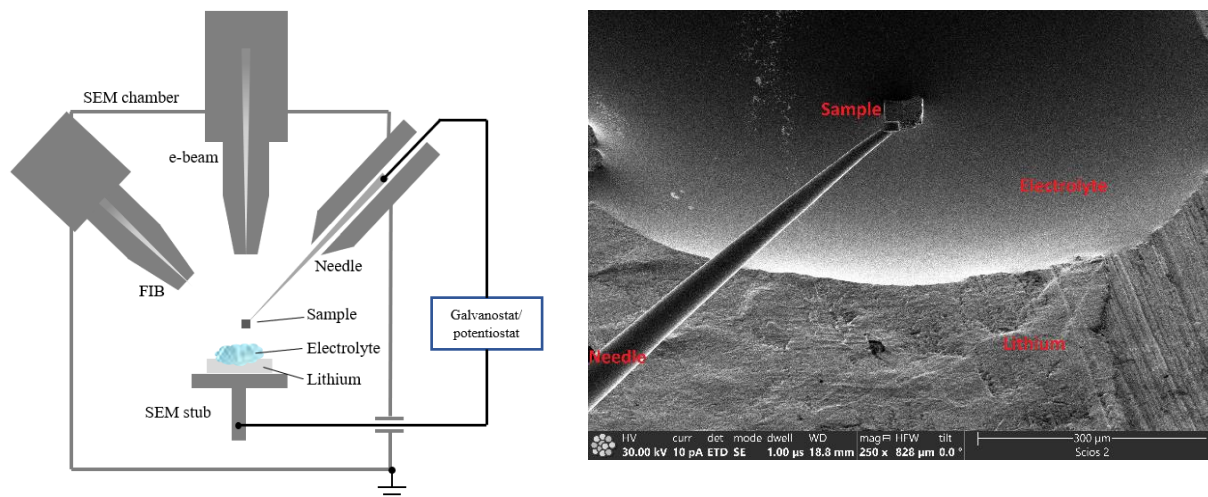
The battery system was created on the surface of the sample stub which serves as the current collector for lithium metal fixed by double-side carbon conductive tape. Sample stub is insulated from the SEM chamber and connected with galvanostat/potentiostat via vacuum feedthrough.

Lithium (99.9% pure ribbon, SigmaAldrich) was stored in an Ar-filled glovebox, and transferred in a sealed glass vial. It was attached to the stub immediately before closing the chamber and dripped with ionic liquid using a pipette. As an ionic liquid, 1-Ethyl-3-methylimidazolium tetrafluoroborate (EMIMBF  $\geq 98\%$ , SigmaAldrich) in a mixture with LiBF<sub>6</sub> salt in 0.5 mol/l concentration was used.

As the investigated sample, a chunk of approx. 50x50x20  $\mu\text{m}$  was cut from an electrode with LTO anode active material coated on aluminum foil produced by CUSTOMCELLS<sup>®</sup> using xenon plasma focused ion beam (PFIB) system Helios Hydra 5 UX DualBeam (Thermo Fisher Scientific). Several chunks were cut on one electrode sheet, which was then transferred to an experimental SEM chamber. The samples were not cut completely, a small „bridge“ for fixation was left. Then, the EasyLift micromanipulation needle was attached to the top-left chunk corner, glued by platinum deposition using MultiChem Gas Delivery System and the „bridge“ was removed (standard lamella preparation procedure). The manipulator needle as well as the stub were insulated from the device frame and connected to a galvanostat/potentiostat.

For electrochemical characterization, BioLogic SP-150 galvanostat/potentiostat was used, which allows precise setting of low currents in the order of nanoamps. This is important because the capacity of the sample is in units of nAh. The half-cell was connected with a shared reference electrode (RE) and the counter electrode (CE) to the lithium (stub), LTO was on the working electrode (needle). The ground electrode was on the microscope frame.

The OCV was measured during the assembly of the cell. One of the chunks glued to the needle was lifted out of the electrode sheet and partially soaked in electrolyte. The CV and GCPL techniques were used for subsequent electrochemical characterization.



**Figure 1:** Experimental scheme of performed experiment (left); SEM image of created half-cell (right)

## 3. RESULTS AND DISCUSSION

When the sample touched the electrolyte surface, the OCV was 1.73 V. After 16 seconds it rose to 1.84 V. After complete immersion, the voltage stabilized at 2.21 V. Then the GCPL was set to a current of  $\pm 5$  nA. However, the discharge current was around -7 nA, so it was set to -2 nA after 73 s. Then, the resulting discharge current was around -4 nA, the charging current was similar. A total of 7 full cycles were performed in different voltage ranges (see detail in Table 1).

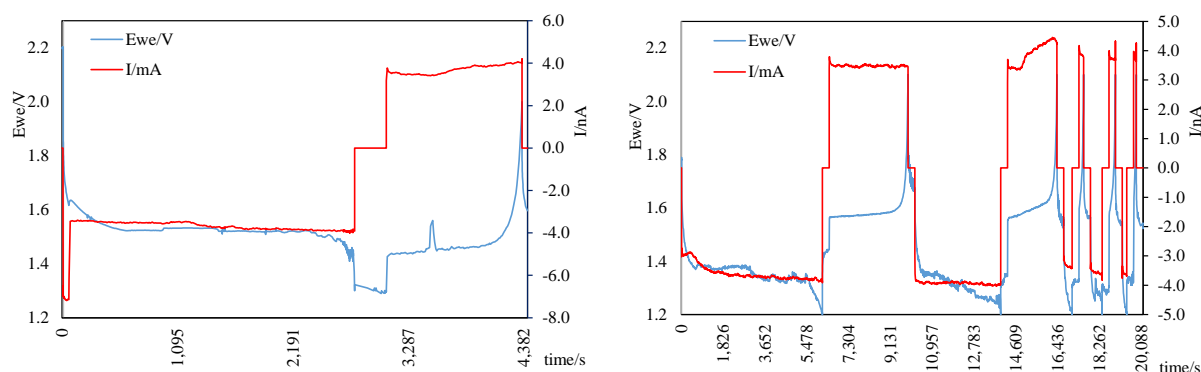


**Table 1:** Cycling conditions and achieved capacity

	Ewe [V]	Q_discharge [nAh]	Q_charge [nAh]
1. cycle	1.3 - 2.0	2.8	1.32
8 hod break			
2. cycle	1.2 - 2.1	6.14	3.33
3. cycle	1.2 - 2.2	4.09	2.34
4. cycle	1.2 - 2.3	0.34	0.22
5. cycle	1.2 - 2.4	0.5	0.29
6. cycle	1.2 - 2.5	0.2	0.12
CV			
7. cycle	1.3 - 2.0	2.03	1.07
discharging	up to 1.0	6.09	–

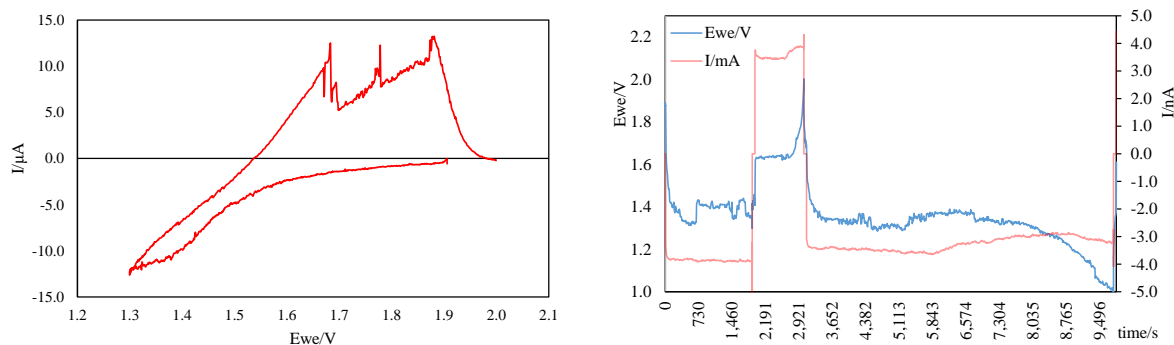
In the case of the first cycle (Figure 2; left), the discharge plateau is clearly visible in the 1.5V range, which is the expected value. However, the signal is noisy and the drop to the 1.3 V limit occurs suddenly. There is a clear decrease in capacity during the subsequent charging. There is also a noticeable small peak, which is probably caused by interference during the taking of a control image.

This was followed by a break of about eight hours, after which 5 more cycles were performed (Figure 3; right). Capacity has risen about twice in the case of the second cycle, but the discharge plateau has shifted lower (charging, on the contrary, higher) and it is not stable. With each cycle, the capacity decreased rapidly.

**Figure 2:** First performed cycle (left); next cycles after 8 hours relaxation (right)

The CV characteristic (0,5 mV/s) also showed the shift of the reduction peak (figure 3; left). In the voltage range of 1.3 V - 2.0 V, only the ramp is visible, but not the maximum. The oxidation peak contains 3 spikes. However, only one is expected for LTO material. The current was probably too high and there were some electrolyte reactions.

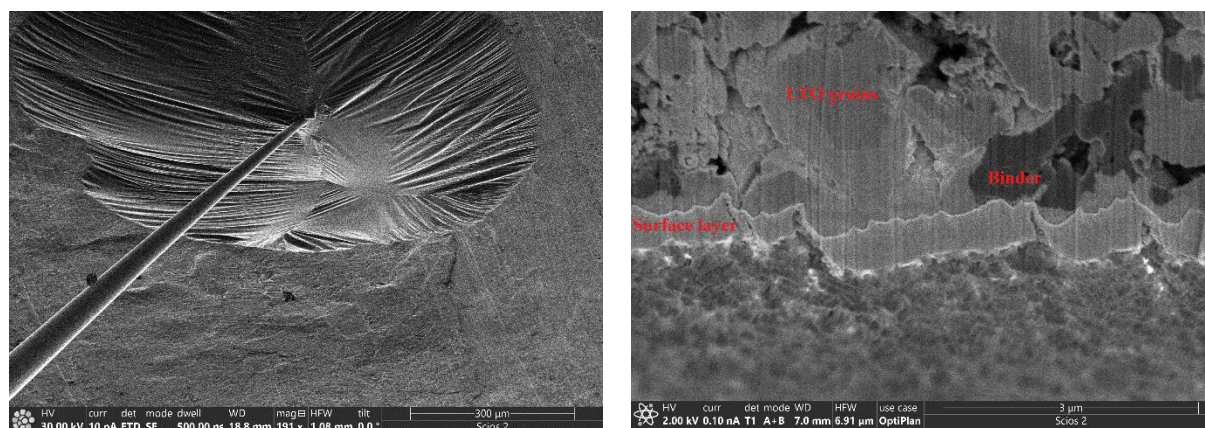
The last cycle (figure 3; right) shows that capacity has been partially restored. However, the discharge ended prematurely, due to strong noise voltage momentarily falling below the set limit of 1.3 V. The cell was, therefore, charged again and the limit of 1.0 V was set during the final discharge. Obviously, the battery is still functional, but in a lower voltage range than expected.



**Figure 3:** CV characteristic (left); final cycle and discharging (right)

After 24 hours, electrolyte side reactions were evident. On its surface was an observable solid matter resembling a shell that could be cracked by moving the needle (figure 4; left). Due to the small thickness, EDS analysis could not reveal composition of the shell.

The cross-section was then prepared using FIB. It is a well observable matrix of grains surrounded by a conductive binder. An approximately 400 nm thick SEI layer was formed on the chunk surface during the experiment.



**Figure 4:** IL surface after 24 hrs (left); sample cross-section (right)

After opening the SEM chamber, the originally transparent electrolyte was yellow-brown and contained small bubbles. The same reaction occurred upon contact of IL with lithium in the glovebox. It is clear that the electrolyte is not suitable for longer cycling if lithium metal is present.

#### 4. CONCLUSION

A micron-sized battery system based on lithium titanate, ionic liquid, and metallic lithium was successfully created and tested in the SEM chamber. The maximum OCV was 2.21 V immediately after assembly. A total of 7 full cycles and one CV measurement were performed, maximum capacity reached 6,14 nAh. To achieve this capacity, it was necessary to let the battery relax for 8 hours. During repeated cycles without relaxation, there is a rapid loss of capacity. After the first cycle, the discharge plateau moved from the original 1.55 V to about 1.3 V. This may be due to an increase in internal resistance during SEI layer formation. This may be due to an increase in internal resistance during SEI layer formation. The measurement is greatly affected by noise. During the experiment, EMIMBF electrolyte side reactions with lithium metal were evident, which probably affected the measurement.

It has been verified that the proposed battery assembly procedure is working, but a different ionic liquid must be used. TFSI in mixture with LiTFSI currently appears to be promising. For the next experiments, it is planned to use the currently tested inert gas transfer system to transport lithium from the glovebox directly to the SEM chamber under an argon atmosphere without exposure to air.

It has been shown that the use of a modified micromanipulation needle as a current collector is possible and advantageous for the analysis of electrode materials.

## ACKNOWLEDGMENT

This work was supported by the specific graduate research of the Brno University of Technology No. FEKT-S-20-6206. The work was developed in cooperation with Thermo Fisher scientific Brno. Yakub Fam, Libor Novák and Tomáš Kazda participated in the preparation of the experiment.

## REFERENCES

- [1] Y. Liang et al., InfoMat, 1, 6-32 (2019).
- [2] Y. Tao et al., Science Advances, 7, 6-32 (2021).
- [3] H. -J. Kim et al., Electronics, 9 (2020).
- [4] Grand View Research, Inc., 2020.
- [5] L. Li and X. Zhang, (I2MTC), p. 1-5, IEEE (2018)
- [6] P. Pietsch and V. Wood, Annual Review of Materials Research, 47, 451-479 (2017)
- [7] Kovachev et al., Batteries, 5 (2019)
- [8] T. Waldmann et al., Journal of The Electrochemical Society, 163, A2149-A2164 (2016)
- [9] H. Liu et al., Journal of Power Sources, 306, 300-308 (2016)
- [10] D. J. Miller et al., Advanced Energy Materials, 3, 1098-1103 (2013)
- [11] D. Chen, et al., Journal of Power Sources, 196, 6382-6387 (2011).
- [12] European Microscopy Congress 2016: Proceedings, Wiley-VCH Verlag GmbH & Co., Weinheim, Germany, (2016).
- [13] J. Y. Huang et al., Science, 330, 1515-1520 (2010)
- [14] X. H. Liu et al., Nano Letters, 11, 2251-2258 (2011)

# Nanoporous Enzyme-Based Impedimetric Biosensor

Z. Košelová<sup>1</sup>

<sup>1</sup>Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Microelectronics, Technická 10, Brno, 612 00, Czech Republic

E-mail: [225675@vut.cz](mailto:225675@vut.cz)

**Abstract**—The ordered nanopores surfaces has been studied in the field of biosensors for the detection of affinity interactions such as antigen-antibody or nucleic acid hybridization. Nanopores are frequently fabricated by anodic oxidation of aluminium. This method is practically simple method of highly ordered pore formation, but is not suitable for massive production. The principle of detection is usually based on the degree of pore blockage by biomolecular interaction that allows a sensitive electrochemical response. Here we show the development and characterization of the nanopore-based enzymatic biosensor for the detection of analytes. We performed and tested the concept of the simple and versatile method of nanopores formation on the electrode using non-conductive polystyrene nanoparticles. The variability in pore size was modulated by nanoparticles diameter (40 and 80 nm). We measured response of our designed biosensor using electrochemical impedance spectroscopy in presence of redox mediator at low potential. The smaller diameter of nanopores increases charge transfer resistance. Biosensors with 80 nm nanoparticles had behaviour similar to reference without nanoparticles, therefore the pore was too wide for this design. The limit of detection from  $R_{ct}$  showed to be at most 0.01 mM for every sample type. This proposed biosensor contributes to the basic knowledge about nanopores sensors, and it could have a potential, for example, in the development of “point of care” devices.

**Keywords**—polystyrene nanoparticles, nanopores, electrochemical impedance spectroscopy, enzyme

## 1. INTRODUCTION

During last decade the interest increased on sensors modification with nanomaterials. Nanopores membranes conducted to biosensors are one of the promising fields [1]. They bring advantages such as enlargement of surface area, provide restriction, and increase selectivity. This turned out to be useful for third-generation sequencing or a direct detection event between affinity partners (e.g. DNA hybridization, antigen-antibody) [2]. The basic step for the creation nanopore-based biosensors is the formation and allocation of a solid-state nanopores system on the sensing surface. There are different techniques for forming nanopores. Some of them, such as ion beam etching, require sophisticated equipment, others are simpler in the requirements (anodization) [3] [4]. The porous anodic alumina has been frequently reported. The anodization of aluminum changes the surface of the metal through oxidation to produce an oxide layer with controllable pore widths, thickness, and length. Even if the method is simple and inexpensive, the reproducibility and massive production can be limiting [4]. Besides, packing of spherical nanoparticles (NPs) in a dense planar arrangement can be an alternative way [5] [6]. The nanopores are considered spaces between NPs, where the nanopore width is determined by NPs diameter. Thus, by specifying NPs material and shape, it is possible to produce a system suitable for a wide range of applications. In this paper, nonconductive spherical polystyrene nanoparticles were used to form porous matrices on the electrodes. NPs are uniform, which supports homogeneity of the formatted monolayer. The affinity phenomena in nanopores are usually measured electrochemically in the presence of redox mediators. Any change in the size of the nanopores (i.e. blocking of pores) influences the molecule flow, and the electrochemical response can be monitored. In this work, we immobilize the enzyme as the biocatalyst on the nanopores matrix to detect the analyte. Enzyme sensors are of high interest as detection tools in clinical and food analysis. However, the electrochemical characterization of enzyme biosensors based on the polystyrene nanoparticles (PS NPs) nanopores

matrices is rarely reported, especially the effect of different nanoparticle size and thus pores on a measured electrochemical signal. In fact, the pore size in our proposed experimental setup did not change during the biochemical detection of the analyte. Instead, the nanopores matrix serves as the stable enzyme support and the production of reduced mediator by enzyme is detected through the mediator re-oxidation on the electrode. To test this concept, electrochemical impedance spectroscopy was used. We applied glucose oxidation (GOD) in the presence of the ruthenium mediator ( $\text{Ru}^{3+}$ ) which was reduced ( $\text{Ru}^{2+}$ ) due to enzymatic reaction with glucose. Then, the calibration curve for the glucose concentration was measured. This well-known principle was applied to observe the effect of NPs matrices and their pore diameter on the enzymatic biosensor. For further analysis of impedance spectra, charge transfer resistance ( $R_{ct}$ ) was chosen. It can be obtained from impedance spectra by fitting the Randles equation [7]. The  $R_{ct}$  is a measure of the difficulty encountered when an electron is shifted from one atom or compound to another atom or compound. Therefore, the redox reaction that occurred on an electrode surface can be measured. This factor is influenced by the layers through which the redox probe penetrates. It could help to monitor changes of the biosensor, which has an NPs matrix with different pore sizes.

## 2.1 CHEMICALS

Carboxylated polystyrene nanoparticles (PS NPs) of diameter  $\approx 40$  and  $80$  nm were purchased from Magsphere Inc. (U.S.A.). Poly-L-lysine hydrobromide (p-lys), glucose oxidase (GOD) from *Aspergillus Niger*, bovine serum albumin (BSA), Hexaammineruthenium(III) chloride (98%) ( $\text{Ru}^{3+}$ ), potassium ferrocyanide and potassium ferricyanide ( $\text{Fe}^{2+}/\text{Fe}^{3+}$ ), glutaraldehyde solution (GO) (25%), were purchased from Sigma-Aldrich (Germany). D-(+)-Glucose monohydrate, KOH,  $\text{H}_2\text{O}_2$  (30%), isopropyl alcohol was purchased from Penta (Czech Republic).

## 2.2 ELECTRODES

The commercial electrodes in a bipotentiometric setup (www.printed.cz, Czech Republic) of a pair of two identical gold disk electrodes were used in this work. The electrode has internal diameter  $400\ \mu\text{m}$  and there is no separate reference or counter electrode. We immobilized the negatively charged PNPs of the diameters  $\approx 40$  nm and  $\approx 80$  nm on the positively charged poly-L-lysine (p-lys) thin layer. Before the modification of the electrodes, the “activation” of the gold electrodes was performed. The gold electrodes were polished with microcloth (Buehler) and isopropyl alcohol to remove the residual photoresist and passivation layers from the surface. The electrodes were then treated with the solution of  $0.5\ \text{M}$  KOH and  $20\%$   $\text{H}_2\text{O}_2$  for  $10$  min.

The gold electrode was functionalized by the adsorption of  $50\ \mu\text{g/mL}$  p-lysine in phosphate-buffered saline (PBS;  $10\ \text{mM}$ , pH  $7.4$ ) for at least  $30$  min to provide amino groups on the gold surface.

After washing and drying steps, a drop of  $1\%$  w/v PS NPs solution was applied on the positive charge - modified electrode for  $40$  min.

The mixture of  $16\ \mu\text{L}$  GOD ( $8\ \text{mg/mL}$ ) and  $25\ \mu\text{L}$  BSA ( $16\ \text{mg/mL}$ ) was then cross-linked with  $2.75\ \mu\text{L}$   $2\%$  glutaraldehyde solution (GO). All were dissolved in PBS ( $10\ \text{mM}$ ; pH  $7.3$ ). The microliter drop was applied to the NPs modified electrodes and let dry. The reference sample was created as  $33\ \mu\text{L}$  BSA ( $13\ \text{mg/mL}$ ) with  $2.75\ \mu\text{L}$   $2\%$  glutaraldehyde solution. The second reference was the cross-linked GOD enzyme with BSA, but without NPs.

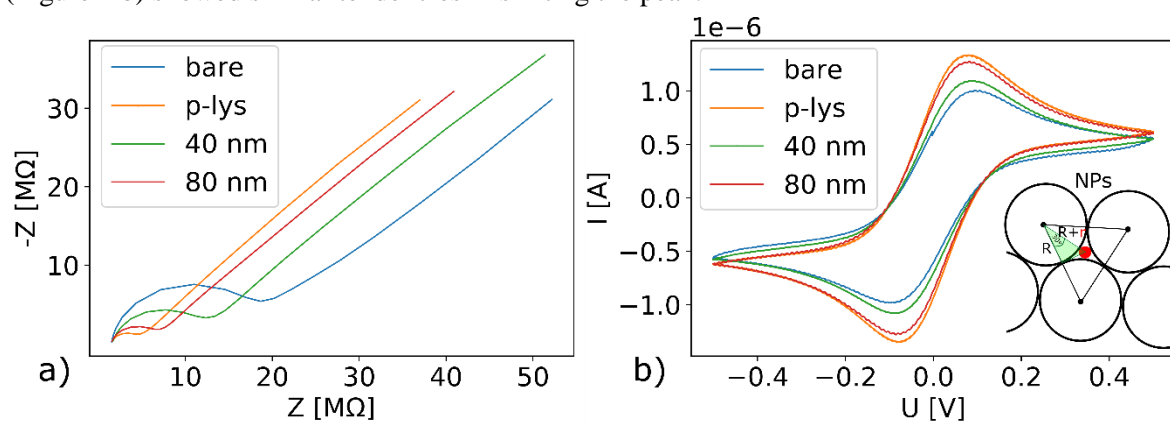
Electrochemical impedance spectra (EIS) and cyclic voltammetry (CV) were recorded using  $\mu\text{AUTOLAB III / FRA2}$  (Metrohm Autolab, Netherlands) analyzer. For EIS enzymatic measurements, the electrode was placed in a beaker with  $2\ \text{mL}$  of  $5\ \text{mM}$   $[\text{Ru}(\text{NH}_3)_6]^3$  dissolved in PBS. The  $50\ \text{mV}$  amplitude potential was applied with a logarithmic distribution of  $20$  individual frequencies from  $100,000$  to  $0.05\ \text{Hz}$ . The first two measurements were made without glucose. Then we started measure glucose concentration series to get a calibration curve. Glucose was added from the solution made of  $0.1\ \text{M}$  D-(+)-Glucose monohydrate and  $5\ \text{mM}$   $[\text{Ru}(\text{NH}_3)_6]^3$  dissolved in PBS.

CV and EIS for confirmation of the modification layers were performed in  $50\ \text{mM}$   $\text{Fe}^{2+}/\text{Fe}^{3+}$  in

10/140~mM PBS (7.4 pH) as a redox probe. CV measurements were made in the range of -0.5 to 0.5 V. The cycle was repeated twice for each sample. The scan rate was 0.1 V/s and the step potential was 2.44 mV.

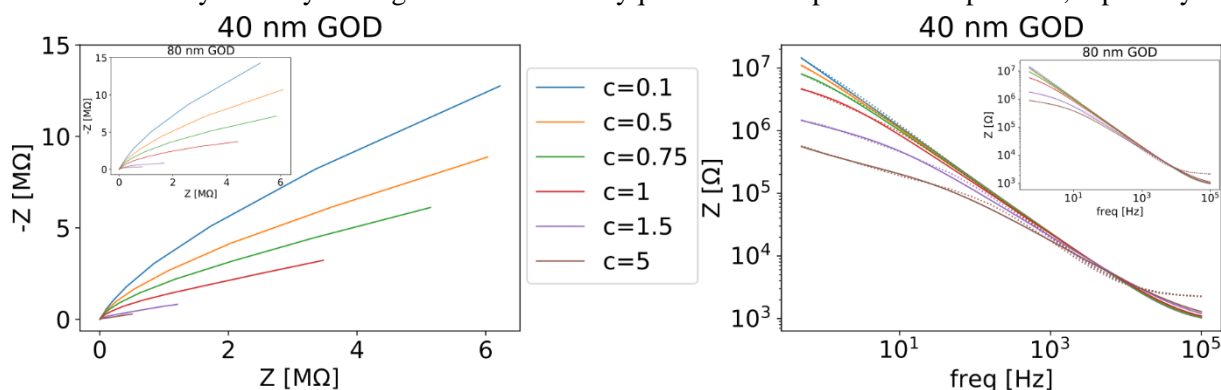
## 2. RESULTS AND DISCUSSION

The immobilization of the NPs monolayer was already evaluated as sufficiently homogeneous, dense, and with reproducible preparation elsewhere [5]. We confirmed the modification layers using impedance spectroscopy and cyclic voltammetry measurements in  $\text{Fe}^{2+}/\text{Fe}^{3+}$  (Figure 1). The result was consistent with previous report [5] [9]. P-lys immobilization caused impedance decrease, because of amino groups presence, which are positively charged and attract negatively charged redox probe. Therefore, the conductivity near the electrode surface increased and the impedance decreased [9]. NPs immobilization caused an increase in impedance. It could be caused by the occupation of positive amino groups with negatively charged NPs and, by the formation of a blocking layer of polystyrene nanopores matrix on the electrode. The NPs packing close to themselves create pores in empty interspace. The close packing of NPs creates pores, as shown in the inset of Figure 1b by red circle. The size of the pores is estimated to be around  $\approx 6.2$  nm and  $\approx 12.4$  nm in diameter for 40 nm and 80 nm NPs, respectively. The smaller pores significantly increased the impedance compared to 80 nm pores. CV measurements (Figure 1 b) showed similar tendencies in shifting the peak.



**Figure 1:** Graph of Nyquist plot (a) and CV (b) for the electrodes after each modification step measured in the presence of the  $\text{Fe}^{2+}/\text{Fe}^{3+}$  redox probe. The inset on the right schematically outlines the geometry used to estimate the nanopore radius ( $r$ ) from the NPs radius ( $R$ ). 40 and 80 nm represents the NPs diameters by which the pores were created.

Then we measured electrodes that had the 40 and 80 nm NPs covered with a film of the GOD enzyme immobilized by the cross-linking method. The GOD cross-linked on p-lys without NPs was used as a reference for behavior without NPs matrix. The protein BSA cross-linked on NPs without enzyme, which still can react on others factors (conductivity, viscosity), was used to distinguish the immobilized effects caused by the enzyme. Figure 2 shows the Nyquist and Bode plots. The impedance, especially at



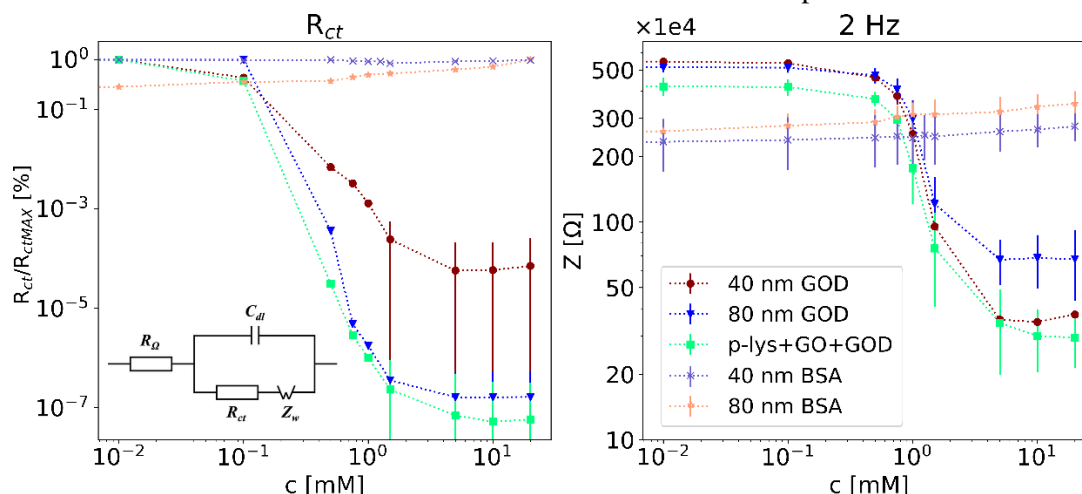
**Figure 2:** Nyquist plot (a) and Bode plot (b) of the chosen glucose concentration responses of GOD on a nanopores layer formed from 40 nm NPs - GOD (inserts show the results for GOD immobilized on 80 nm NPs monolayer). 40 nm and 80 nm represent the NPs diameters by which the pores were created.



lower frequencies, decreased with increasing glucose concentration. The shape of the curve in the Nyquist graph was changing and, for high glucose concentration (about 1.5 mM), an arc began to form. Then in those cases visible lines on figure 2 were diffusion tails.

The decrease in impedance corresponds to the reaction occurring on the electrode surface. The current increased as a result of the re-oxidation of  $\text{Ru}^{2+}$  produced by the GOD reaction with glucose on the electrode. The changes in the electrochemical behavior of the modified electrodes were further expressed by the charge transfer resistance  $R_{ct}$ , representing the difficulty encountered when an electron is shifted from one atom or compound to another (Figure 3 a).  $R_{ct}$  (and its standard deviation  $\sigma_c$ ) was calculated by fitting the data using the Randles equation. Due to the easier flow through the pores of redox molecules in 80 nm GOD compared to 40 nm GOD, it was expected that  $R_{ct}$  reaches lower values, which turned out to be true. This applies to 0.5 – 1 mM glucose concentration. Higher values have a large uncertainty. Between 80 nm GOD and p-lys+GO were just small differences, which can suggest, that roughly 12.4 nm pores are overly wide to play a significant effect.  $R_{ct}$  is linear only in concentration in the log-log scale,  $\log(y) = a \log(x) + b$ . That can be modified into:  $y = x^a B + c$ . Typically, the calibration curve is linear, but there are studies concerning non-linear functional relationships [10][11][12][13].

The calibration curve for the impedance measured at selected frequencies can be obtained by linear fitting  $y = ax + b$  (Figure 3 b shows 2 Hz as example). It is necessary to choose lower frequencies for the transfer of massive ions to manifest. For all samples, including BSA, the  $R_{ct}$  standard deviations for concentrations up to 1 mM are much smaller than the difference of two adjacent  $R_{ct}$  values. The limit of detection (LOD), defined as the glucose concentration for which the signal changes by  $3.2\sigma_0$  from the value at zero concentration of  $R_{ct}$  measurement is  $\approx 0.01$  mM, as shown in Table 1. The fitting parameters for  $R_{ct}$  and impedance at a 2 Hz voltage frequency from the calibration curve and their calibration area were also shown. Due to  $\text{Ru}^{2+}$  were not created for reference BSA samples, there was absence of calibration curve and therefore the table did not contain their parameters.



**Figure 3:**  $R_{ct}$  parameter dependence on glucose concentration. The values were normalized by maximal  $R_{ct}$  value for given type of sample.  $R_{ct}$  was calculated by fitting Randles equation corresponding to showed circuit (a). Graph (b) shows impedance measurement for 2 Hz voltage frequency as example of impedance linear calibration. 40 nm and 80 nm represents NPs diameters by which pores were created.

**Table 1:** The table shows the calibration area, parameters from the calibration curves, and LOD for the  $R_{ct}$  and impedance measurements at 2 Hz frequency. a and b for  $R_{ct}$  is from  $y = x^a B + c$ , slope for impedance is from  $y = ax + b$

Sample	$R_{ct}$				Z for 2 Hz		
	Calibration area [mM]	a	B [ $\Omega \cdot \text{mM}^{-1}$ ]	LOD [mM]	Calibration area [mM]	a [ $\Omega \cdot \text{mM}^{-1}$ ]	LOD [mM]
40 nm GOD	0.1 – 1.9	-2.57	$2.6 \cdot 10^6$	0.01	0.3 – 1.6	$-3.7 \cdot 10^6$	0.06
80 nm GOD	0.1 – 1.2	-4.88	$6.5 \cdot 10^6$	0.01	0.4 – 1.7	$-3.6 \cdot 10^6$	0.06
Plys+GO	0.1 – 1.7	-5.84	$4.8 \cdot 10^6$	0.01	0.3 – 1.6	$-3.0 \cdot 10^6$	0.02

### 3. CONCLUSION

There are already simple methods suitable for the fabrication of nanopores, such as oxidation of aluminum. Although this method is capable of highly ordered pore formation, it is not suitable for massive production and the reproducibility could be low. This paper worked with the packing of spherical nanoparticles (NPs) in a dense planar arrangement, which can be an alternative method. The negative charge of the nanoparticles causes the formation of the nanoparticle monolayer due to the repulsive effect, as was previously shown [5] [8]. Therefore, without complex facilities can be created nanopores nonconductive layer, at which it is possible to immobilize enzyme. We showed that the nanopore layer has an impact on the biosensor response in impedance measurement. The smaller diameter of the nanopores caused an increase in the  $R_{ct}$  parameter, especially in the range of 0.5 mM to 1 mM glucose concentrations. For higher glucose concentrations,  $R_{ct}$  had large uncertainties and possibly could reach as small values as samples with 80 nm NPs. The LOD for  $R_{ct}$  was smaller than 0.01 mM, so we are able to measure change for low concentrations. The better arrangement (e.g. smaller pores, or different enzyme immobilization techniques) may bring even better improvement of the calibration area. Further exploration of the properties of nanopores biosensors could be useful, for example, in the development of “point of care” devices.

### ACKNOWLEDGMENT

Created with financial support from project IGA – ID 6215 - New approaches in the use of modern micro- and nanoelectronics and project IGA – ID 7861 - Development of nanopores electrochemical enzyme-based biosensor.

### REFERENCES

- [1] X. Yang *et al.*, “Mesoporous Materials–Based Electrochemical Biosensors from Enzymatic to Nonenzymatic,” *Small*, vol. 17, no. 9, p. 1904022, Mar. 2021, doi: 10.1002/smll.201904022.
- [2] A. Santos, T. Kumeria, and D. Losic, “Nanoporous anodic aluminum oxide for chemical sensing and biosensors,” *TrAC - Trends in Analytical Chemistry*, vol. 44. Elsevier B.V., pp. 25–38, Mar. 01, 2013, doi: 10.1016/j.trac.2012.11.007.
- [3] Q. Chen and Z. Liu, “Fabrication and Applications of Solid-State Nanopores,” *Sensors*, vol. 19, no. 8, p. 1886, Apr. 2019, doi: 10.3390/s19081886.
- [4] S. Manzoor, M. W. Ashraf, S. Tayyaba, and M. K. Hossain, “Recent progress of fabrication, characterization, and applications of anodic aluminum oxide (AAO) membrane: A review,” Dec. 2021, Accessed: Mar. 02, 2022. [Online]. Available: <http://arxiv.org/abs/2112.08450>.
- [5] J. Sopoušek, J. Věžník, P. Skládal, and K. Lacina, “Blocking the Nanopores in a Layer of Nonconductive Nanoparticles: Dominant Effects Therein and Challenges for Electrochemical Impedimetric Biosensing,” *ACS Appl. Mater. Interfaces*, vol. 12, no. 12, pp. 14620–14628, Mar. 2020, doi: 10.1021/acsami.0c02650.
- [6] A. de la Escosura-Muñiz, M. Espinoza-Castañeda, M. Hasegawa, L. Philippe, and A. Merkoçi, “Nanoparticles-based nanochannels assembled on a plastic flexible substrate for label-free immunosensing,” *Nano Res.*, vol. 8, no. 4, pp. 1180–1188, Apr. 2015, doi: 10.1007/s12274-014-0598-5.
- [7] F. Lisdat and D. Schäfer, “The use of electrochemical impedance spectroscopy for biosensing,” *Analytical and Bioanalytical Chemistry*, vol. 391, no. 5. Springer, pp. 1555–1567, Jul. 16, 2008, doi: 10.1007/s00216-008-1970-7.
- [8] J. Sopoušek, J. Humlíček, A. Hlaváček, V. Horáčková, P. Skládal, and K. Lacina, “Thick nanoporous matrices of polystyrene nanoparticles and their potential for electrochemical biosensing,” *Electrochim. Acta*, vol. 368, p. 137607, Feb. 2021, doi: 10.1016/j.electacta.2020.137607.
- [9] K. Lacina, J. Sopoušek, V. Čunderlová, A. Hlaváček, T. Václavěk, and V. Lacinová, “Biosensing based on electrochemical impedance spectroscopy: Influence of the often-ignored molecular

- charge,” *Electrochem. commun.*, vol. 93, pp. 183–186, Aug. 2018, doi: 10.1016/j.elecom.2018.07.015.
- [10] J. Martin, A. R. Gracia, A. G. Asuero, J. Martin, A. R. Gracia, and A. G. Asuero, “Fitting Nonlinear Calibration Curves: No Models Perfect,” *J. Anal. Sci. Methods Instrum.*, vol. 7, no. 1, pp. 1–17, Mar. 2017, doi: 10.4236/JASMI.2017.71001.
- [11] Y. H. Yun *et al.*, “A nanotube array immunosensor for direct electrochemical detection of antigen–antibody binding,” *Sensors Actuators B Chem.*, vol. 123, no. 1, pp. 177–182, Apr. 2007, doi: 10.1016/J.SNB.2006.08.014.
- [12] Y. Fong, K. Sebestyen, X. Yu, P. Gilbert, and S. Self, “nCal: an R package for non-linear calibration,” *Bioinformatics*, vol. 29, no. 20, p. 2653, Oct. 2013, doi: 10.1093/BIOINFORMATICS/BTT456.
- [13] R. R. Mustafa, R. Sukor, S. Eissa, A. N. Shahrom, N. Saari, and S. M. Mohd Nor, “Sensitive detection of mitragynine from *Mitragyna speciosa* Korth using an electrochemical immunosensor based on multiwalled carbon nanotubes/chitosan- modified carbon electrode,” *Sensors Actuators B Chem.*, vol. 345, p. 130356, Oct. 2021, doi: 10.1016/J.SNB.2021.130356.

# Lithographic thin-film structures based on Electrochromic materials: Case study by scanning probe microscopy

Denis Misiurev

Brno University of Technology, The Czech Republic

E-mail: xmisiu00@vutbr.cz

**Abstract** – The main object of this study will be based on producing and characterization of lithographic thin-film structures produced based electrochromic materials. Atomic force microscopy will be used as a main method of evaluation the morphology of lithographic structures. The optical properties of eslectrochromic materials will be analyze using scanning near-field optical microscopy by applying different electrical potential.

**Keywords** – AFM, SNOM, electrochromic materials, thin-films.

## 1. INTRODUCTION

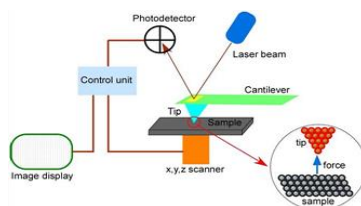
The electrochromic materials possess the ability of changing their optical properties by applying different electrical potential (usually 1–5V). By changing optical properties (transmission, reflection, polarization), the electrochromic materials change their coloration. Consider the ability of changing optical properties, the lithographic structures based on electrochromic materials have been found as an essential component of quantum computers.

Quantum computers in their functionality use optical phenomena such as superposition, diffraction, quantum entanglement and polarization to enhance data processing. The benefits of quantum entanglement are extremely important for numerous advancing calculations.

Physical phenomena such as interference, reflection and entanglement will be characterized by scanning near-field optical microscopy to find optimal parameters to enhance the performance of electrochromic thin-films. In addition, scanning near-field optical microscopy can detect distribution of light energy. The morphology will be provided by atomic force microscopy.

## 2. ATOMIC FORCE MICROSCOPY

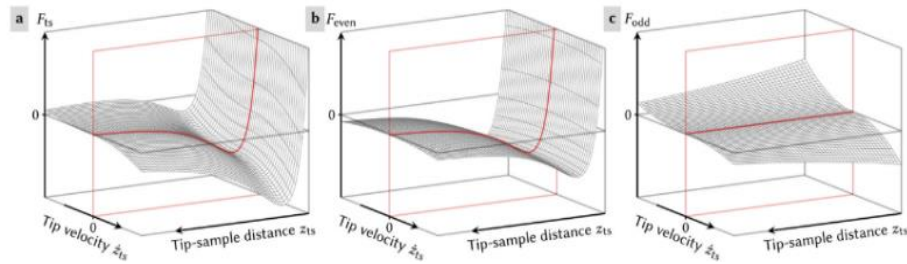
Atomic force microscopy (AFM) is important and widely spread tool in area of surface science and obtaining molecular/nano scale image [1]. The resolution of image depends on sharpness of the needle. AFM system can provide image with resolution beyond diffraction limit [2]. The tool shows extreme sensitivity as a profilometer to measure surface topology and morphology [1]. Main advantage of AFM in comparison to SPM, AFM can obtain image of wide range [3] of conductive and nonconductive materials [4]. AFM can also obtain more precise 3D image [5] and does not require special conditions to initiate analysis [1].



**Figure 1:** Schematic representation on AFM system [4]

### 3. FUNCTION PRINCIPLE

The working principal of AFM is based on mechanical [6] scanning surface by sharp needle (probe). The cantilever is vital part of AFM. The needle is extremely sharp and attached to the end of the cantilever [5]. The force of interaction between probe and sample is subject of investigation. The moment of interaction of the tip and sample can be approximated with help of three variables: deflection, the maximal and the phase difference and frequency. These variables allow to obtain information about probe–sample interaction (distribution of attached force) [7].

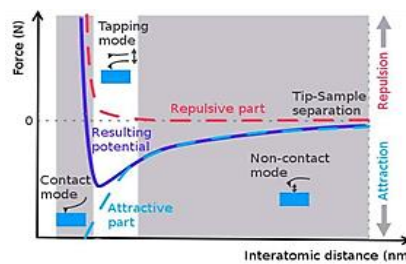


**Figure 2:** probe sample interaction curve [8]

The probe interacts with underlying [6] surface of tested material and simultaneously being registered by photodetector. Photodetector is represented by photodiode or laser which records movement of probe [6]. The AFM allows not only to obtain very precise topography of surface but also various physical properties [6]. AFM can operate in different modes [4]. Operated mode is dependent on interaction distance between the probe and surface [4].

### 4. LENNARD-JONES CURVE

The moment of interaction of the probe with the scanning sample can be described using Lennard-Jones curve. The curve is represented by potential energy of intermolecular distance [9]. As probe gets closer to the sample the potential energy (repulsive force) decreases very quickly and then has slower increase (attractive force) [9]. The Lennard-Jones curve is divided into several different part [10]. Every part describes different moments of interaction between the probe and sample. The part there the potential energy of intermolecular interaction decreased is corresponded with repulsive resalting force [10]. The repulsive force is related to the Pauli principle, stopping pair electrons from intersecting.



**Figure 3:** Lennard-Jones curve [8]

The 0 point means what the probe got in the contact with the surface of the sample. In other words, the repulsive force is equal to attractive. Right after the initiating the contact the probe will interact the surface, which is represented by local minimum of the Lennard-Jones curve. At this point, attractive force will slowly increase [8].

### 5. SCANNING NEAR-FIELD OPTICAL MICROSCOPY

Although AFM possesses outstanding imaging capabilities, it only delivers tomographic information about the sample surface. To obtain information about the optical properties of surfaces on the nanometer scale scanning near-field optical microscopy (SNOM) has been employed for several years

[11]. The SNOM is simple, relatively complex, self-compactness system, which provides optical and topographic image of sample surface.

The SNOM developed in 1980s has gained huge popularity in a field of researching surface sciences due to SNOMs unique advantages.

## 6. ADVANTAGES OF SNOM

The main advantage of SNOM is considered breakthrough of diffraction limit [12] of optical image. The typical resolution of near-field microscope is 50–100 nm and depends on size of the sample, probe, working mode [13]. Such resolution cannot be achieved in far-field microscopy because it exceeds diffraction limit which is given by wavelength.

The considerable advantage of SNOM is versatility of measuring scale of sample. The sample can be measured in large and small scale in high resolution. Lower measuring scale leads to shorter measure period [12].

SNOM can also provide additional information about optical properties of measured surface. In addition to optical contrast, SNOM allows to obtain chemical and topographical contrast [14].

Lastly, SNOM is capable to detect light energy distribution. The detection provides correspondence of light distribution and structure layout.

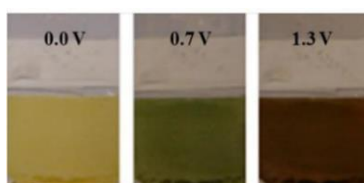
## 7. ELECTROCHROMIC MATERIALS

Electrochromism is a unique [15] ability of materials reversibly changes their optical properties under applying electrical potential [17].

The ability of changing optical properties makes them incredible for producing a numerous devices and applications such as smart coating for windows [16], reflecting surfaces or camouflage, energy harvesting, temperature sensors, display modules [18], charging elements etc [15].

There are main parameters which are concerned for evaluation electrochromic materials: speed of switching coloration, performance, ability to recreate optical image, resistance to environmental conditions. The efficiency of electrochromic materials is measured using time which is required to change of bleached color to different color.

The continuous development of electrochromic materials brings up new materials which possessed multi-colored electrochromic properties [19].



**Figure 4:** Example of electrochromic materials color changing capabilities [19]

Electrochromic materials can be used as an essential part of thin-coating and furthermore for changing optical transmission signal [20]. The integration of electrochromic materials into thin-films for quantum computers has been used to obtain higher performance of transmitted signal.

## 8. PHYSICAL PRINCIPLE OF ELECTROCHROMISM

Electrochromism is related to changing coloration of electrochromic materials under influence of voltage or electric field [22]. The effect is caused by dual transferring and eviction of electrons(cathode) and ions (anode) [25]:

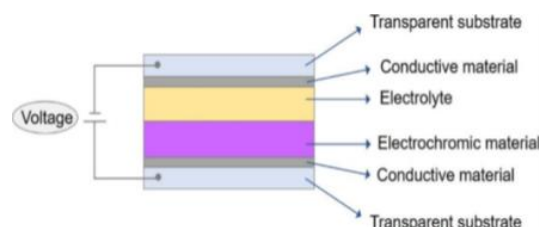


[25] As typical ions can be ions of silver, lithium etc.



## 9. DESIGN DESCRIPTION OF POSSIBLE STRUCTURES

Most of electrochromic films are produced in the form of multi-layer structure. The basic structure of electrochromic device consists of glass substrate and conductive medium. The substrates are located on both sides of material and have function of electrode and counter electrode. By applying voltage between electrodes, the conductive medium changing their coloration due to transferring of electrons and ions movement across medium [23].



**Figure 5:** Layer layout of electrochromic material film [23]

## 10. SAMPLES PREPARATION

The main materials for preparing electrochromic material are mostly oxides of different metals. Metal oxides such as wolfram, molybdenum, vanadium and titanium oxides have stronger electrochromic effect among others and primarily used in preparation of electrochromic materials.

The most important aspect of this work will include preparation of lithographic thin-film structures based on Electrochromic materials. The absence of technology or organization for preparing such structures in the Czech Republic causes additional complication to the overall research. As a potential solution of the issue involves long-term cooperation with foreign universities across the Europe.

There are many technics which allowed producing lithographic thin-film structures. However, the deeper description of the technology will be provided based on the primary method of foreign university.

In case, it would be unsuccessful to establish cooperation with foreign university, samples will be order externally.

## 11. CONCLUSION

With the faster developing world economy [24], the questions regarding the energy efficiency and decreasing pollution are getting more and more important. Investment and research new material which has multi or combination of properties getting more popular nowadays [24].

## 12. ACKNOWLEDGEMENT

This work was supported by the Czech foundation Agency (GA 19-17457S).

## REFERENCES

- [1] Scanning Probe Microscopy, Applications [online], CJ Roberts, MC Davies, SJB Tendler, and PM Williams available on: [http:// www.sciencedirect.com](http://www.sciencedirect.com).
- [2] Quantitative Electromechanical Atomic Force Microscopy [online], Liam Collins, Yongtao Liu, Olga S. Ovchinnikova and Roger Proksch available on: [http:// www.sciencedirect.com](http://www.sciencedirect.com).
- [3] Polynomial force approximations and multifrequency atomic force microscopy [online], Daniel Platz, Daniel Forchheimer, Erik A. Tholén and David B. Haviland available on: <https://www.sciencedirect.com/>
- [4] Physica E [online], Aparna Deshpande a, Brian J. LeRoy available on: <https://www.sciencedirect.com>
- [5] Polymer Testing [online], Hua Wu, Yan Qiao available on: [https:// www.sciencedirect.com](https://www.sciencedirect.com)

- [6] Atomic force microscopy for the characterisation of pinning effects of seawater micro-droplets in n-decane on a calcite surface [online], G.C. Savulescu a, M. Rücker, A. Scanziani, R. Pini, A. Georgiadis, P.F. Luckham available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [7] Quantitative atomic force microscopy [online], Hagen Söngen available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [8] AFM Nano eye – Development of an education oriented high resolution profilometer [online], Ioan Alexandru Ivan, Claudie Petit, Ion Valentin Gurgu, Rosario Toscano available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [9] Adhesion study between micron-scale graphite particles and rough walls using the finite element method [online], Qi Sun, Wei Peng, Xiao Hai, Suyuan Yu available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [10] Double-beam modeling and experiments of resonance behaviors of AFM-based nanowires [online], Ji'an Li , Tingjun Wang , Xing Liu , Bo Chen, Qi Xu, Chaoming Wang , Yinghui Li available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [11] Multifunctional AFM/SNOM Cantilever Probes: Fabrication and Measurements [online], M. Stopka, D. Drews, K. Mayr, M. Lacher, W. Ehrfeld, T. Kalkbrenner, M. Graf, V. Sandoghdar, J. Mlynek available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [12] A new positioning approach of fiber probe for near-field scanning optical microscope [online], Shi-Biao Xiang, Tian-Hao Zhang, Xu Xiang, Chun-Ping Zhang available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [13] Raman, AFM and SNOM high resolution imaging of carotene crystals in a model carrot cell system [online], Anna Rygula, Tomasz Oleszkiewicz, Ewa Grzebelus, Marta Z. Pacia, Malgorzata Baranska, Rafal Baranski available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [14] Live endothelial cells imaged by Scanning Near-field Optical Microscopy (SNOM): capabilities and challenges[online], Katarzyna Bulat<sup>1</sup>, Anna Rygula, Ewelina Szafraniec, Yukihiro Ozaki, and Malgorzata Baranska available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [15] Electrochromism: An emerging and promising approach in (bio)sensing technology [online], Mohammad Amin Farahmand Nejad, Saba Ranjbar, Claudio Parolo, Emily P. Nguyen, Ruslan Álvarez-Diduk, Mohammad Reza Hormozi-Nezhad, Arben Merkoçi available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [16] Multicolored and durable electrochromism in water soluble naphthalene and perylene based [online], diimides, Sayan Halder, Susmita Roy, Chanchal Chakraborty available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [17] Preparation and electrochromic properties of tungsten oxide and iridium oxide porous films [online], Fusong Jiang, Tao Zheng, Yong Yang available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [18] Preparation and characterization of spray- deposited efficient Prussian blue electrochromic thin film [online], M.H. Elshorbagy<sup>a,b</sup>, R. Ramadana, K. Abdelhady available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [19] Polymeric complexes of transition metal ions as electrochromic materials: Synthesis and properties [online], Radosław Banasz, Monika Wałęsa-Chorab available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [20] Electrochromic materials and devices for energy efficiency and human comfort in buildings: A critical review [online], C.G. Granqvist , M.A. Arvizu , I. Bayrak Pehlivan, H.-Y. Qu, R.-T. Wen, G.A. Niklasson available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [21] NiCo-mixed hydroxide nanosheets as a new electrochromic material with fast optical response [online], diimides, Sayan Halder, Susmita Roy, Chanchal Chakraborty available on: [http:// https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [22] Constructing Alternated Heterobimetallic [Fe(II)/Os(II)] Supramolecular Polymers with Diverse Solubility for Facile Fabrication of Voltage-Tunable Multicolor Electrochromic Devices [online],

- Manas Kumar Bera, Yoshikazu Ninomiya, and Masayoshi Higuchi available on: [http://  
https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [23] Electrochromic coordination nanosheets: Achievements and future perspective [online], Manas Kumar Bera, Sudhanya Mohanty, Siddhartha Shankar Kashyap, Sanjib Sarmah available on: [http://  
https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [24] A review of NiO-based electrochromic-energy storage bifunctional material and integrated device [online], Junkai Wang, Xiangtao Huo, Min Guo, Mei Zhang available on: [http://  
https://www.sciencedirect.com](http://https://www.sciencedirect.com)
- [25] Electrochromic Material [online], C.Moretti, X.Tao, L.Koehl V.Koncar available on: [http://  
https://www.sciencedirect.com](http://https://www.sciencedirect.com)

# Jsme skupina E.ON, těší nás

Jsme jedním z největších energetických koncernů u nás i ve světě. Centrálu máme v Německu, ale najdete nás už v 15 zemích Evropy, včetně České republiky. Zakládáme si na tom, že naše energie můžou šetřit peníze i přírodu a čím dál víc využíváme obnovitelné zdroje.

## Kdo patří do naší rodiny

### E.ON Energie, a.s.

- Obchoduje s elektřinou a plynem, zajišťuje marketing a komunikaci a stará se i o výrobu energií.
- Pro zákazníky připravuje řešení na míru v oblasti fotovoltaiky, tepelné techniky a elektromobility.

### EG.D, a.s.

- Distributor energií, který vlastní a provozuje rozvodnou síť elektřiny zejména na jihu Čech a Moravy a rozvodnou síť plynu na jihu Čech.
- Zajišťuje připojení odběrných míst k síti a stará se o dopravu energií k zákazníkům.

### E.ON Česká republika, s. r. o.

- Funguje jako podpora výše uvedeným společnostem. Zajišťuje jim služby, jako je účetnictví, právo nebo HR, a na starost má i zákaznickou péči.

## Na přírodě nám záleží



E.ON Energie je jedničkou ve výkupu zelené elektřiny v České republice.



EG.D pomocí bezpečnostních prvků na sloupech vysokého napětí chrání ptáky před úrazem.



V budovách E.ONu i ve všech dobíječkách elektromobilů využíváme zelenou energii.

# Nastartuj svou kariéru

#spolujsmeeon



## Studentské programy

- Letní energetická akademie – 25. – 29. 7. 2022
- Letní technické brigády – od června do září
- Stipendijní program
- Diplomové/bakalářské práce



## Absolventské programy

- Trainee program – od září
- Junior technik – od září

[www.kariera.eon.cz](http://www.kariera.eon.cz)

 E.ON Kariéra

 E.ON Kariéra

**eg.d e.on**



# Identification of horizontal genes transfer elements across strains inhabiting the same niche using pan-genome analysis

J. Schwarzerová<sup>1,2</sup> and D. Čejková<sup>1</sup>

<sup>1</sup> Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication (FEEC), Brno University of Technology (BUT), Brno, Czech

<sup>2</sup> Molecular Systems Biology (MOSYS), University of Vienna, Vienna, Austria

E-mail: [Jana.Schwarzerova@vut.cz](mailto:Jana.Schwarzerova@vut.cz), [cejkovad@vut.cz](mailto:cejkovad@vut.cz)

**Abstract**—Tracing horizontal gene flux across strains in farm animals is one of the important steps for research focused on detection and genomic enzymology of genes conferring antibiotic resistance. In this study, we have built the comprehensive computational methodology for the detection of horizontal genes transfer elements via pan-genome analysis. In total, 133 anaerobes isolated from chicken gastrointestinal tract were examined for the presence of traits of horizontal transfer. The shared genes from all isolates, so called core genome genes were identified and characterised in order to assign the function to the gene within individual bacterial cell and within community of cells. This study provides an evidence that horizontal transmission frequently occurs not only between closely related bacteria, but also between distant taxonomical groups. Hence chickens are known primary reservoirs of antibiotic resistance genes, and the dissemination of these genes to other bacterial pathogens often leads to life-threatening infections, even within human population. Thus, the research on this subject, and the associated results are of a great importance for public health.

**Keywords** — Comparative Genomics, Core Genome, Antibiotics Resistance, Chicken Microbiome

## 1. INTRODUCTION

Horizontal gene transfer (HGT) plays major role in bacterial evolution. HGT is also known as ‘the non-genealogical transmission of genetic material from one organism to another’ [1]. HGT is an important driving force that modulate bacterial genomes, and thus plays a pivotal role in evolution of prokaryotes. HGT is often mediated via mobile genetic elements (MGE) which are able to acquire and harbor foreign genetic material, including antibiotic resistance genes. Bacteria carrying such cargo have evolutionary advantage if particular antibiotics is present in the niche, they overgrow other bacteria. Later, the MGE can be spread vertically and horizontally and additional genes can be acquired by the same bacteria. Bacteria harboring many antibiotic resistance genes are called multidrug resistance bacteria (MRB). MRB, often human pathogen associated with nosocomial infections, can be resistant to many known antibiotics; e.g. methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant *Enterococcus faecalis* represent one of major threats for public health [2]. Therefore, it is very important to trace such genes as well as their vehicles, MGE, in the bacterial community.

The study used whole-genome sequencing data to trace horizontal gene-flux across strains and provide insight into species evolution. One of the most used methods in multi-genome studies is pan-genome analysis. Pan-genome is the term that was inferred in the study by Tettelin et al. [3]. The authors discovered that different strains of *Streptococcus sp.* might differ substantially in their gene content and total gene pool of a species might be orders of magnitudes larger than the gene content of any single strain. Pan-genome analysis provides information on genomic diversity of the investigated bacteria, determining core (conserved, that can characterize the biological function in large microbial clades), accessory (dispensable genes in different species) and unique (strain-specific) gene pool of a species [4]. Our study aims to determine core and accessory sequences from 133 gastrointestinal tract (gut) anaerobes isolated from chicken caecum in pure cultures and to detect gene elements playing role as mediator in HGT.



## 2. MATERIALS

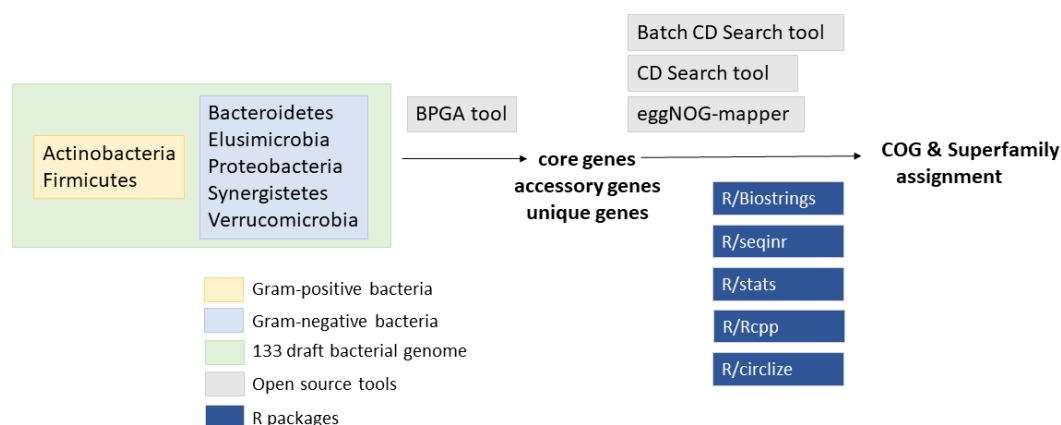
Dataset originated from the study by Medvedcky et al. [5]. In the study, 204 novel bacterial isolates from chicken caecum were isolated via cultivation on the Wilkins-Chalgren anaerobe agar under anaerobic growth conditions. The genomes were sequenced using Illumina NextSeq 500 platform [6]. Raw sequencing reads were quality trimmed using Trimmomatic [7] and assembled by IDBA-UD [8]. Our follow-up study used the annotated genomes data that are deposited in NCBI under accession number PRJNA377666 [5]. In total, it is 133 draft genomes from gut anaerobes isolated from chicken caecum in pure cultures.

## 3. METHODS

Methodology carried out in this study is based on pan-genome analysis. The whole analysis is performed on seven different strains such as Actinobacteria, Bacteroidetes, Elusimicrobia, Firmicutes, Proteobacteria, Synergistetes and Verrucomicrobia. Firstly, we divided these strains into two different groups. The first group includes Gram-positive, the second group includes Gram-negative bacteria group. Pan-genome analysis was performed for both groups. Throughout the study, gene and protein sequences were analyzed. However, the term core genes are widely used, so we will use names for core genes and core protein sequences interchangeably.

The pan-genome analysis was performed by Bacterial Pan Genome Analysis (BPGA) tool [9]. Besides defining the core, accessory and unique genome gene pools, BPGA also enables additional features for downstream analyses, core/pan/MLST (Multi Locus Sequence Typing) phylogeny, exclusive presence/absence of genes in specific strains, subset analysis; atypical G + C content analysis of core, accessory and unique genes, to name a few [9]. Thus, the BPGA tool shown new insight into the analysis of chicken gut genomes.

The identified core genome genes from each group were functionally annotated, i.e. superfamily and Clusters of Orthologous Groups (COG) were assigned to every gene to predict their function within bacterial cell and/or community. Firstly, the core genome genes were analysed using Batch CD Search tool [10]. The superfamilies were visualized using Venn diagram. The overlapping region containing shared superfamilies of both groups was analyzed in more detail using CDD Search tool [11]. The COG assignment was performed using eggNOG-mapper tool [12]. The results were analysed and visualized using R, expressly, R/ Biostrings [13], R/ seqinr [14] and R/ stats packages [15] were applied. The whole pipeline of methodology is shown in Figure 1.



**Figure 1:** The methodology used in the study; in yellow are marked Gram-positive bacteria which represents Actinobacteria and Firmicutes; in blue are marked Gram-negative bacteria represented by Bacteroidetes, Elusimicrobia, Proteobacteria, Synergistetes and Verrucomicrobia phyla.

## 4. RESULTS & DISCUSSION

The original dataset was divided into two groups. Gram-positive bacterial represents first group and include 99 draft genomes. Gram-negative bacteria represented the second group and include 34 draft genomes. In total, 133 draft genomes were analyzed.



## Pan-genome analysis

In total, 365,298 genes were identified, 259,514 genes were identified in Gram-positive and 105,784 were identified in Gram-negative bacteria. The highest number of core genes were in group of Gram-positive bacteria. It included 2,475 core genes, see Table I. The number of core genes from Gram-negative bacteria was significantly lower including 340 core genes. The low number of core genes in Gram-negative bacteria might be caused because the group was consisted of taxonomically distant bacterial genomes which were examined, so the members of this group likely do not share many genes. Therefore, they may bear only hundred gene in common. Thus, there is very different distribution based on standard deviation from average GC content, on contrary to conventional pan-genome analysis.

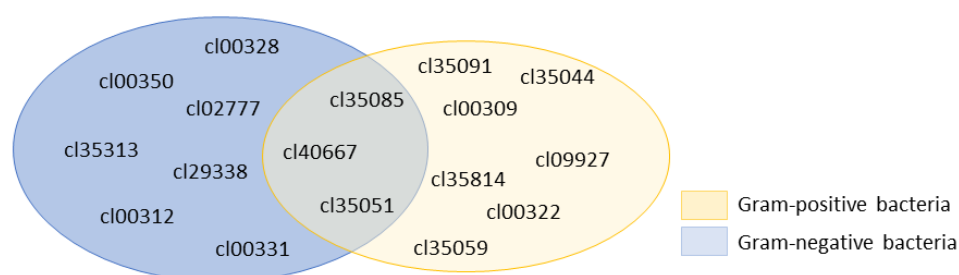
In addition, pan-genome analysis also identified accessory and unique genes. However, also accessory genes can be mediated by and/or associated with HGT. The accessory genes will be further analysed in follow-up studies. Summary of pan-genomes analysis is shown in Table I.

**Table I:** Number of sequences from pan-genome analysis using BPGA tool. The default setting for pan-genome analysis was used:  $2 \times$  standard deviation (SD) from average GC content.

Group of genomes	Core genes	Accessory genes	Unique genes
Gram positive	2,475	217,317	39,722
Gram negative	340	79,403	26,041

## Determination of horizontally transferable gene elements

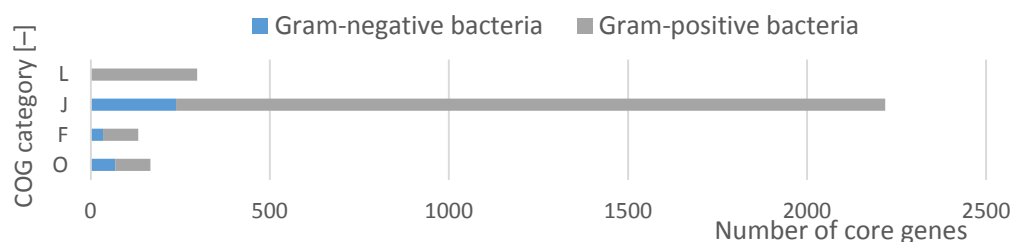
The core genes were used as input in Batch CD Search tool. In total, 10 superfamilies were assigned to 302 core genes in Gram-negative bacteria, as well as 10 superfamilies were assigned to 900 core genes in Gram-positive bacteria. This intersection represents 3 superfamilies, especially cl3508, cl40667 and cl35051, see Figure 2. Superfamily cl35085 represents molecular chaperone DnaK. cl40667 represents ribosomal protein L2 and cl35051 is an elongation factor EF-Tu. In total, 17 superfamilies were identified in all analysed draft genomes. The rest 613 genes belonged to gene of unknown function.



**Figure 2:** Venn diagram of identified superfamilies identified via Batch CD Search tool.

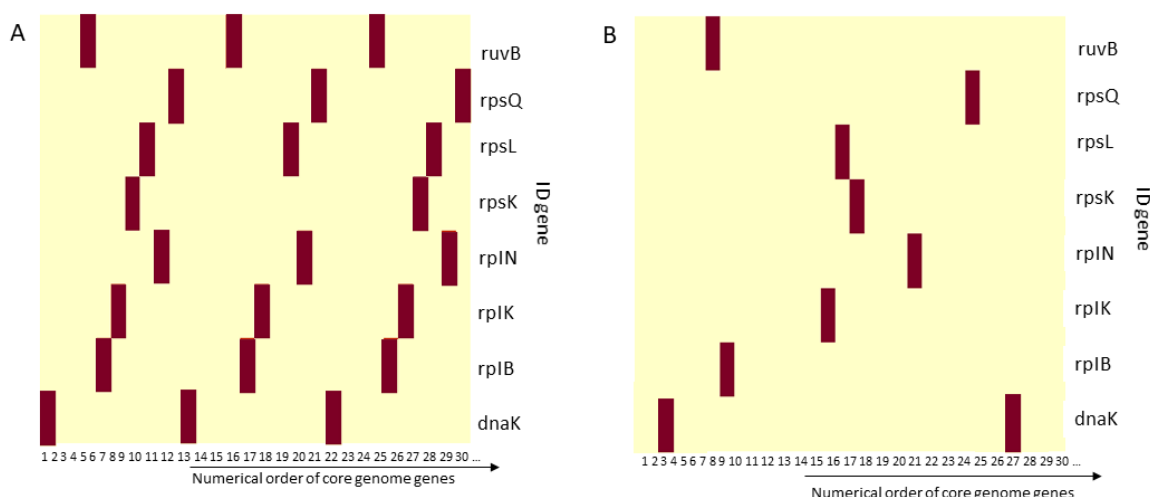
In more detail, we focused on the intersection set. The next analysis was performed by standard CD search tool using Blastp [16]. The verification of the associated superfamily and assigned gene function was performed using visualization from CD search tool.

In addition, the COG assignment was performed for core gene sequences using eggNOG-mapper tool. On summary, we identified 25 diverse clusters in Gram-positive bacteria and 10 diverse clusters in Gram-negative bacteria. In total, we determined 4 different COG categories, see Figure 3. The most common category is J, group of genes involved in translation, ribosomal structure and biogenesis function.



**Figure 3:** COG distribution using eggNOG-mapper. L: Replication and repair category, J: Translation, ribosomal structure and biogenesis, F: Nucleotide metabolism and transport O: Post-translational modification, protein turnover, and chaperone.

In the last, the association between individual core genes and COGs were visualised by heatmap. For clarity, illustration of the analysis containing first thirty core gene sequences are depicted in Figure 4. Considerable amount of determined elements was connected to horizontal gene transfer analysis. For example, *tuf* gene is identified as gene with possible HGT [17], or *dnaK* gene is also often mentioned in literature about HGT [18][19]. Moreover, *dnaK* gene is connected to research focused to multidrug resistance [20].



**Figure 4:** Heatmaps represent examples of identified genes obtained by eggNOG-mapper in chicken microbiome and their prevalence across core sequences. (A) represents example for Gram-negative bacteria; (B) represents Gram-positive bacteria.

## 5. CONCLUSION

Advantages of high-throughput low-cost sequencing technologies and metagenomic techniques have shift the research interest from single or few genome analyses to large-scale comparison studies. It allows us to pave new avenues for novel research challenges, such as the detection of horizontal gene transfer in diverse bacteria. One of such challenge comprised our study focusing on identification of transferable genes and elements in bacterial genomes, not between closely related but also between distant bacterial members inhabiting the same environment.

Our study used pan-genome analyses to determine core-genome of chicken microbiome for Gram-negative and Gram-positive bacteria. These core-genomes genes were used in subsequent analysis in which functional COG and superfamily categories were assignment to each gene. Especially, functional COG and superfamily categories were assignment to each gene. We identify 10 superfamilies in Gram-positive and Gram-negative bacteria using CD Batch tool. Nevertheless, we also identify 10 ID genes in Gram-negative bacteria and 25 ID genes in Gram-positive bacteria using eggNOG-mapper that is spaced across both of core genome genes. These identified genes include genes which connects to HGT, which is a promising way to open further studies.

## ACKNOWLEDGMENT

This work has been supported by grant project GACR 22-16786S.

## REFERENCES

- [1] BOTO, Luis. Horizontal gene transfer in evolution: facts and challenges. *Proceedings of the Royal Society B: Biological Sciences*, 2010, 277.1683: 819-827.
- [2] ARGEMI, Xavier, et al. Comparative genomic analysis of *Staphylococcus lugdunensis* shows a closed pan-genome and multiple barriers to horizontal gene transfer. *BMC genomics*, 2018, 19.1: 1-16.
- [3] TETTELIN, Hervé, et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial “pan-genome”. *Proceedings of the National Academy of Sciences*, 2005, 102.39: 13950-13955.
- [4] CHAUDHARI, Narendrakumar, et al. BPGA-an ultra-fast pan-genome analysis pipeline. *Scientific reports*, 2016, 6.1: 1-10.
- [5] MEDVECKY, Matej, et al. Whole genome sequencing and function prediction of 133 gut anaerobes isolated from chicken caecum in pure cultures. *BMC genomics*, 2018, 19.1: 1-15.
- [6] PAIJMANS, Johanna LA, et al. Sequencing single-stranded libraries on the Illumina NextSeq 500 platform. *arXiv preprint arXiv:1711.11004*, 2017.
- [7] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20.
- [8] PENG, Yu, et al. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*, 2012, 28.11: 1420-1428.
- [9] CHAUDHARI, Narendrakumar M.; GUPTA, Vinod Kumar; DUTTA, Chitra. BPGA-an ultra-fast pan-genome analysis pipeline. *Scientific reports*, 2016, 6.1: 1-10.
- [10] MARCHLER-BAUER, Aron, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic acids research*, 2010, 39.suppl\_1: D225-D229.
- [11] MARCHLER-BAUER, Aron; BRYANT, Stephen H. CD-Search: protein domain annotations on the fly. *Nucleic acids research*, 2004, 32.suppl\_2: W327-W331.
- [12] CANTALAPIEDRA, Carlos P., et al. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Molecular biology and evolution*, 2021, 38.12: 5825-5829.
- [13] PAGES, Hervé, et al. Biostings: String objects representing biological sequences, and matching algorithms. *R package version*, 2016, 2.0: 10.18129.
- [14] CHARIF, Delphine, et al. Package ‘seqinr’. 2021.
- [15] *Documentation for package ‘stats’ version 4.2.0: The R Stats Package* [online]. Available from: <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/00Index.html>
- [16] WHEELER, David L., et al. Database resources of the national center for biotechnology information. *Nucleic acids research*, 2007, 36.suppl\_1: D13-D21.
- [17] KACAR, Betül, et al. Functional constraints on replacing an essential gene with its ancient and modern homologs. *MBio*, 2017, 8.4: e01276-17.
- [18] PETITJEAN, Céline, et al. Horizontal gene transfer of a chloroplast DnaJ-Fer protein to Thaumarchaeota and the evolutionary history of the DnaK chaperone system in Archaea. *BMC evolutionary biology*, 2012, 12.1: 1-14.
- [19] HAGHI, Morteza, et al. Detection of heat shock protein (DnaK, DnaJ and GrpE) horizontal gene transfers among *Acanthamoeba polyphaga*, *Acanthamoeba polyphaga mimivirus* (APMV), amoeba-infecting bacteria and sputnik virophage. *Int J Adv Biotechnol Res*, 2016, 7: 1618-1622.
- [20] MUJAWAR, Shama, et al. Variant analysis from bacterial isolates affirms DnaK crucial for multidrug resistance. In: *International Work-Conference on Bioinformatics and Biomedical Engineering*. Springer, Cham, 2020. p. 237-248.

# Calculating the Cortical Thickness of Brain by using electrical and mathematical models for Optimal Transcranial Direct Current Stimulation.

Ali Hamza<sup>1</sup>, Martin Mezl<sup>2</sup>, and Dr. Saima Zafar<sup>3</sup>

<sup>1</sup>doctoral student, Brno University of Technology, Czech Republic

<sup>2</sup>Assistant Professor, Brno University of Technology, Czech Republic

<sup>3</sup>Professor NUCES, Pakistan

E-mail: [ali.hamza@vut.cz](mailto:ali.hamza@vut.cz), [mezl@vut.cz](mailto:mezl@vut.cz), [saima.zafar@nu.edu.pk](mailto:saima.zafar@nu.edu.pk)

**Abstract**—transcranial direct current stimulation (tDCS) is one of the most well-known brain stimulation techniques. In this technique, a constant low intensity (4mA to 5mA) current provides to the brain to modulate the excitability of neurons. There are two categories of tDCS, and both are used for diagnosing different neural disorders, one is anodal stimulation and the other one is cathodal stimulation. Anodal stimulation is used for patients of stroke, epilepsy and depression. Whereas cathodal stimulation is helpful for the obesity, drug addiction and hypertension. The aim of this research was to improve the efficiency of tDCS which depends on the intensity and magnitude of current applied for stimulation. It is important to note that there are numerous variables that can alter the intensity of the current. Some of these variables are hair thickness, head size, skull, tissues thickness, cortical fluid thickness, age, cortical thickness of the brain. In this article, the cortical thickness of the brain is estimated, and the current intensity is calculated for a patient. The Magnetic resonance images (MRI) data of 8 male subjects are analyzed, the raw MRIs are preprocessed, and layers of the head are segmented using MATLAB tool statistical parametric mapping (SPM). After the segmentation of these layers, we calculated the combined thickness of these layers to calculate the desired value of current. The authenticity of our results is compared with the standard current values. Our results shows that the calculated current intensity is closely matched with the standard values and is a direct function of the cortical thickness of brain which is computed from the MRIs. We claim that using our computation, the stimulation current is more accurate and is patient dependent instead of applying the same current to every patient.

**Keywords**—tDCS, Cortical thickness, Cortical fluid thickness

## 1. INTRODUCTION

The advancement in the study of the human central nervous system has been a great interest in scientific experimentation. In the past decade experimentation on tDCS became popular for modulating the excitability of neurons and motor areas. Almost 1500 research articles published between the years 2011 to 2021s, whereas from the period 2000 to 2005 the number was only 65 to 70. These number suggests that there is a lot of research taking place in this field. It gives us a booster that there is an ample space for the researchers to do research in this field. This work focuses on the efficiency enhancement of the stimulation techniques especially tDCS, but for a better understanding we must know the basics of how stimulations work. After that, we can be able to take a deep study on the stimulation techniques [8].

There are many stimulation techniques that neurologists used for the diagnosing of neural disorders. The crux of these stimulation techniques will discuss in the following section. In the next section, we will discuss the brain structure that is directly related to the efficiency of stimulation techniques.

tDCS was an innovative technique for modulating neuroplasticity in clinical and cognitive neuroscience early in its growth. One of the frequently used techniques provides tDCS with an energy of 1–2 mA (0.029–0.057 mA / cm<sup>2</sup>) through plate electrodes positioned on the scalp with a fluid flowing from the anodal to the cathodal panel [19]. Recent trials that implement tDCS in a multitude of study and clinical environments have demonstrated their capacity to enhance or interrupt cognitive functions and improve a variety of neurological and psychiatric impairments. In humans, anodal cortical activity lasts 5 to 30 minutes, causing enhanced excitability for hours after the stimulus. Our focus will be on increasing the

efficiency of stimulation techniques in order to diagnose various neural disorders. Before getting into details that how we can enhance the efficiency of stimulation. We must first understand about the stimulation techniques.

tDCS is a type of neuro-stimulation that utilizes the electrode on the body with the steady small immediate flow. It is designed to help mental disorders like depression, stroke, and epilepsy. tDCS is also useful for healthy people's cognitive enhancement. Several assessments have shown cognitive improvements [18]. Moreover, it can be helpful in treating neuropathic pain after spinal cord injury, schizophrenia, Alzheimer's disease. One element of tDCS is its capacity even after the stimulation has been finished to create cortical modifications. The length of the change depends on the length of the stimulation and intensity of the stimulation. Stimulation is either depolarizing the neuron's dormant membrane or hyperpolarizing it. The brain membrane is hyperpolarized when anodal tDCS is supplied which improves the neuronal unexcitability and enables the rapid cell shooting. Whereas in cathodal stimulation we deexcites the neural activities and improve diseases like hypertension and drug addiction.

## 2. METHODOLOGY

In neurology models, most of the time the data we get is in the form of magnetic resonance images (MRIs). We cannot use these MRIs for our analysis without applying necessary image processing techniques, because there is noise and other artefacts present in these images [2]. So, to remove the noise and artefacts we will undergo a data pre-processing steps.

- Data acquisition and reconstruction
- Distortion correction
- Slice timing correction
- Motion correction and normalization
- Temporal filtering
- Spatial smoothing

Gaussian distribution is used to remove the noise from the images.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

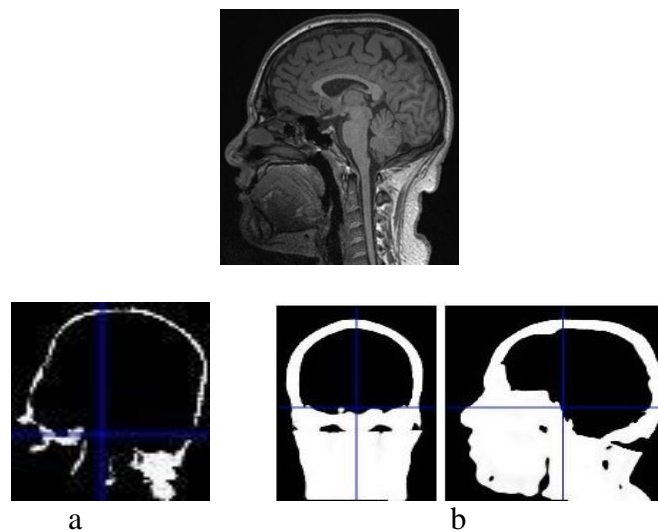
where  $x, \mu$  and  $\sigma$  are the intensity, the mean value, the standard deviation, and the mean respectively. The procedure for removing artefacts is to convolve the MRI signal with a Gaussian beam of having a specific width. This beam is called the Gaussian kernel, the shape of this kernel is like a bell curve. The smoothness of data depends on the width of the bell curve. When we analyzing the single subject our kernel size or width is close to 4mm but when we doing an analysis of multiple subjects this width should be in between 6mm to 10mm depends on inter-subject variability.

## 3. SEGMENTATION

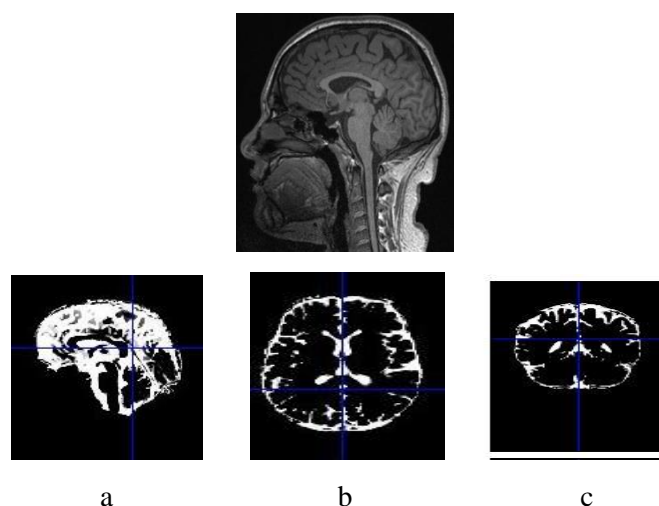
After doing all preprocessing steps the most important and final step of our methodology is segmentation. The segmentation of MRIs is most important and critical step in the analysis of clinical application. Image segmentation is generally used for measuring the brain's structure. Why segmentation is an essential task in study of brain imaging because it can affect the whole analysis [2]. In our work we do the region-based segmentation with the help of MATLAB tool SPM 12 using voxel-based morphometry (VBM). The images for each subject are categorized into a number of different tissue types. The tissue types are defined according to tissue probability maps, which define the prior probability of finding a tissue type at a particular location. Typically, the order of tissues is grey matter, white matter, CSF, bone and soft tissues

The basic aim of segmentation is to split an image into its different regions and these regions will not overlap and all the images are independent of each other in respect of information, color, region and intensity [2]. This will be helpful in finding the thickness of each layer separately and after that these

thickness values will be used in compiling the results. The process will completely be understandable with the help of figure1 and figure 2.



**Figure 1:** Original image is segmented into skull region (a) and whole structure of head other than brain (b).



**Figure 2:** Original image is segmented into three parts of brain white matter (a) cerebrum (b) and grey matter (c)

#### 4. RESULTS

In this study we always take our measurements in an average form because it is impossible to measure the exact thickness of any layer of human head because it is highly unfolded in nature so in order to measure the thickness of different layers of brain [14], we segmented the MRIs after the pre-processing into its different layers and after that we measure the thickness of each layer. Finally, we combine the total thickness of these individual layers and make them usable for our calculations. we measure the thickness of cerebrum and total thickness of cerebrum and cranium bones of 24 to 50 years male subject. We are basically interested in finding the total thickness of cranium (skull bone) and cerebrum because this thickness plays a major role in finding the current value for the stimulation technique. It is important to note that the value of current that we will calculate is depends on the total thickness of head [5]. For this purpose, we take 36 to 44 slices of MRIs and measure the length of each slice. Using mathematical modeling of linear approximation [1].



### Calculation of Current

The standard values of resistivity for each layer of brain are as follows.

Resistivity of brain's white matter = 700 ohm-cm

Resistivity of brain's Gray matter = 330 ohm-cm

Resistivity of spinal cord and cerebellum = 628 ohm-cm

Resistivity of cerebrospinal fluid = 65 ohm-cm

Resistivity of scalp = 250 ohm-cm

Resistivity of skull = 4000 ohm-cm

Total resistivity of all layers =  $\rho = (700+330+650+64+260+4000)$  ohm-cm = 6004 ohm-cm

Resistivity in meters = 60.04 ohm-m

As we know that

$$R = \rho \frac{L}{A}$$

R = Resistance = ?

$\rho$  = total resistivity = 60.04 ohm-m

L1 = Thickness of hairs = 4mm

$$L0 = \frac{1}{N} \sum_{i=1}^N \text{length of slice}_i$$

L = Total thickness = L0 + L1

A = Area of layers =  $A = \pi L^2$

$$R = \rho \frac{L}{A}$$

In tDCS we use battery of 9 volts

According to ohm's law

$$I = \frac{V}{R}$$

Note that all measurements in calculations are in millimeters. The results are shown in table 1.

**Table I:** Results of current values for 8 subjects

Cortical thickness of patient	Total Resistance (R)	Current value for each patient
11.90 mm	1605.99 ohm	5.60 mA
11.48 mm	1664.74 ohm	5.40 mA
11.24 mm	1700.00 ohm	5.29 mA
11.78 mm	1622.62 ohm	5.54 mA
11.72 mm	1630.65 ohm	5.51 mA
11.33 mm	1686.78 ohm	5.33 mA
11.44 mm	1670.57 ohm	5.38 mA
11.36 mm	1682.33 ohm	5.34 mA

## 5. CONCLUSION

The values of current that we measure are ranging from 5.29 mA to 5.6 mA, which are very close to recent stimulation experimentations. The maximum current nowadays neurologist given to their patients is 5 mA [8]. So, the latest literature proves the authenticity of our results. One thing is here to note that we analyze the data of male patients because we do not have any female patient in our data set. The value of current must be high in case of female patient because of the density of hairs and their skull is comparatively thicker than males.

If we want to make these stimulation techniques more effective, it will be good for the patient's health and time to calculate its thickness so according to that we should give the desire current to its head.

tDCS can be used to temporarily and reversibly modulate cognitive states and actions. tDCS is becoming popular tool for investigating brain behavior relationships [13]. There are number of factors that involves in the enhancement of its efficiency. But the aim of this thesis was to provide a guideline about the efficiency of tDCS as we mentioned in introductory portion that efficiency of stimulation is directly depends on the cortical thickness of brain and its respective current value [8]. For this purpose, we find the thickness of everyone according to this we take decision about the value of current that will provide during the stimulation.

In results section we find the thickness of 8 healthy male subjects and according to their brain structure and orientation we will give them a respective current. When we analyzed our results, we concluded that if the thickness is high then we will give a current with higher magnitude and vice-versa.

## ACKNOWLEDGMENT

Thanks to my supervisor Martin Mezl and Dr Saima Zafar helped me immeasurably throughout the research process whose turn of phrase & strategic guidance brought elegance to my fractured ideas or thoughts. During this work, I developed a tremendous amount of respect and appreciation for both my supervisors.

## REFERENCES

- [1] Despotović, I., Goossens, B., & Philips, W. (2015). MRI segmentation of the human brain: challenges, methods, and applications. *Computational and mathematical methods in medicine*, 2015.
- [2] Lan, T., Erdogmus, D., Hayflick, S. J., & Szumowski, J. U. (2008, October). Phase unwrapping and background correction in MRI. In *2008 IEEE Workshop on Machine Learning for Signal Processing* (pp. 239-243). IEEE.
- [3] Tan, K. C., Kim, T. H., Chun, S. I., Shin, W. J., & Mun, C. W. (2008). A simple, fast, and robust phase unwrapping method to unwrap MR phase images. In *4th Kuala Lumpur International Conference on Biomedical Engineering 2008* (pp. 487-490). Springer, Berlin, Heidelberg.
- [4] Latikka, J., & Eskola, H. (2019). The electrical conductivity of human cerebrospinal fluid in vivo. In *World Congress on Medical Physics and Biomedical Engineering 2018* (pp. 773-776). Springer, Singapore. Väisänen, J., Wendel, K., Seemann, G., Malmivuo, J., & Hyttinen, J. (2009). Sensitivities of bipolar subcutaneous and cortical EEG leads. In *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany* (pp. 267-270). Springer, Berlin, Heidelberg.
- [5] Sadleir, R. J., & Argibay, A. (2007). Modeling skull electrical properties. *Annals of Biomedical engineering*, 35(10), 1699-1712.
- [6] Weller, R. O., Sharp, M. M., Christodoulides, M., Carare, R. O., & Møllgård, K. (2018). The meninges as barriers and facilitators for the movement of fluid, cells and pathogens related to the rodent and human CNS. *Acta neuropathologica*, 135(3), 363-385.
- [7] Asan, A. S., Gok, S., & Sahin, M. (2019). Electrical fields induced inside the rat brain with skin, skull, and dural placements of the current injection electrode. *PloS one*, 14(1), e0203727.

- [8] Liu, A., Vöröslakos, M., Kronberg, G., Henin, S., Krause, M. R., Huang, Y., ... & Berényi, A. (2018). Immediate neurophysiological effects of transcranial electrical stimulation. *Nature communications*, 9(1), 1-12.
- [9] Krause, M. R., Vieira, P. G., Csorba, B. A., Pilly, P. K., & Pack, C. C. (2019). Transcranial alternating current stimulation entrains single-neuron activity in the primate brain. *Proceedings of the National Academy of Sciences*, 116(12), 5747-5755.
- [10] Yazdan-Shahmorad, A., Kipke, D. R., & Lehmkuhle, M. J. (2011). Polarity of cortical electrical stimulation differentially affects neuronal activity of deep and superficial layers of rat motor cortex. *Brain stimulation*, 4(4), 228-241.
- [11] Conde, V., Vollmann, H., Sehm, B., Taubert, M., Villringer, A., & Ragert, P. (2012). Cortical thickness in primary sensorimotor cortex influences the effectiveness of paired associative stimulation. *Neuroimage*, 60(2), 864-870.
- [12] Schnack, H. G., Van Haren, N. E., Brouwer, R. M., Evans, A., Durston, S., Boomsma, D. I. & Hulshoff Pol, H. E. (2014). Changes in thickness and surface area of the human cortex and their relationship with intelligence. *Cerebral cortex*, 25(6), 1608-1617.
- [13] Zhuang, Y., Zeng, X., Wang, B., Huang, M., Gong, H., & Zhou, F. (2017). Cortical surface thickness in the middle-aged brain with white matter hyperintense lesions. *Frontiers in aging neuroscience*, 9, 225.
- [14] Seiger, R., Ganger, S., Kranz, G. S., Hahn, A., & Lanzenberger, R. (2018). Cortical Thickness Estimations of FreeSurfer and the CAT12 Toolbox in Patients with Alzheimer's Disease and Healthy Controls. *Journal of Neuroimaging*, 28(5), 515-523.
- [15] Shafi, M. M., Westover, M. B., Fox, M. D., & Pascual- Leone, A. (2012). Exploration and modulation of brain network interactions with noninvasive brain stimulation in combination with neuroimaging. *European Journal of Neuroscience*, 35(6), 805-825.
- [16] Selimbeyoglu, A., & Parvizi, J. (2010). Electrical stimulation of the human brain: perceptual and behavioral phenomena reported in the old and new literature. *Frontiers in human neuroscience*, 4, 46.
- [17] Miyaguchi, S., Onishi, H., Kojima, S., Sugawara, K., Tsubaki, A., Kirimoto, H., Tamaki, H. and Yamamoto, N., 2013. Corticomotor excitability induced by anodal transcranial direct current stimulation with and without non-exhaustive movement. *Brain research*, 1529, pp.83-91.
- [18] Power, H.A., Norton, J.A., Porter, C.L., Doyle, Z., Hui, I. and Chan, K.M., 2006. Transcranial direct current stimulation of the primary motor cortex affects cortical drive to human musculature as assessed by intermuscular coherence. *The Journal of physiology*, 577(3), pp.795-803.
- [19] Bastani, A. and Jaberzadeh, S., 2012. Does anodal transcranial direct current stimulation enhance excitability of the motor cortex and motor function in healthy individuals and subjects with stroke: a systematic review and meta-analysis. *Clinical neurophysiology*, 123(4), pp.644-657.

# Anisotropic Forward Modeling of Electroencephalogram in MATLAB - CST Studio Suite Interface

D. Kuřátko<sup>1</sup>, and Z. Raida<sup>1</sup>

<sup>1</sup>Brno University of Technology, Czech Republic

E-mail: [xkurat01@vutbr.cz](mailto:xkurat01@vutbr.cz), [raida@vut.cz](mailto:raida@vut.cz)

**Abstract**—A forward model of an electroencephalogram (EEG) is a necessary part of an EEG source analysis and plays a crucial role in accuracy of brain waves localization. The paper is aimed to exploit CST Studio Suite (CST) for an anisotropic forward model computation. Since the CST computes electric field intensity only, the line integration is used for the recalculation to the electric potential in MATLAB. The integration method and the accuracy of CST are validated by two software tools, and the comparison shows consistent electric potentials. Finally, the simplified anisotropic rat brain models are developed and simulated. Nevertheless, the anisotropic models compared with isotropic one show a rather weak effect on computed electric potentials.

**Keywords**—Forward model, rat brain, CST Studio Suite, anisotropic tissue

## 1. INTRODUCTION

In the brain research, an active brain area is necessary to be reconstructed [1]. Such active brain areas can be localized by an EEG source analysis that solves two mathematical problems [1]. The first one, the forward model [1], computes electric potentials on a brain surface; potentials are generated by dipole sources defined by known parameters. The second one, the inverse problem [1], exploits the forward model, and focuses on the reconstruction of dipolar sources that can explain the measured EEG.

In this paper, we want to exploit CST Studio Suite (Dassault Systems, Illinois, USA) controlled by MATLAB for forward modeling that allows an inclusion of anisotropic tissues to the simulations. In the past, the volume conductor model was assumed to be isotropic. Nevertheless, the anisotropic behavior of tissues was shown in [1], [2]. Generally, the white matter [1], [2] is considered as the main anisotropic compartments of the head. The anisotropy of the white matter is caused by the directionality of the axons that are mostly oriented along the brain surface [1]. In [2], authors reported scares effect of the white matter to the EEG source analysis. In [3], authors observed non-improving human EEG source analysis, when the anisotropic white matter was considered. On the other hand, the significant effect of an anisotropic white matter on the EEG source analysis was reported in [4]. Recently, [5] showed a higher accuracy of brain waves localization for deep sources and lower value of signal to noise ratio when anisotropic tissues were involved. Further, [6] concluded that anisotropic tissues should be considered when the high accuracy is required.

Since the influence of anisotropic tissues to forward model is still unclear, we decided to exploit CST for anisotropic simulations.

## 2. SETTING OF FORWARD MODEL

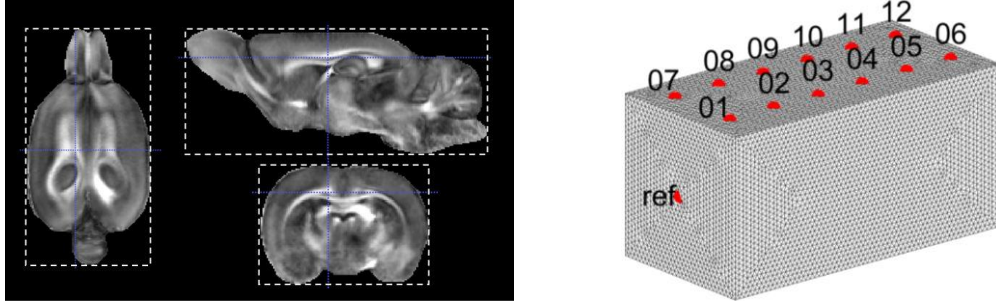
For the forward model computations, we need [1]:

- The volume conductor model defined by the conductivity.
- The dipolar source with a given location, orientation, and a dipole moment.
- Electrodes for sensing electric potential originating from dipolar sources.

The simple volume conductor model is described by a sphere in case of a human forward model. Since the presented paper is focused on a rat and the rat brain cannot be approximated by the sphere sufficiently. Hence, the cuboid shape was selected as a simplified 3D model of the rat brain. The

procedure of a model estimation is demonstrated in Fig. 1 (a) on diffuse tensor images (DTI). The final meshed model with dimensions 30x15x15 mm is shown in Fig. 1 (b). The conductivity of the volume conductor model was set to 0.33 S/m that is a conventional value used in forward models [1].

The output of the forward model corresponds to the electric potential that is measured on the brain surface by electrodes [1]. Here, we used twelve sensing point electrodes; their positions are shown in Fig. 1 (b). Note that all electric potentials are measured with respect to the reference electrode placed on the side wall of the 3D model.



**Figure 1:** The simplified rat brain model: initial schematic for model estimation (a), final meshed rat brain model (b).

The brain wave sources are generally modeled as dipoles [1]. Here, we used one dipole in the origin of the coordinate system that corresponds to the center of the model. Moreover, the dipole was excited by the dipole moment  $1 \text{ nA} \cdot \text{m}^{-1}$  and was oriented perpendicularly to the brain surface.

### 3. MATLAB-CST INTERFACE

Nowadays, forward models are based on the quasi-static approximation of Maxwell equations [1]. Nevertheless, the CST offers solver allowing full-wave simulations:

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad (1)$$

$$\nabla \times \left( \nabla \times \frac{\mathbf{A}}{\mu} \right) + j\omega\sigma\mathbf{A} - \omega^2\varepsilon\mathbf{A} = \mathbf{J}. \quad (2)$$

Here,  $\mathbf{A}$  represents the magnetic vector potential,  $\mathbf{B}$  corresponds to the magnetic induction vector,  $\omega$  denotes angular frequency ( $\omega=2\pi f$ ),  $\sigma$  is conductivity,  $\varepsilon$  denotes permittivity,  $\mathbf{J}$  represents the primary current, and  $\mu$  is permeability of vacuum since relative permeability of biological tissues equals to one.

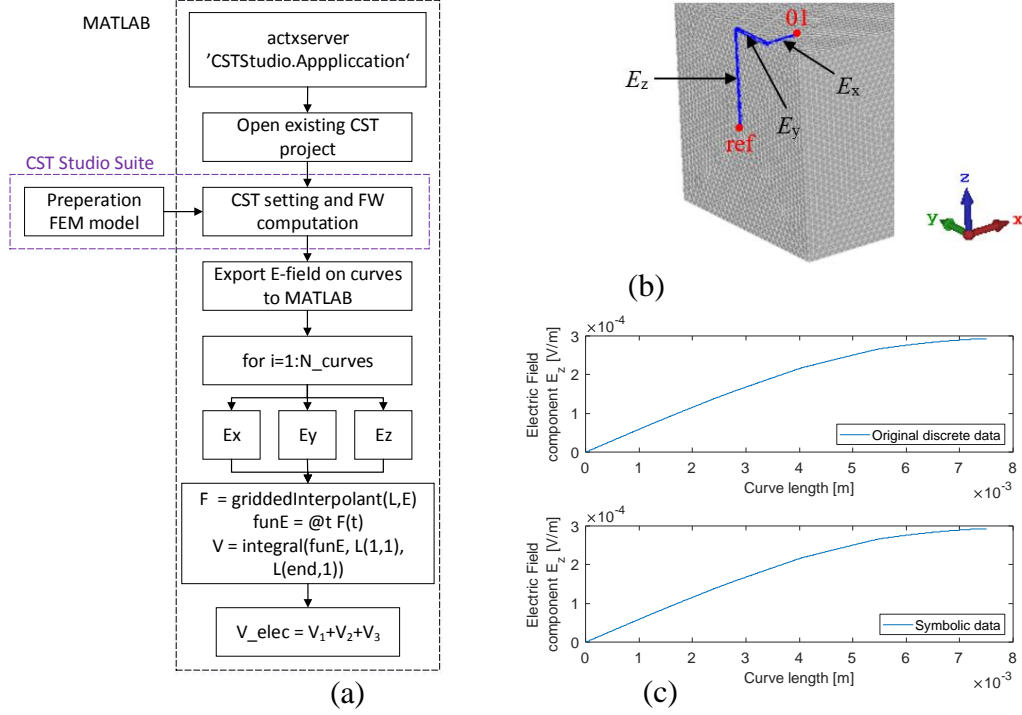
The CST can be controlled from MATLAB by the ActiveX framework, that simplifies data handling. The general flowchart of controlling CST by MATLAB is depicted in Fig. 2 (a). The interface is based on an existing CST project, where the forward model and simulation setup is set. MATLAB mediates communication with CST for data post-processing. Since the CST computes the electric field intensity and forward models are represented in electric potentials, line integral is necessary to be evaluated. Generally, the electric potential is measured with respect to the reference electrode. Thus, the integration is performed on lines that connect the reference electrode with the sensing one.

The process is based on 1D lines. Thus, the integration through the 3D space has to be defined by three lines; one line can be changed in one direction only. After that, the corresponding component of electric field intensity can be measured on an appropriate line. As the result of electric field integration, the electric potential is obtained. Since the integration is defined by three lines, the potential on the electrode is given by the sum of three electric potentials. Finally, an example of an electric potential computation on the electrode 01 is illustrated in Fig. 2 (b).

The presented forward model is defined by twelve sensing electrodes. Since one electrode requires three lines, the integration has to be based on 36 lines in total. Thanks to the MATLAB-CST interface, the data can be exported to MATLAB by invoke command, and the integration can be performed. Since the MATLAB command integral requires symbolic data and CST can export only discrete

points, the data conversion is necessary. Such conversion can be accomplished in MATLAB by the function `griddedInterpolant`.

From the comparison of the original discrete data with the symbolic one for the electric field component  $E_z$  (see Fig. 2c)), the correct data conversion is obvious. Note that the whole process runs in the **for** cycle (Fig. 2a). In addition, the data has to be sorted depending on the electric field component.



**Figure 2:** Procedure for forward model computation: The MATLAB-CST interface (a), integration lines for electric potential computation on electrode 01 (b), example of data conversion for electric field component  $E_z$  (c).

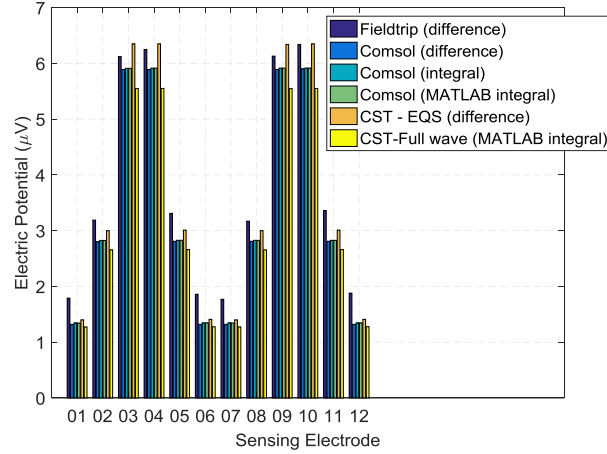
To validate results from CST and integration in MATLAB, we performed simulation in COMSOL Multiphysics (COMSOL Inc., Burlington, USA) and Fieldtrip [7]. The cooperation of SimBio toolbox [7] with FieldTrip is necessary to perform FEM simulations. Unlike the full-wave computation in CST, the FieldTrip and COMSOL solve Poisson's equation:

$$\nabla \cdot (\sigma \nabla U) = \mathbf{J}. \quad (3)$$

Here,  $U$  denotes electric potential,  $\mathbf{J}$  represents the primary current and  $\sigma$  is conductivity. Obviously (eq. 3), the resulting quantity is electric potential. Therefore, the potential on sensing electrodes is given by the potential difference of two points. Since the integration is feasible in COMSOL, we can construct the same integration lines as in CST and integrate electric field intensity components in COMSOL. Finally, the CST offers electro quasi-static solver (EQS), that is based on eq. 3 and results in electric potentials.

The comparison of all software tools and methods for electric potential computation is shown in Fig. 3. Comparisons obviously show that all software tools result in consistent electric potentials. The low differences are caused by equations that are computed by individual software tools. Note that all software tools are based on finite element method (FEM). The difference of two electric potentials and integration of lines between these points leads to the same results, since the COMSOL (difference) and COMSOL (integral) showed the same electric potentials on all electrodes. When integrating the same data in MATLAB, the electric potential is the same with COMSOL (integral). Finally, the above conclusions proved the correct computation of CST and integration in MATLAB.





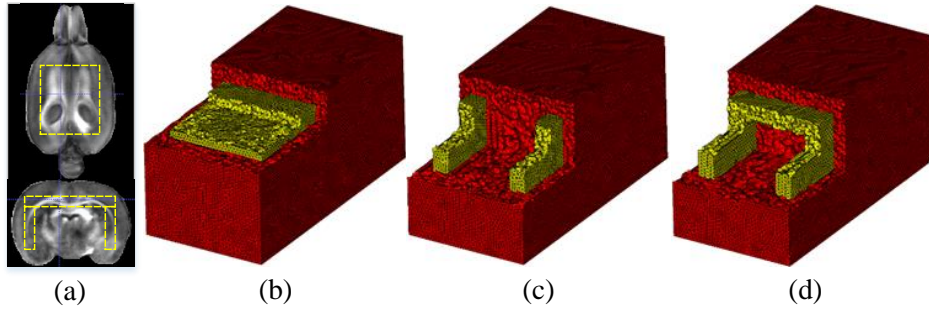
**Figure 3:** Electric potential on all electrodes computed by CST, COMSOL and FieldTrip.

#### 4. RAT BRAIN ANISOTROPY

Starting from DTI images presented in Fig. 4 (a), we approximated the most anisotropic tissues by blocks illustrated with yellow dashed lines. In DTI, the anisotropic tissues are represented by white color and rather isotropic ones by gray color. Based on the approximation of DTI, we constructed three 3D models differing in an anisotropic part only:

1. Horizontal part, Fig. 4 (b).
2. Vertical parts, Fig. 4 (c).
3. Combination of horizontal part and vertical one that create inverted U shape model, Fig. 4 (d).

The anisotropic parts were set to the values  $\sigma_{\parallel} = 0.28$  S/m and  $\sigma_{\perp} = 0.11$  S/m [2]. Except these anisotropic parts, the rest of the model was defined by the isotropic conductivity 0.33 S/m. The rest of the simulation setup remain the same as in section 3. Note that part of the 3D models in Fig. 4 is hidden to show the volume mesh used in simulations. Moreover, the anisotropic parts are illustrated by yellow color and isotropic ones by red color.

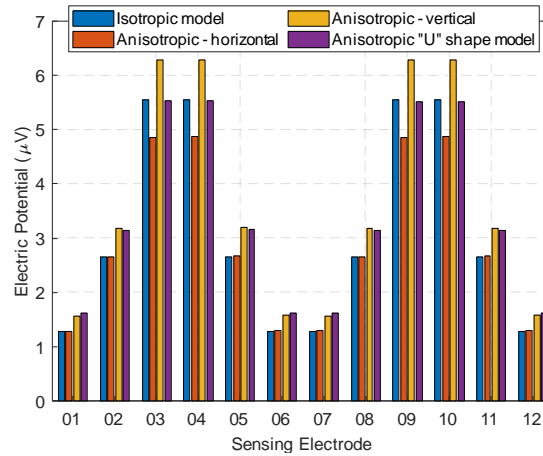


**Figure 4:** 3D models for anisotropic forward model: initial schematic based on DTI for model estimation (a), horizontal model (b), vertical model (c), inverted U-shape model (d).

The comparison of anisotropic simulations is shown in Fig. 5 together with an isotropic model. When the horizontal part was included only, the electric potential obviously decreased on electrodes 03, 04 and on symmetric ones 09, 10 with respect to the isotropic simulation. Since the perpendicular conductivity is lower than the longitudinal one and the dipole is oriented perpendicularly to the anisotropic part, the decrease in electric potential on electrodes 03, 04, 09, 10 is reasonable.

When the vertical part is included only, the dipole can radiate more energy towards the brain surface. That leads to the increase of the electric potential on all the electrodes with respect to the isotropic model. Finally, the inverted U-shape anisotropic model combines both the effects – the increase of the electric potential on all the electrodes except of the electrodes closest to the dipole (03, 04, 09, 10).

The simulations show that the anisotropy can influence the accuracy of the forward model, but the effect is rather weak. Moreover, the influence on the anisotropy should be analyzed with respect to the dipole orientation. Finally, we can confirm the conclusion from the paper [6]: the anisotropy should be involved, when high accuracy is required.



**Figure 5:** Influence of anisotropic tissues to the forward model.

## 5. CONCLUSION

In this paper, we demonstrated the use of the CST software for anisotropic forward model simulations. In order to recalculate the electric field intensity to the electric potential, the line integration in MATLAB was used. The procedure for the integration and the accuracy of CST was compared with two software packages. The comparison showed quite consistent electric potentials on all the electrodes and for all the software packages and methods. Originating from DTI, three anisotropic models were developed, and simulations were compared with an isotropic forward model. The comparison showed a rather weak effect of the anisotropy. In addition, the influence of the anisotropy depends on the dipole orientation. The future work will be focused on simulations with realistic anisotropic rat head models.

## ACKNOWLEDGMENT

This work was supported by the BUT specific research program (the project no. FEKT-S-20-6526).

## REFERENCES

- [1] J. Vorwerk, J. Cho, S. Rampp, et. al. "A Guideline for Head Volume Conductor modelling in EEG and MEG". *NeuroImage*, 2014, vol. 100, p. 590–607.
- [2] T. Goto, R. Hatanaka, T. Ogawa, et al. "An Evaluation of the Conductivity Profile in the Somatosensory Barrel Cortex of Wistar Rats", *Journal of Neurophysiology*, vol. 104, no. 6, pp. 3388–3412, 2010.
- [3] W. H. Lee, Z. Liu, B. A. Mueller, et. el. "Influence of white matter anisotropic conductivity on EEG source localization: Comparison to fMRI in human primary visual cortex", *Clinical Neurophysiology*, vol. 120, no. 12, pp. 2071–2081, 2009.
- [4] H. Hallez, B. Vanrumste, P. V. Hese, et. al. "Dipole estimation errors due to differences in modeling anisotropic conductivities in realistic head models for EEG source analysis", *Phys. Med. Biol.*, vol. 53, no. 7, pp. 1877–1894, 2008.
- [5] N. Samadzadehaghdam, B. MakkiAbadi, and S. Masjoodi, "Evaluating the Impact of White Matter Conductivity Anisotropy on Reconstructing EEG Sources by Linearly Constrained Minimum Variance Beamformer", *ABE*, vol. 9, no. 0, pp. 53–61, 2020.
- [6] D. Güllmar, J. Haueisen, and J. R. Reichenbach, "Influence of anisotropic electrical conductivity in white matter tissue on the EEG/MEG forward and inverse solution. A high-resolution whole head simulation study", *NeuroImage*, vol. 51, no. 1, pp. 145–163, 2010.
- [7] J. Vorwerk, R. Oostenveld, M. C. Piastra, et. al. "The FieldTrip-SimBio pipeline for EEG forward solutions", *BioMedical Engineering OnLine*, vol. 17, no. 1, p. 37, 2018.

# Mitochondria morphology and membrane potential under stress conditions

Masoumeh Ezati<sup>1</sup>, Inna Zumberg<sup>1</sup>, Vratislav Cmiel<sup>1</sup>

<sup>1</sup> Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technická 3082/12, 61600 Brno, Czech Republic

E-mail: [243468@vutbr.cz](mailto:243468@vutbr.cz) , [zumberg@vutbr.cz](mailto:zumberg@vutbr.cz) , [cmiel@vut.cz](mailto:cmiel@vut.cz)

**Abstract** — Mitochondria, organelles found in the cytoplasm of almost all eukaryotic cells, generate large quantities of energy in the form of adenosine triphosphate (ATP). In addition to producing energy, mitochondria store calcium for cell signaling activities and play an important role in maintenance of ionic balance. Mitochondria, however, are highly sensitive to any kind of stress in which they mainly response by disturbance of respiration, reactive oxygen species (ROS) production and release of cytochrome c into the cytoplasm. Osmotic stress, in particular, generates ROS that degrade lipids, proteins, and DNA. High levels of salt concentration can cause an imbalance in cellular ion homeostasis that results in ion toxicity and osmotic stress. This study aims to investigate possible effects of KCl and NaCl on Bone marrow-derived (MSCs) mitochondria membrane potential (MMP) and morphology. The results indicated that KCl and NaCl of salt concentration can cause an imbalance in cellular ion homeostasis that results in ion toxicity and osmotic stress. This study aims to investigate possible effects of KCl and NaCl on MSCs mitochondria membrane potential (MMP) and morphology. The results indicated that KCl and NaCl decreased the potential and changed the morphology of mitochondria membrane compared to cells growing in normal condition.

**Keywords** — Mitochondria membrane potential, Mitochondria morphology, Osmotic stress, Mitochondria fusion.

## 1. INTRODUCTION

Mitochondria are membrane-bound cell organelles that generate most of the chemical energy needed to power the cell's biochemical reactions. Chemical energy produced by the mitochondria is stored in a small molecule called adenosine triphosphate (ATP). Cells use this energy to synthesize proteins, power mechanical motion, sustain homeostasis, and perform other life-supporting functions. Mitochondria contain a double membrane. The outer membrane allows large molecules to flow into the mitochondrial intermembrane space, and the highly invaginated inner membrane, which has a large surface area, is responsible for oxidative phosphorylation. Diverse environmental stresses often induce similar kinds of cellular damage. For example, many, or even most, environmental stresses induce oxidative stress and protein denaturation. As a consequence, diverse stresses often illicit similar cellular adaptive responses, such as the production of stress proteins, up-regulation of oxidative stress protectors, and accumulation of protective solutes. In many cases, mitochondria are key sites of damage during environmental stress, especially mitochondrial electron transport. The mitochondria are highly networked and can change shape and size in response to stress. They also have been shown to join together in response to stress [1], which is known as mitochondria fusion. Mitochondrial fusion play critical role in maintaining functional mitochondria when cells experience metabolic or environmental stresses. Fusion helps mitigate stress by mixing the contents of partially damaged mitochondria as a form of complementation [2], [3].

Researches show that high levels of salt concentration can cause an imbalance in cellular ion homeostasis that results in ion toxicity and osmotic stress. This stress generates reactive oxygen species (ROS) that degrade lipids, proteins, and DNA. Potassium chloride (KCl) is a naturally occurring salt

derived from the ground or sea. It's a potassium-based salt that food manufacturers mostly use to replace sodium chloride, or table salt. KCl may affect the integrity of mitochondria by breaching the electrostatic force between the lipids and proteins. the mitochondrial membrane, cristae, and the matrix proteins appear altered under the influence of KCl. Sodium chloride (NaCl), also known as salt, is an essential compound our body uses to absorb and transport nutrients, maintain blood pressure, maintain the right balance of fluid, transmit nerve signals and contract and relax muscles. Human body needs NaCl to function, but too little or too much salt can be harmful to your health. Hypertonicity, induced by high NaCl, decreases cell volume, increases cytosolic osmolality, and changes mitochondrial osmotic equilibrium, which could affect mitochondrial function [4].

The objective of the present study is to investigate the mitochondrial membrane potential and morphology of MSCs in stress response during different ionic stress caused by KCl and NaCl. To find the optimum non-toxic concentrations of the salts, MTT [3–4,5-dimethylthiazol-2yl (2,5diphenyl-2H-tetrazoliumbromide)] assay was used. The cells then stained with JC-10 dye (AAT Bioquest, Inc) for investigating the MMP and MitoLite Red FX600 (AAT Bioquest, Inc) for determining the cells mitochondrial morphology. The morphology of the stained cells were studies using confocal fluorescent microscopy. The results suggested that KCl and NaCl caused reduction in MMP and morphological changes in the mitochondria membrane. The results also indicated that KCl may induce the mitochondria fusion under different stress conditions after 3 and 24 hours.

## **2. MATERIALS AND METHODS**

### **2.1 CELL CULTURE**

MSCs was used for this study and cultured in DMEM supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin at 37°C in a 5% CO<sub>2</sub> (carbon dioxide). The culture medium was changed every 2 days. For biological experiments, cultured cells were detached by trypsinization, suspended in a new culture medium and used for designed experiments.

### **2.2 CELL VIABILITY ASSAYS**

To examine cell viability, MSCs cells were plated 24 hours before the experiment in 96-well plates at a density of  $3 \times 10^4$  cells/well. The cells were then treated for 24 hours with KCl and NaCl at several concentrations. Then, MTT assay was used to study the viability of MSCs. For the measurement, medium was removed and replaced with 100  $\mu$ L of MTT reagent (Sigma-Aldrich, United States) (5 mg/ml) and left at 37°C for 3 h; MTT solution was removed and the formazan crystals were then dissolved in solubilizing solution and transferred into a new 96-well plate. The plate then was scanned at 570 and 650 nm (for background subtraction) using a Portable Biotek Epoch Microplate Spectrophotometer.

### **2.3 MEASUREMENT AF MITOCHONDRIAL MEMBRANE POTENTIAL**

Effects of KCl and NaCl on mitochondrial membrane potential in MSCs cells was investigated by JC-10 staining, which is an indicator of cell health. MSCs were incubated with 1  $\mu$ M of the dye at 37°C for 30 minutes. JC-10 is capable of selectively entering into mitochondria, and reversibly changes its color from green (monomeric form) to orange (aggregate form) as membrane potentials increase.

### **2.4 MITOCHONDRIAL MORPHOLOGY**

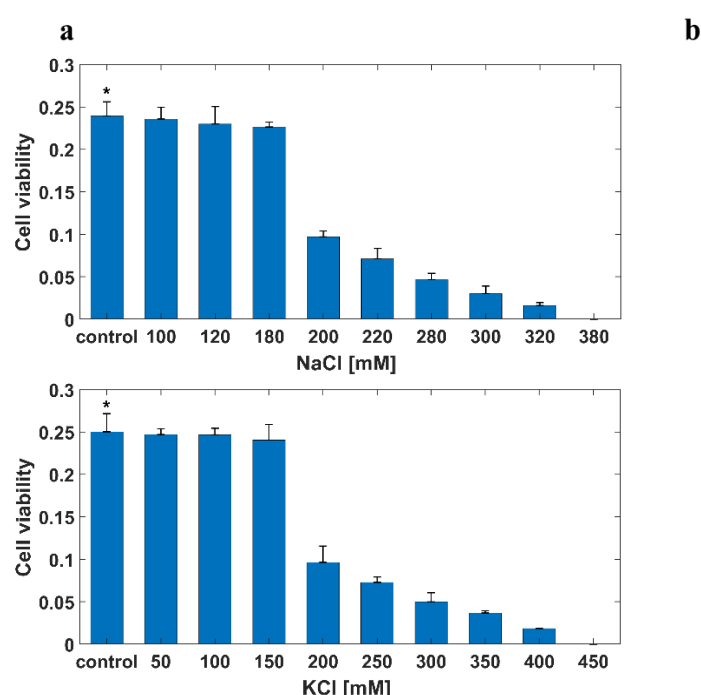
Mito-Lite Red FX600 reagent was used for determining mitochondrial morphology. Cells were plated at  $5 \times 10^4$  cells on confocal dish for quantitating fluorescence. The medium was removed, and cells were washed with PBS before staining. Cells were incubated with 2  $\mu$ l of 500 X MitoLite stock solution in

1000  $\mu$ l of HBSS for 20 min at 37°C. MitoLite Red FX600 has excitation/emission maxima of approximately 580/598 nm.

### 3. RESULTS AND DISCUSSION

#### 3.1 CELL VIABILITY ASSAYS

Mitochondria are fundamental organelles in animal and plant cells for energy supply by synthesis of ATP via oxidative and are thus highly sensitive to impact of stress, disease or ageing. In this experiment different concentrations of KCl and NaCl were used. The cell viability after KCl and NaCl exposed was determined by MTT test against the MSCs. According to **Fig.1 a** and **b** cell death was begun after incubation with the solution of 200 mM NaCl and KCl respectively, which indicates the cytotoxicity of these solutions. After 24 hours of incubation, MSCs were death completely in the medium contained 380 mM NaCl and 450 KCl. Therefore, 180 mM NaCl and 150 mM KCl were used for MMP measurement and mitochondrial morphology analysis.

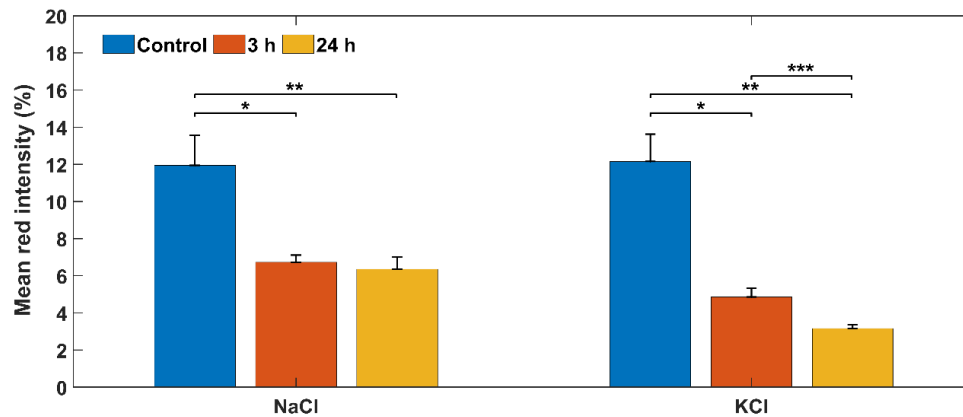


**Figure 1:** Cell viability on MSCs with MTT assay after 24 h treatment of MSCs with various concentrations of NaCl (a) and KCl (b). The control shows MSCs cells without the exposure to NaCl and KCl. Data are presented as mean  $\pm$  SD; n = 3. \* $p$ <0.05 control compared to different concentration of NaCl and KCl (the unequal variances two-sample t-test, or Welch's t-test was used).

#### 3.2 STUDY OF MITOCHONDRIAL MEMBRANE POTENTIAL

One of the most common methods to detect mitochondrial dysfunction is to monitor the mitochondrial membrane potential. Mitochondria are critical oxygen sensors linked to various protective effects, such as enhancement of antioxidant defense, cell survival and anti-apoptosis to determine the mitochondrial membrane potential in MSCs (control) and MSCs exposed KCl and NaCl was assessed using membrane-permeant dual-emission potential-sensitive JC-10 dye. Confocal image analysis consisted of calculating mean red intensity (amount of J-aggregates) in each group using the MATLAB software. According to confocal images in the MSC cells, increasing osmolality by adding NaCl causes decreases mitochondrial membrane potential and reduces cell volume after 3 and 24 hours compared to control. However, there was no significant difference between the samples exposed to NaCl after 3 and 24 h. The same results were observed with the cells treated with KCl, however, KCl effect on MMP was

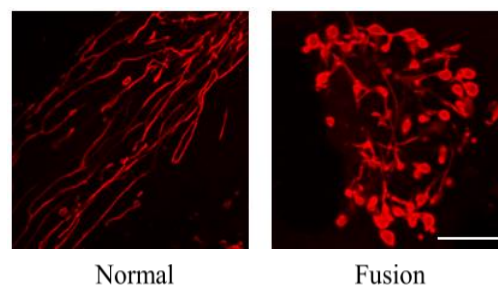
significantly higher than that of NaCl. In contrast to the cells treated with NaCl, a significant difference was observed between the exposed cells to KCl after 3 and 24 h. (**Fig 2**).



**Figure 2:** Mitochondrial membrane potential assessment of MSCs exposed to 180 mM NaCl and 150 mM KCl for 3 and 24 h. Data are presented as mean  $\pm$  SD; n = 3. \* $p$ <0.05 control compared to NaCl and KCl after 3 h, \*\* $p$ <0.05 control compared to NaCl and KCl after 24 h, \*\*\* $p$ <0.05 KCl (3h) compared to KCl (24h). The unequal variances two-sample t-test, or Welch's t-test was used.

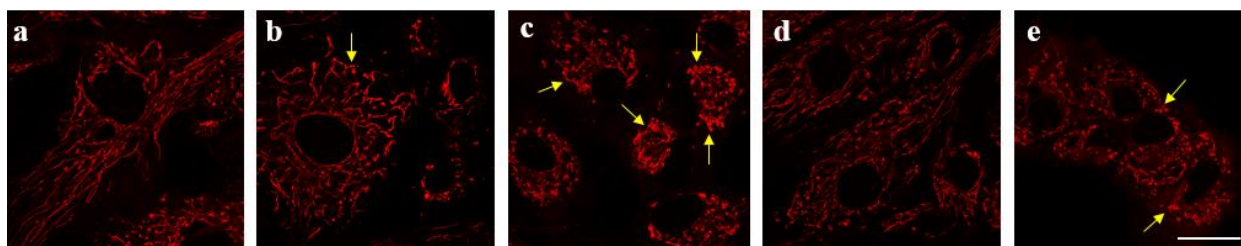
### 3.3 MITOCHONDRIAL MORPHOLOGY

Mitochondria are highly dynamic organelles and attain different shapes reflecting different cellular states, during the lifetime of a cell. In most mammalian cells, mitochondria are tubular in shape under normal conditions but may attain various forms during various cellular perturbations. During apoptosis mitochondria are extensively fragmented and form small punctuate or round structures, while during necrosis mitochondria usually swell and become distended. Fusion is another behavior of cell mitochondria during abnormal condition [5] (**Fig 3**). Fusion helps mitigate stress by mixing the contents of partially damaged mitochondria as a form of complementation and as a result maintain one of the most important vital functions, namely respiration. According to our confocal images, after 3 hours of incubation with KCl and NaCl, no significant changes in the cells mitochondria morphology was observed compared to control (**Fig 4 a**). However, in some regions the early stage fusion was seen. After 24 hours of incubation, the morphological changes were significantly noticeable. The mitochondria were completely round shape and fusion between them was obvious (**Fig 4 b, c, d, e**). These results are in line with previous reports which suggested that mitochondrial fusion and formation of protuberances of the outer mitochondrial membrane are induced by elevated levels of KCl and NaCl. According to the literature [6], the higher concentration of KCl may affect the integrity of mitochondria by breaching the electrostatic force between the lipids and proteins. In addition, adding NaCl decreases mitochondrial membrane potential by causing NADH redistribution out of mitochondria [7].



**Figure 3:** Confocal images of cells in normal (left) and abnormal condition (right). Scale bar 8  $\mu$ m.





**Figure 4:** Confocal fluorescence microscopic images of mitochondria in MSCs in control (a), 3 h (b) and 24 h (c) after treatment with KCl and 3 h (d) and 24 (e) after treatment with NaCl. Yellow arrows show locations where mitochondria fusion can be seen. Scale bar 30  $\mu\text{m}$ .

#### 4. CONCLUSION

This study aims to investigate the possible morphological and membrane potential changes of MSCs cells mitochondria under stress conditions. In this regard, the cells exposed to NaCl and KCl provided the osmotic condition for the cells. Based on the obtained results, it was found that both NaCl and KCl strongly decreased mitochondrial membrane potential after 24 hours. In addition, the mitochondria morphology analysis confirmed that in some cases, the mitochondria normal structure (tubular shape) changed to round shape under the stress condition. The results also confirmed that putting mitochondria under environmental stress condition may lead to their fusion to large networks may help the organism to cope with the stress situation.

#### REFERENCES

- [1] B. Westermann, "Mitochondrial fusion and fission in cell life and death," *Nature Reviews Molecular Cell Biology*, vol. 11, no. 12, pp. 872–884, Dec. 2010, doi: 10.1038/nrm3013.
- [2] R. J. Youle and A. M. van der Bliek, "Mitochondrial Fission, Fusion, and Stress," *Science*, vol. 337, no. 6098, pp. 1062–1065, Aug. 2012, doi: 10.1126/science.1219855.
- [3] A. S. Rambold, B. Kostelecky, N. Elia, and J. Lippincott-Schwartz, "Tubular network formation protects mitochondria from autophagosomal degradation during nutrient starvation," *Proceedings of the National Academy of Sciences*, vol. 108, no. 25, pp. 10190–10195, Jun. 2011, doi: 10.1073/pnas.1107402108.
- [4] P. Sharma, A. B. Jha, R. S. Dubey, and M. Pessarakli, "Reactive Oxygen Species, Oxidative Damage, and Antioxidative Defense Mechanism in Plants under Stressful Conditions," *Journal of Botany*, vol. 2012, pp. 1–26, Apr. 2012, doi: 10.1155/2012/217037.
- [5] Ma. L. Escobar, O. M. Echeverría, and G. H. Vázquez-Nin, "Necrosis as Programmed Cell Death," in *Cell Death - Autophagy, Apoptosis and Necrosis*, InTech, 2015. doi: 10.5772/61483.
- [6] S. Mishra and R. Mishra, "Molecular Integrity of Mitochondria Alters by Potassium Chloride," *International Journal of Proteomics*, vol. 2015, pp. 1–12, Dec. 2015, doi: 10.1155/2015/647408.
- [7] L. Michea, C. Combs, P. Andrews, N. Dmitrieva, and M. B. Burg, "Mitochondrial dysfunction is an early event in high-NaCl-induced apoptosis of mIMCD3 cells," *American Journal of Physiology-Renal Physiology*, vol. 282, no. 6, pp. F981–F990, Jun. 2002, doi: 10.1152/ajprenal.00301.2001.

# Optimization of wavelet transform in the task of intracardiac ECG segmentation

R. Ředina<sup>1,2</sup>

<sup>1</sup>Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology, Brno, Czech Republic

<sup>2</sup>International Clinical Research Centre (ICRC), St. Anne's University Hospital, Brno, Czech Republic

E-mail: [195715@vut.cz](mailto:195715@vut.cz)

**Abstract**— My work deals with the selection of an appropriate wavelet transform setting for feature extraction from intracardiac ECG recordings. The studied signals were obtained during electrophysiological examinations at the Department of Pediatric Medicine, University Hospital Brno. In this paper, several wavelets are tested for feature extraction which is followed by adaptive thresholding to detect atrial activity from the extracted features. The procedure is evaluated using the F-score. Although the presented procedure does not appear to be overall effective for intracardiac signal segmentation, it certainly does not reject the use of wavelet transforms in combination with advanced machine learning, neural network, or deep learning techniques.

**Keywords**— ECG, Intracardiac ECG, Atrial activity, Wavelet transform, Adaptive threshold, F-score

## 1. INTRODUCTION

The beginning of ECG scanning at the end of the previous century marked a milestone in the understanding of heart disease. Over time, computer technology and, in recent years, state-of-the-art machine learning techniques have become involved in analyzing this signal alongside specialist physicians. We are now on the dawn of another milestone as the clinical relevance of electrophysiology itself is beginning to grow. The output of today's electrophysiological examination is the mapping of electrical potentials directly from the heart cavities. The evaluation of intracardiac ECG (*iECG*) signals is now performed by a clinical specialist and there is therefore a place for the involvement of computer technology.[1]

The first task that must be performed is the implementation of a segmentation algorithm that can find the key parts in the intracardiac signals. Only a correctly segmented signal can be used for subsequent analysis.

Since this is a completely new type of signal, a critical approach to its analysis is needed. From the available literature, wavelet transform has been reported as one of the effective methods for segmentation of surface ECG. Therefore, we can accept the premise that this method could provide good results also in the case of intracardiac signals. In this paper, the application of the method as well as the optimization of the adjustable parameters were addressed to achieve the highest possible segmentation quality.

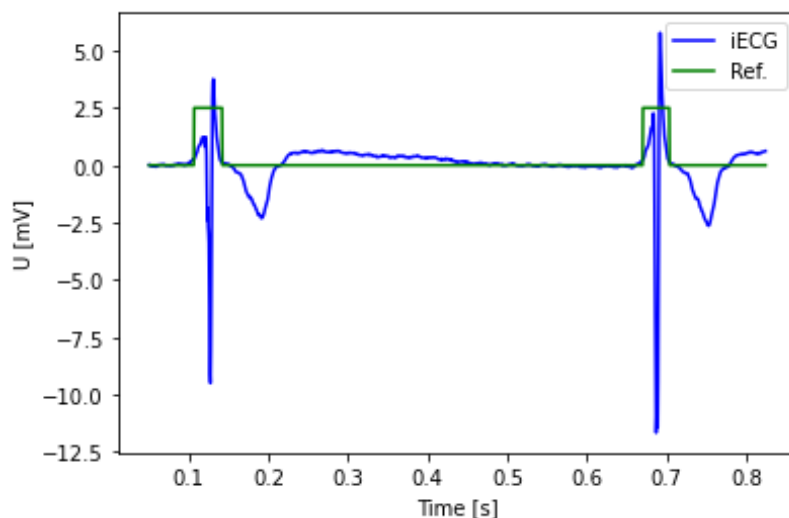
## 2. DATA

The experimental data used in this study were collected at the Department of Pediatric Medicine, University Hospital Brno. A total of 100 patients underwent electrophysiological examination and signals of different lengths (five to twenty seconds) were recorded for a total of 348 recordings.

Each recording consisted of a classic surface twelve-lead ECG (I, II, III, avR, avL, avF, V1 to V6). In addition, there are five intracardiac leads that are imaged with a catheter inserted in the coronary sinus (cs1-2 – cs9-10). It should be noted that the leads are numbered on the catheter from its distal end, so lead cs1-2 is inserted deepest in the coronary sinus and lead cs9-10 just behind its mouth in the right atrium. Thus, the depolarizing wave propagating from the sinoatrial node down the right atrium was first recorded in lead cs9-10 while lead cs1-2 was only pick up the last phase of atrial depolarization.

It is these five leads that were crucial for this work. As the signals are acquired in the atria, the main objective was to mark atrial activity in the signal.

An example of the acquired signal can be seen in Figure 1 below. The atrial activity (AA) is represented here by a high frequency oscillation that could resemble a QRS complex compared to a conventional surface ECG. The ventricular activity is captured in the form of a subsequent low-frequency wave as a far-field signal. The signal annotation was created by an experienced expert at an international clinical research center.



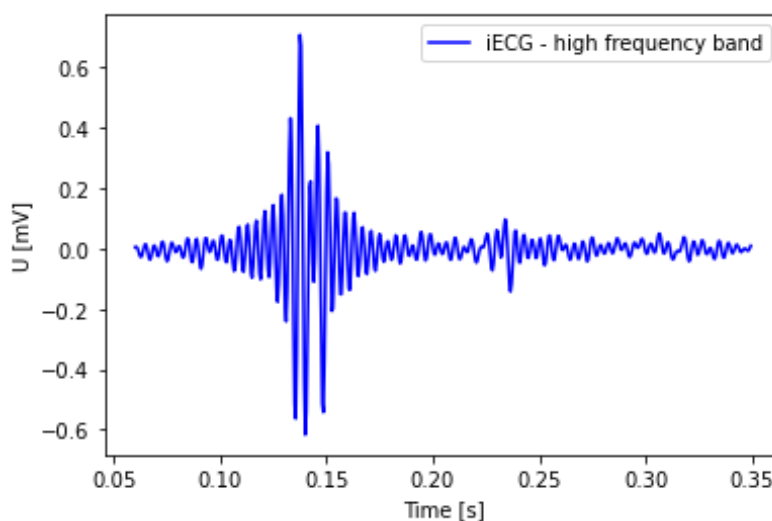
**Figure 1:** iECG signal with marked atrial activity (Ref.)

### 3. METHODS

The wavelet transform proved to be an effective tool for surface ECG analysis [2]. The wavelet transform performs a time-frequency dependent decomposition of the signal into a pre-selected number of frequency bands. The key role is played by individual wavelets, which can be understood as digital filters through which the signal is decomposed into different frequency bands [3]. In this paper dyadic wavelet transform is used.

The whole procedure described above is summarized in Figure 2 below. In this work, wavelets from different families were successively investigated, such as Haar, Daubechies, Biorthogonal, Coiflets or Symlets.

The decomposition level was set to level 3 by default. Atrial activity is picked up in the intracardiac leads by a high-frequency oscillation. However, the specific frequency distribution depends on specific anatomical conditions and catheter position, which may not always be the same.



**Figure 2:** High-frequency band of decomposed iECG signal

The band containing the highest frequencies was then selected and this was passed on for further analysis. Signal segmentation was performed based on adaptive thresholding [4]. The value of the calculated wavelet transform was monitored in the floating window. The moment it exceeded three times the standard deviation, the position was evaluated as belonging to atrial activity.

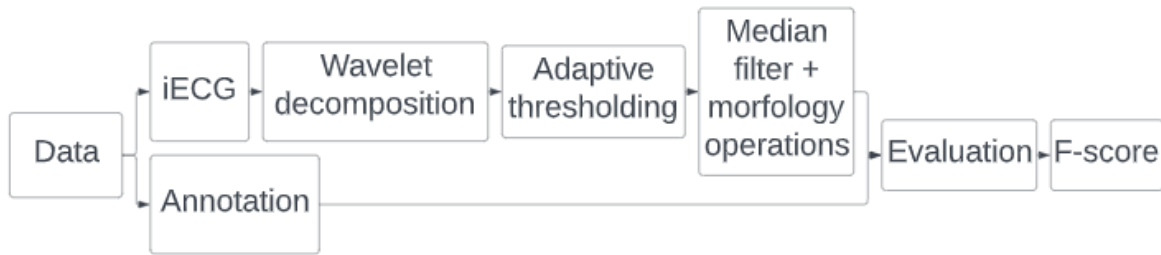
Since not only one but a total of 5 leads are examined, it is important to consider the position at which each lead is measured. As it follows from the description of the data above, the beginning of depolarization can be expected in lead cs9-10 and its end in lead cs1-2. Therefore, the beginning of the detection was searched in the first mentioned lead and its end in the second one.

The evaluation of the algorithm was performed by calculating the F-score for each wavelet. The F score was calculated as the ratio of twice the total non-zero hits to the sum of twice the correctly determined values and all the false results. The formula is presented below:

$$F = \frac{2TP}{2TP + FN + FP'} \quad (1)$$

Where TP are *true positive* predictions, FP are *false positive* predictions and FN are *false negative* predictions.

After the thresholding is done, it is necessary to perform postprocessing of the acquired data. Due to potential lone detections, a median filter was included in the postprocessing to remove this impulse noise. The median filter was implemented in a sliding window of 17 samples. To maximize the effect, a morphological dilation operation was applied to the signal to eliminate false negative detections. The dilation was set to be able to cover false detections of 14 samples in length. Median filtering was then performed, followed by signal erosion to cancel out any dilation effect and further remove even larger false positive sections. The erosion parameter was set to 14. After erosion, median filtering was again performed. At the end of postprocessing, dilation was included with a parameter of 7 to cancel out the previous twice as large erosion. The results before postprocessing are shown in Figure 3.



**Figure 3:** Block diagram of the algorithm

#### 4. RESULTS

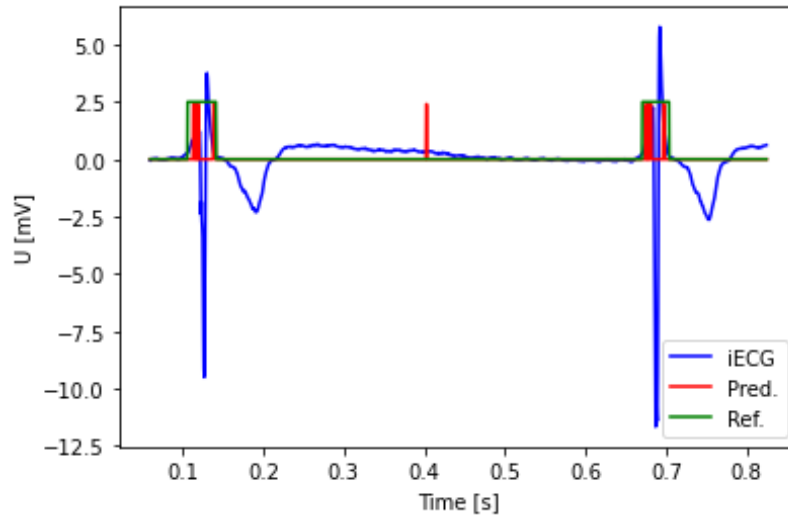
**Table I:** F-score of detection for every wavelet form

Wavelet	F-score	Wavelet	F-score
haar	0.183	<b>db4</b>	<b>0.235</b>
coif1	0.188	db5	0.195
coif2	0.187	bior1.3	0.179
coif3	0.110	bior1.5	0.205
coif4	0.078	bior2.2	0.197
db2	0.201	sym3	0.205
db3	0.205	sym4	0.184

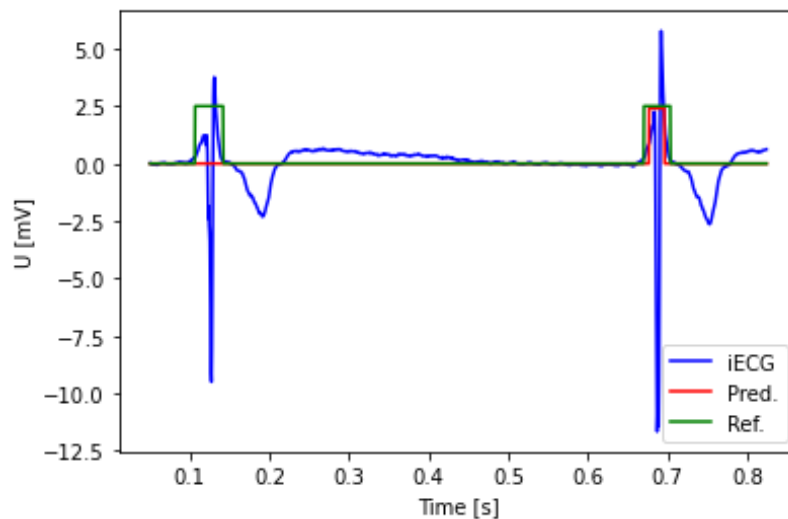
The segmentation results depending on the type of wavelet used are summarized in Table I. The best

result in the context of this thesis was achieved for the db4 wavelet, where the F-score reached 0.235. This result alone is not satisfactory. However, in the context of the other results, it is possible to accept the hypothesis that not all wavelets are suitable for symptom extraction in each experimental setup.

Table I clearly shows that the db4 wavelet performs best in this experimental setup. This is likely associated with similar wave morphology and atrial activity on iECG. The db4 wavelet is followed by bior1.5 and db3. One possible outcome is captured in Figure 3. The underestimation of the length of the segmented section is probably due to the fact that the beginning of the wave is more gradual and therefore does not have such a pronounced response after the wavelet transform.



**Figure 4:** Segmentation results before postprocessing procedure. Green reference signal. Algorithm output in red.



**Figure 5:** Final segmentation results. Green reference signal. Algorithm output in red.

## 5. CONCLUSION

The aim of this study was to assess the influence of different types of waves on the result of segmentation of intracardiac ECG recordings. On the other hand, the work did not aim to create the best possible algorithm for segmenting the signals. By comparing the results from different settings of the feature extraction step, it was concluded that the db4 wavelet from the Daubechies family would be the best possible choice in this context.

Although the presented procedure cannot be interpreted as an efficient segmentation algorithm, there is no reason to reject the use of the wavelet transform for feature extraction, applicable in more complex segmentation algorithms. With the use of modern machine learning techniques or artificial neural networks, the wavelet transform can be a valuable tool.

With further refinement of the thresholding itself, better results could be achieved as well as adjusting the median filtering and other parameters in postprocessing. The whole procedure contains several optimizable variables.

## **ACKNOWLEDGMENT**

Brno Ph.D. Talent Scholarship Holder – Funded by the Brno City Municipality.

## **REFERENCES**

- [1] EISENBERGER, Martin, Alan BULAVA a Martin FIALA. Základy srdeční elektrofyzologie a katéetrových ablací. Praha: Grada, 2012. ISBN 9788024736778.
- [2] Hejč J, Vítek M, Ronzhina M, Nováková M, Kolářová J. A Wavelet-Based ECG Delineation Method: Adaptation to an Experimental Electrograms with Manifested Global Ischemia. *Cardiovasc Eng Technol*. 2015 Sep;6(3):364-75. doi: 10.1007/s13239-015-0224-z. Epub 2015 Apr 8. PMID: 26577367.
- [3] Addison, P. S. Wavelet transforms and the ECG: a review. *Physiol. Meas.* 26:155–199, 2005.
- [4] LU, Xuanyu, Maolin PAN a Yang YU. QRS Detection Based on Improved Adaptive Threshold. *Journal of Healthcare Engineering* [online]. 2018, 2018, 1-8 [cit. 2022-03-13]. ISSN 2040-2295. Dostupné z: doi:10.1155/2018/5694595



# Characterization of osteolytic lesions by low-dose spectral CT in myeloma patients

M. Nohel<sup>1</sup>, J. Jan<sup>1</sup>, and J. Chmelik<sup>1</sup>

<sup>1</sup>Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology, Brno, Czech Republic

E-mail: [xnohel04@vutbr.cz](mailto:xnohel04@vutbr.cz), [jan@vut.cz](mailto:jan@vut.cz), [chmelikj@vut.cz](mailto:chmelikj@vut.cz)

**Abstract**—This paper presents a preliminary study of characterization of focal osteolytic lesions in low-dose spectral CT in myeloma patients. Spectral CT with energy decomposition into two energies allows us to use post-processing software for creating several parametric maps as well as so-called monoenergetic images. The paper includes a demonstration of the different contrast of lytic lesions in the spine and a comparison of healthy vertebrae and vertebrae affected by focal lytic lesions. It is shown that lytic lesions are better recognizable on a 40 keV monoenergetic image compared to conventional CT.

**Keywords**—Spectral CT, Multiple myeloma disease, Skeletal imaging, lytic lesions

## 1. INTRODUCTION

A hematological disease characterized by clonal proliferation of plasma cells in the bone marrow is known as multiple myeloma. One of the most common manifestations of the disease is skeletal involvement. It is usually manifested by the presence of osteolytic bone lesions [1]. Sometimes, plasma cells can proliferate outside the bone marrow, and then we call this extramedullary disease [2]. In most patients, multiple myeloma is preceded by a stage called monoclonal gamopathy of undetermined significance (MGUS). Some patients may develop an intermediate stage referred to as smoldering myeloma. The presence of bone lesions is one of the diagnostic criteria that indicate the progression of MGUS to multiple myeloma. Unlike multiple myeloma with advanced skeletal osteolytic involvement and other symptoms, including anemia, hypercalcemia, and renal failure, MGUS, and smoldering myeloma are asymptomatic conditions. [3, 4]

Early initiation of therapy requires early detection of osteolytic lesions – a very common manifestation of the disease. In addition to magnetic resonance (MRI) and hybrid imaging methods (especially PET/CT), the use of low-dose CT is of fundamental importance today [5]. The bone marrow can be imaged using MRI [6]. PET/CT then can provide information regarding the activity of skeletal lesions [7]. The importance of low-dose CT lies in the possibility of obtaining especially morphological information regarding the presence of osteolytic lesions. On the other hand, compared to PET/CT, the examination is burdened with lower radiation exposure, however, when compared to MRI, the advantage is a shorter duration of the examination and better availability [8]. The radiation dose of low-dose CT is lower than that of standard examination protocol, although in comparison with a radiographic examination the radiation dose is higher [9,10]. However, with modern dual-energy CT and optimal protocol settings, it is possible to achieve a similar value of dose as in radiographic examination [11, 12]. Nevertheless, considering the benefits, low-dose CT has its place in the current guidelines of international societies dealing with multiple myeloma [5].

Currently, the so-called spectral CT (sCT) with energy decomposition into two energies by two-layer detectors from Philips Healthcare is used for an increasing number of indications. Unlike conventional CT (cCT) examination, sCT can separate signals from photons with different energy (technical design varies from manufacturer to manufacturer by using two different X-ray energies (dual-energy CT) or two-layer detectors with different sensitivity for different X-ray energies (sCT)). This allows us to use post-processing software for creating several parametric maps as well as so-called monoenergetic images [13].

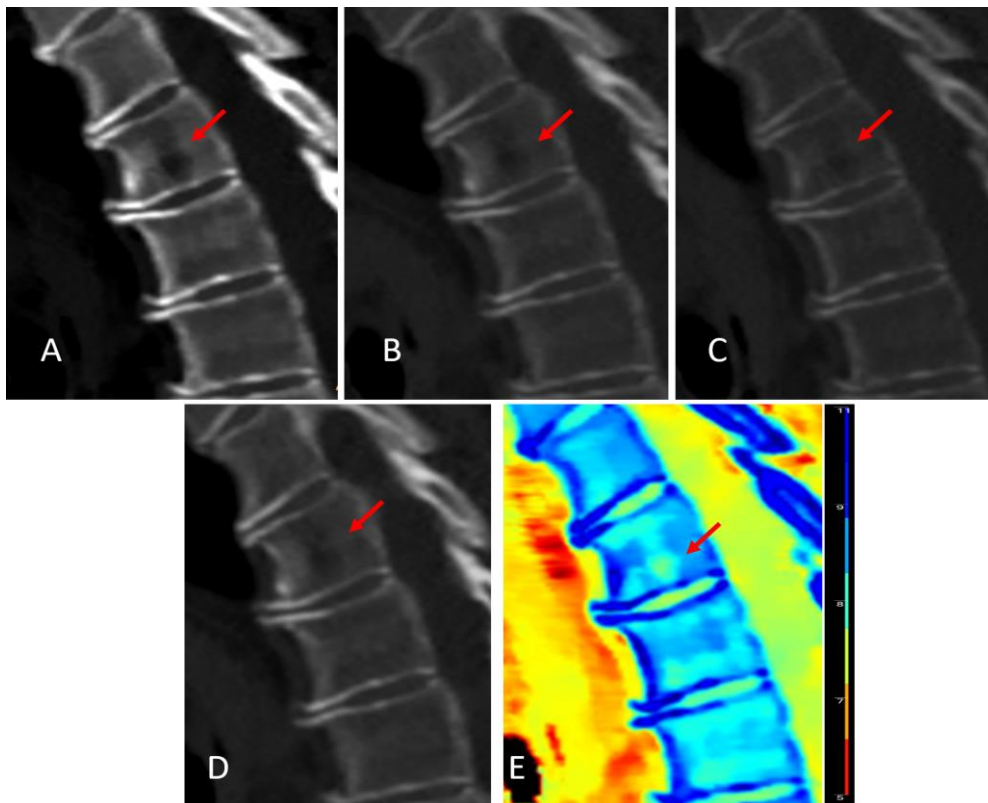
## 2. DATA

In this work, an anonymized database of oncological patients with multiple myeloma disease examined by spectral CT scanner was used. The data were obtained through Philips Healthcare IQon spectral CT in cooperation with University Hospital Brno, Department of Radiology and Nuclear Medicine, where parametric images were created. The database contains scans of six patients, where the monoenergetic images on 40, 80, and 120 keV are reconstructed for each patient. Then conventional CT scan with peak tube voltage 120 keV and an effective proton number map are available. Two patients have a focal lesion in vertebrae, two have diffuse infiltrated lesions and two are without any lesion. In this paper, only data from patients with focal lytic lesions in vertebrae were used for the preliminary study.

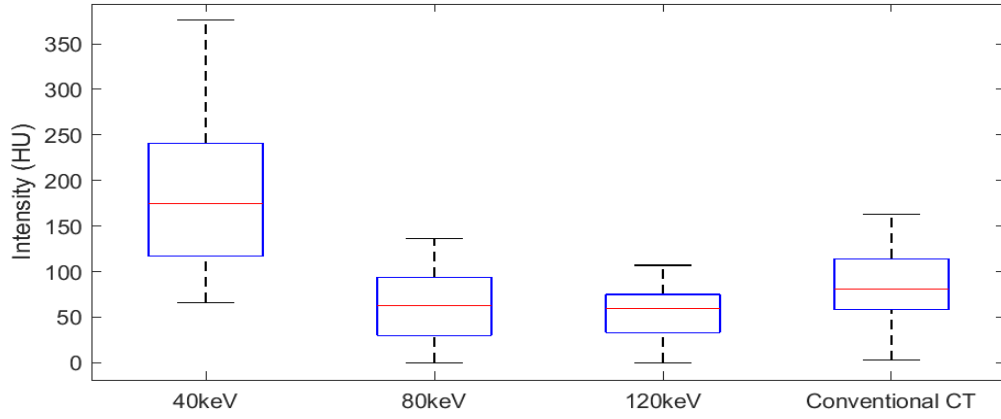
A commercially available workstation with the Intellispace Portal client (version 12.1), Philips Healthcare, was used to visualize the individual scans and parametric images. Access to this software is possible thanks to the cooperation with Philips Healthcare company. The implementation and analysis of the data were performed in the MathWorks MATLAB 2021b programming environment. The Image processing toolbox for visualization of data was used and the Statistics toolbox was used for statistical analysis.

## 3. FOCAL LESION IN DIFFERENT MONOENERGETIC IMAGES

Spectral CT has the advantage of being able to reconstruct parametric images at different monoenergetic levels. This makes it possible to see the difference in tissue density with greater contrast using lower monoenergetic levels. This can be seen in Figure 1, where for a monoenergetic image at 40 keV the lytic lesion in the third thoracic vertebra (Th3) is best seen compared to higher energy images or a conventional CT scan. Spectral CT also allows the creation of an effective proton number map, in which the lytic lesion can be easily recognizable. Monoenergetic images at lower energy levels and an effective proton number map can be very helpful for lesion detection. Figure 2 shows a boxplot of the segmented lesion on individual monoenergetic images compared to a conventional CT scan. We can see the monoenergetic image at 40keV in which the lesion has a much larger variance of density values than in the other images.



**Figure 1:** Illustration of the focal lesion in a third thoracic vertebra (red arrow) in images obtained from spectral CT on 40 (A), 80 (B), and 120 (C) keV monoenergetic images, conventional CT (D) with peak tube voltage 120 keV and effective proton number map (E).

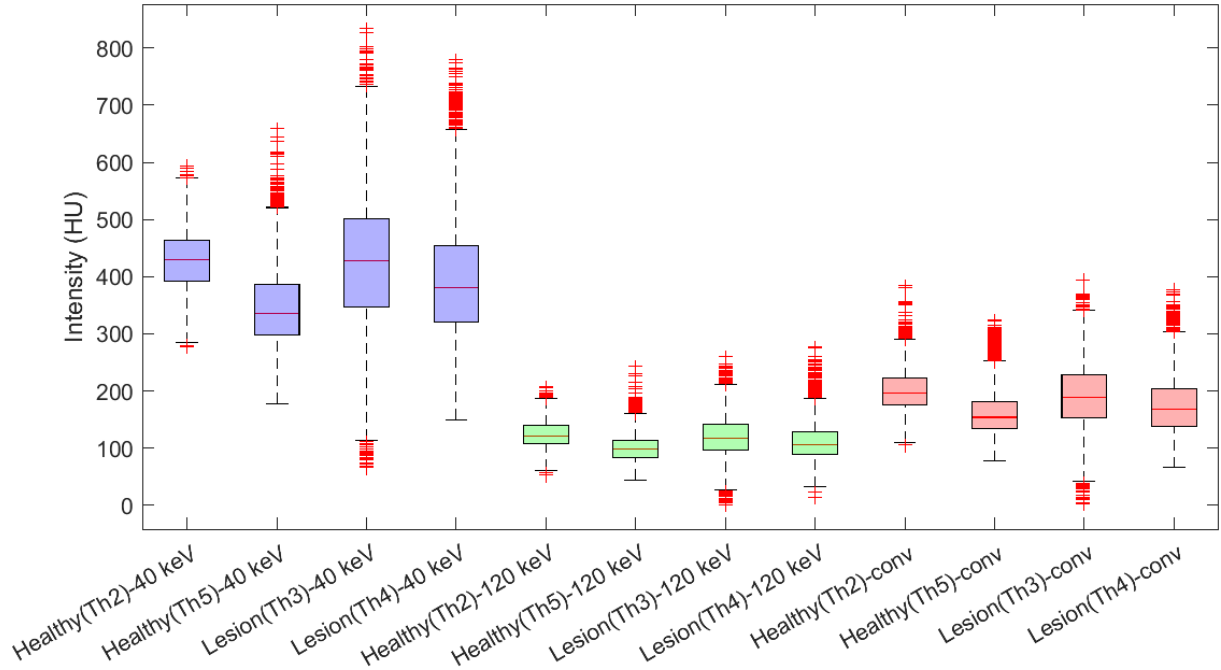


**Figure 2:** Comparison of the box-and-whisker plot of the focal lesion in a third thoracic vertebra in 40, 80, and 120 keV monoenergetic images and conventional CT with peak tube voltage 120 keV. The red line indicates the median value in the dataset, the blue box indicates interquartile range and the black lines indicate minimum and maximum in the dataset.

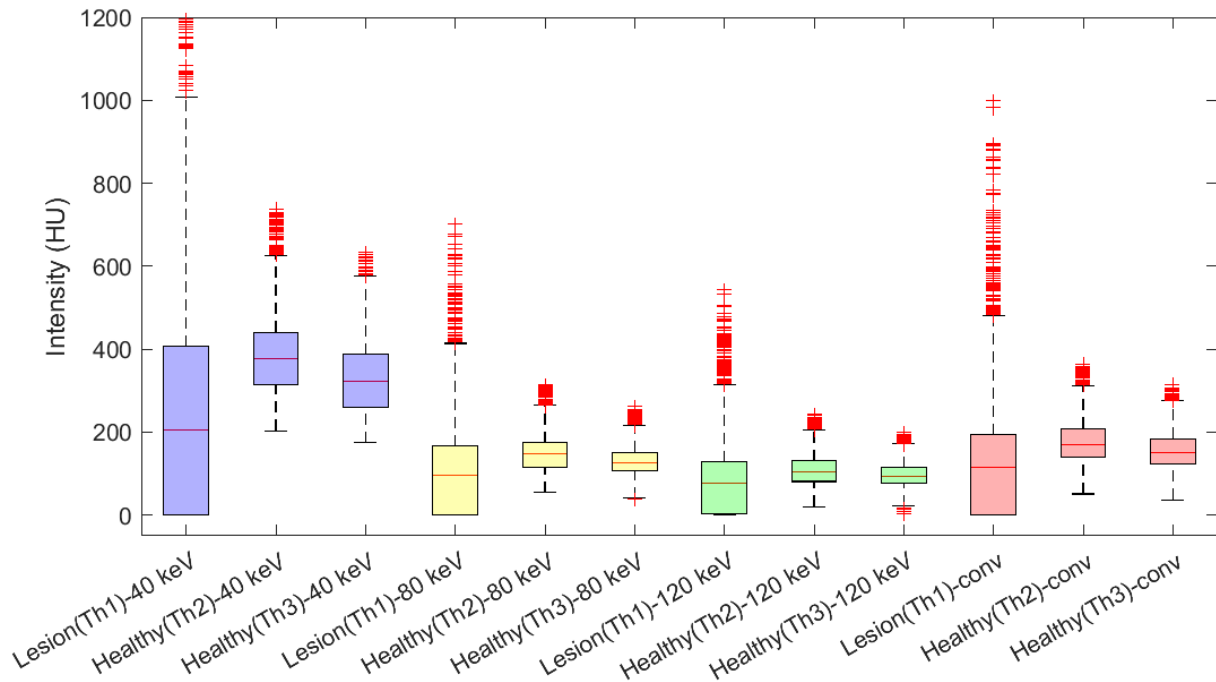
#### 4. DIFFERENCE BETWEEN HEALTHY AND AFFECTED VERTEBRAE

Initial experiments on available spectral CT data compared vertebrae from patients with focal vertebral lesions with healthy vertebrae from the same patient. For individual patients affected by focal lesions, vertebral bodies containing the focal lesion and healthy vertebrae were manually segmented. The region of interest was the trabecular part of the vertebral body. Boxplots of individual vertebrae in various monoenergetic images and conventional CT are shown in Figures 3 and 4. Figure 3 shows data from patient one, where the second thoracic vertebra (Th2) and Th5 do not contain lesions, and vertebrae Th3 and Th4 contain focal lesions. Figure 4 shows data from patient two, where the Th2 and Th3 vertebrae do not contain lesions and the Th1 vertebra contains a focal lesion.

The graphs below clearly show that the variance of the vertebrae affected by focal lesions (Th3 and Th4 in Fig. 3 and Th1 in Fig. 4) is significantly greater than that of healthy vertebrae. The variance is highest for the 40 keV monoenergetic image where the lesions are most visible.



**Figure 3:** Box-and-whisker plot of segmented trabecular part of body vertebra in healthy thoracic vertebrae (Th2, Th5) and thoracic vertebrae affected by focal lesions (Th3, Th4) on 40 (blue) and 120 (green) keV monoenergetic images and conventional CT (red) for patient one. The red line indicates the median value in the dataset, the rectangle box indicates interquartile range, the black lines indicate minimum and maximum and the red marks are outliers in the dataset.



**Figure 4:** Box-and-whisker plot of segmented trabecular part of body vertebra in healthy thoracic vertebrae (Th2, Th3) and thoracic vertebra affected by focal lesions (Th1) on 40 (blue), 80 (yellow), and 120 (green) keV monoenergetic images and conventional CT (red) for patient two. The red line indicates the median value in the dataset, the rectangle box indicates interquartile range, black lines indicate minimum and maximum and red marks are outliers in the dataset.

Statistical evaluation of data of healthy and diseased vertebrae was performed. First, the null hypothesis that the variance of the healthy and diseased vertebrae values are equal or similar was tested. Values were tested on 40 and 120 keV monoenergetic images and conventional CT images. Using Bartlett's test, which tests the agreement of variances, we rejected this hypothesis for all three images because the alpha value is less than 0.05. Thus, it can be said that there is a statistical difference in the homogeneity of variance for healthy and affected vertebrae in the individual parametric images.

However, using Levene's test, we did not reject the hypothesis because the value is greater than 0.05. This test showed that monoenergetic levels of 40 and 120 keV have the same significance ( $p=0.294$ ) and have a higher significance than conventional CT ( $p=0.289$ ).

Next, an ANOVA test was performed to compare the variances of healthy and affected vertebrae with the null hypothesis that the mean value does not depend on vertebral lesion involvement. The p-value came out to be 0.01 and hence the hypothesis is rejected.

## 5. CONCLUSION

This contribution deals with the characterization of focal osteolytic lesions on spectral CT images. The differences in contrast in the generated parametric images with energies of 40, 80, and 120 keV compared to conventional CT scan with peak tube voltage 120 keV are shown. An effective proton number map is also shown, where the lesion is very well visible. Density values from manually segmented healthy vertebrae and vertebrae affected by focal lesions were shown in box-and-whisker plots for single monoenergetic images and conventional CT. Monoenergetic images from spectral CT, especially at lower energies, seem to be better for the visualization of lytic lesions due to the higher contrast between the lesion and the surrounding tissue in these images.

## ACKNOWLEDGMENT

This paper and the research behind it would not have been possible without the support of Philips Healthcare company and University Hospital Brno, Department of Radiology and Nuclear Medicine.

## REFERENCES

- [1] R. Silbermann and G. D. Roodman, "Myeloma bone disease: Pathophysiology and management", *Journal of Bone Oncology*, vol. 2, no. 2, pp. 59-69, 2013, doi 10.1016/j.jbo.2013.04.001
- [2] M. H. Jagosky and S. Z. Usmani, "Extramedullary Disease in Multiple Myeloma", *Current Hematologic Malignancy Reports*, vol. 15, no. 2, pp. 62-71, 2020, doi 10.1007/s11899-020-00568-3
- [3] B. Jamet, C. Bailly, T. Carlier, C. Touzeau, A. -V. Michaud, M. Bourgeois, P. Moreau, C. Bodet-Milin, and F. Kraeber-Bodere, "Imaging of Monoclonal Gammopathy of Undetermined Significance and Smoldering Multiple Myeloma", *Cancers*, vol. 12, no. 2, 2020 doi 10.3390/cancers12020486
- [4] B. R. Madhira, V. M. Konala, S. Adapa, S. Naramala, P. M. Ravella, K. Parikh, and T. C. Gentile, "Recent Advances in the Management of Smoldering Multiple Myeloma", *World Journal of Oncology*, vol. 11, no. 2, pp. 45-54, 2020, doi 10.14740/wjon1245
- [5] J. Hillengass, S. Usmani, S. V. Rajkumar, B. G. M. Durie, M. -V. Mateos, S. Lonial, C. Joao, K. C. Anderson, R. García-Sanz, E. Riva, J. Du, N. van de Donk, J. G. Berdeja, E. Terpos, E. Zamagni, R. A. Kyle, J. San Miguel, H. Goldschmidt, S. Giralt, S. Kumar, N. Raje, H. Ludwig, E. Ocio, R. Schots, H. Einsele, F. Schjesvold, W. -M. Chen, N. Abildgaard, B. C. Lipe, D. Dytfeld, B. M. Wirk, M. Drake, M. Cavo, J. J. Lahuerta, and S. Lentzsch, "International myeloma working group consensus recommendations on imaging in monoclonal plasma cell disorders", *The Lancet Oncology*, vol. 20, no. 6, pp. e302-e312, 2019, doi 10.1016/S1470-2045(19)30309-2
- [6] F. I. Baffour, K. N. Glazebrook, S. K. Kumar, and S. M. Broski, "Role of imaging in multiple myeloma", *American Journal of Hematology*, vol. 95, no. 8, pp. 966-977, 2020, doi 10.1002/ajh.25846
- [7] A. G. Ormond Filho, B. C. Carneiro, D. Pastore, I. P. Silva, S. R. Yamashita, F. D. Consolo, V. T. M. Hungria, A. F. Sandes, E. G. Rizzatti, and M. A. C. Nico, "Whole-Body Imaging of Multiple Myeloma: Diagnostic Criteria", *RadioGraphics*, vol. 39, no. 4, pp. 1077-1097, 2019, doi 10.1148/rg.2019180096
- [8] E. Zamagni, P. Tacchetti, and M. Cavo, "Imaging in multiple myeloma: How? When?", *Blood*, vol. 133, no. 7, pp. 644-651, Feb. 2019, doi 10.1182/blood-2018-08-825356
- [9] F. Cretti and G. Perugini, "Patient dose evaluation for the whole-body low-dose multidetector CT (WBLDMDCT) skeleton study in multiple myeloma (MM)", *La radiologia medica*, vol. 121, no. 2, pp. 93-105, 2016, doi 10.1007/s11547-015-0573-6
- [10] R. Hemke, K. Yang, J. Hussein, M. A. Bredella, and F. J. Simeone, "Organ dose and total effective dose of whole-body CT in multiple myeloma patients", *Skeletal Radiology*, vol. 49, no. 4, pp. 549-554, 2020, doi 10.1007/s00256-019-03292-z
- [11] S. Zensen, D. Bos, M. Opitz, J. Haubold, M. Forsting, N. Guberina, and A. Wetter, "Radiation exposure and establishment of diagnostic reference levels of whole-body low-dose CT for the assessment of multiple myeloma with second- and third-generation dual-source CT", *Acta Radiologica*, vol. 63, no. 4, pp. 527-535, 2022, doi 10.1177/02841851211003287
- [12] S. Suntharalingam, C. Mikat, A. Wetter, N. Guberina, A. Salem, P. Heil, M. Forsting, and K. Nassenstein, "Whole-body ultra-low dose CT using spectral shaping for detection of osteolytic lesion in multiple myeloma", *European Radiology*, vol. 28, no. 6, pp. 2273-2280, 2018, doi 10.1007/s00330-017-5243-8
- [13] R. Forghani, B. De Man, and R. Gupta, "Dual-Energy Computed Tomography", *Neuroimaging Clinics of North America*, vol. 27, no. 3, pp. 371-384, 2017 doi 10.1016/j.nic.2017.03.002

# Analyses of apedal locomotion systems based on ferroelastomers

S. Kouakou<sup>1</sup>, V. Böhm<sup>2</sup>, and K. Zimmermann<sup>3</sup>

<sup>1</sup>Brno University of Technology

<sup>2&3</sup>Ilmenau University of Technology

E-mail: [244245@vut.cz](mailto:244245@vut.cz), [valter.boehm@tu-ilmenau.de](mailto:valter.boehm@tu-ilmenau.de), [Klaus.Zimmermann@tu-ilmenau.de](mailto:Klaus.Zimmermann@tu-ilmenau.de)

**Abstract**—In this paper, the movement behavior of amoeboid locomotion system is investigated and the theoretical proof of the locomotion of the system is provided with the finite element method. It is shown that not only the speed of locomotion but also its direction can be influenced by the drive frequency. Depending on the drive frequency, a movement from the home position and a subsequent movement in opposite directions can be achieved. In addition, high speeds of movement can be achieved in a limited frequency range.

**Keywords**— Legless Locomotion System, Amoeboid Robot, Technical Mechanics

## 1. INTRODUCTION

The development of biologically inspired movement systems is a current research topic. Conventional locomotion systems, which are predominantly based on forms of movement with feet and wheels, are not /or only suitable to a limited extent in many applications; such as the possibility of locomotion in different external environmental conditions. The increased mobility desired for future robotic systems requires in many areas, the use of new and unconventional principles of movement that have been already proven in nature and also adapted to environmental conditions. In recent times, biological locomotion methods have therefore been implemented more and more in technical systems. The aim is not to reproduce the highly complex biological system in detail, but rather to adopt fundamental principles with a strong abstraction. A traditional anchored research focus at the Department of Mechanic Engineering at the Technical University of Ilmenau is based on the development of bionically inspired legless locomotion systems. A current topic is the investigation of amoeboid locomotion [1].

The principle of amoeboid locomotion essentially offers two significant advantages. Firstly, the possibility of combining the functions of locomotion and manipulation, and secondly, the possibility of locomotion in restricted territories due to its contour-variable and flexible shape. Motivated by these two fascinating properties, this paper investigates the amoeboid locomotion system of robots based on magneto-sensitive elastomers body. The change of the body shape of some of these robots is done as a single unit [2]; however, it depends on self-configuration [3, 4] for other robots. The self-configuring robots are particularly complex and extremely sensitive to errors resulting from the more complicated structure. The Robots based on the shape change of a single unit are only at the design stage or their development has been abandoned due to various drawbacks.

## 2. METHOD AND EQUATIONS

The first approaches to describing the movement of unicellular organisms are already known from the 18th century. However, the microscopic techniques of the time could not yet explain how locomotion takes place. Thanks to technical progress, amoeboid locomotion can be explained macroscopically. The different stages of locomotion include the formation of pseudopodia, followed by attachment to the base via certain predetermined focal contacts of the cell membrane. After that, a certain cell body regions contracts and the final detachment of some focal contacts in the posterior region occurs [5] as shown in the **Figure 1: F0**. The transition between the different stages is smooth and the impression of a cell sliding on a surface is created. The explanation of the training principle of amoeboid locomotion is still not clear today. The pressure flow hypothesis and the frontal zone contraction hypothesis [6] are the two most common theories in biology. The pressure flow hypothesis is based on the perceptible contraction

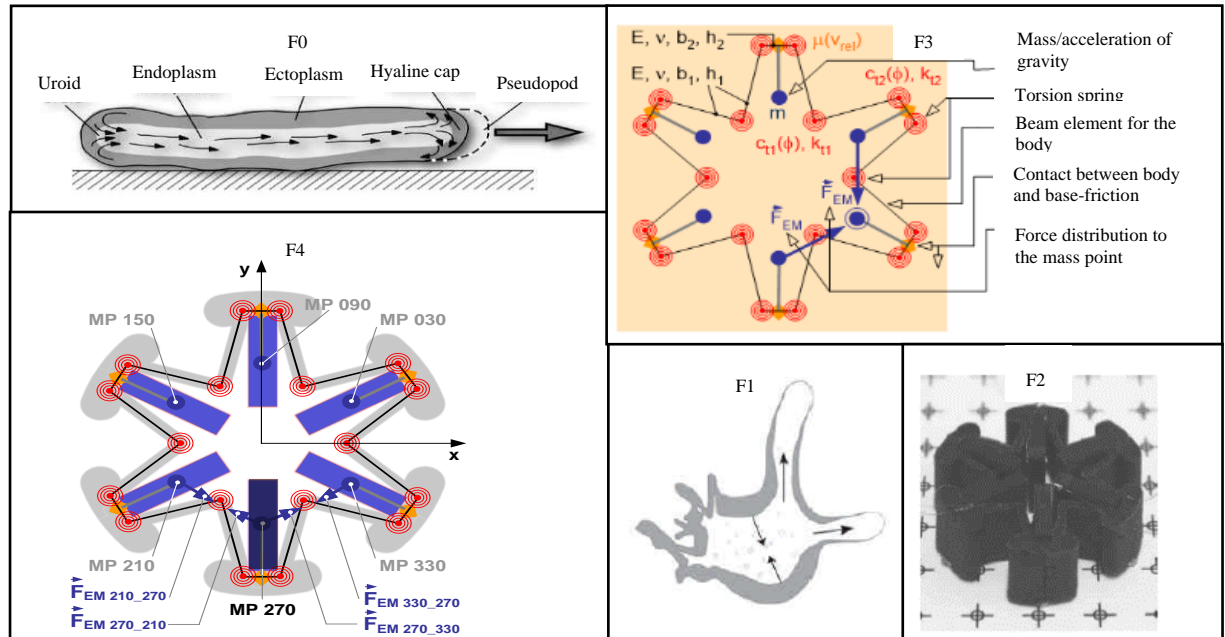


of the back of the cell. As the amoeba moves, the repetitive folding of the membrane at the posterior end in the uroid region can be clearly observed as presented in **Figure 1: F1**. The front zone contraction hypothesis assumes that the cell is pulled forward by the sol-gel transformation in the front zone. From the biological model, the following options occur in the realisation of the technical locomotion system:

- I: Fluid flow (pressure gradient) and change in compliance of a variable contour body
- II: Change in compliance of a body with variable contour subjected to internal pressure
- III: locally varied deformation of a body with variable contour

Between the three possibilities from I to III, there is an increasing abstraction and therefore a simpler technical implementation. At the beginning of the developments in the field, the feasibility of implementation belonging to group III was considered.

The objective was that the system should move in the plane. The friction in the contact area between the moving system and the environment (base) must be symmetrical. Due to the simple power supply, the principle of an electric drive was chosen from the start. In order to ensure the simplicity of the structure, electromagnets integrated in a flexible body were chosen as the drives. The material used for the body is a magnetic ferroelastomer. The electromagnetic driving force is generated when the system is activated between the excited and non-excited coils as well as between the excited coils and the surrounding body. This driving force generates a locally variable deformation of the flexible body which causes the adhesion limit to be exceeded in the contact area. This means that the displacement of the system is caused by a dynamic and periodic driving force. The system shown in **Figure 1: F2** was developed on the basis of these considerations. It consists of a symmetrical ferroelastomer body and 6 coils of the size  $(11.5 \times 11.5 \times 3 \text{ mm})$  arranged at  $60^\circ$  to each other and glued into the body. The large number of coils ensures that the relative position of the coils is close enough to each other, which increases the electromagnetic forces and thus the energy efficiency. It can be expected that the direction of movement of the system can be determined by the choice between the coils during activation.



**Figure 1:** F0: Locomotion of an amoeba; F1: Contraction of the body of an amoeba [7]; F2: real system; F3: Modelling with the element types used in ANSYS® [8]; F4: Drive activated with 1 coil.

Transient dynamic analyses are mainly performed to describe the behaviour of the system. This method ensures the best possible dynamic description in the transient phase of the system. The iterative problem-solving procedure due to non-linearities as well as the fine time resolution in transient dynamic calculations require a considerable amount of computational time. Therefore, a 3D representation of the system is not used in the studies. The system is represented geometrically using beam elements, spring-damper elements and mass point elements, as shown in **Figure 1: F3**. In the model, the coils are considered as mass points. Their connection to the spring body is done by means of beam elements, which are assumed to be rigid. The spring body is connected to the beam elements in the region of large

bends and in the region of transitions of different cross sections with discrete spring-damper elements.

In order to apply the force load, information on the relative position of the coils to each other is required for each calculation step (characterised by "i") during the activation of the coils. For this purpose, position vectors are formed. For each mass point, two position vectors with X and Y components are formed for each step time, with which the direction and distance of adjacent mass points are described. In **Figure 1: F4**, one coil (MP270) is activated. It is also possible to activate two coils at the same time.

Calculation of the components of the position vectors:

$$\mathbf{R}_{\text{Pmsp}_X \text{Sppm}} = (\mathbf{U}_{\text{XiPm}} + \mathbf{Pos}_{\text{pmX}}) - (\mathbf{U}_{\text{XiSp}} + \mathbf{Pos}_{\text{SpX}}), \quad (1)$$

$$\mathbf{R}_{\text{Pmsp}_Y \text{Sppm}} = (\mathbf{U}_{\text{YiPm}} + \mathbf{Pos}_{\text{pmY}}) - (\mathbf{U}_{\text{YiSp}} + \mathbf{Pos}_{\text{SpY}}), \quad (2)$$

The indices Pmsp is related to the mass point considered and Sppm to the adjacent mass point.

The distance between neighbouring mass points ( $\text{Abst}_{\text{Pm}_\text{Splor}}$ ) is calculated as the following:

$$\text{Abst}_{\text{Pm}_\text{Splor}} = \sqrt{\mathbf{R}_{\text{Pmsp}_X \text{Sppm}}^2 + \mathbf{R}_{\text{Pmsp}_Y \text{Sppm}}^2}, \quad (3)$$

Based on the distance between adjacent mass points, the magnitude of the electromagnetic force is determined using the following exponential function:

$$F(\text{Abst}_{\text{Pm}_\text{Splor}}) = \text{Spz} \cdot F_0 \cdot \exp(\text{Abst}_{\text{Pm}_\text{Splor}} \cdot \text{Dcf}), \quad (4)$$

The variable Spz is used to enable or disable the force on the mass point and  $F_0$  is the initial force.

The X and Y components of the force are calculated as follows:

$$F_X(\text{Abst}_{\text{Pm}_\text{Splor}}) = \frac{F(\text{Abst}_{\text{Pm}_\text{Splor}}) \cdot \mathbf{R}_{\text{Pmsp}_X \text{Sppm}}}{\text{Abst}_{\text{Pm}_\text{Splor}}}, \quad (5)$$

$$F_Y(\text{Abst}_{\text{Pm}_\text{Splor}}) = \frac{F(\text{Abst}_{\text{Pm}_\text{Splor}}) \cdot \mathbf{R}_{\text{Pmsp}_Y \text{Sppm}}}{\text{Abst}_{\text{Pm}_\text{Splor}}}, \quad (6)$$

This results in the counterforce components for the adjacent mass points:

$$F_X(\text{Abst}_{\text{Splor}_\text{Pm}}) = \frac{F(\text{Abst}_{\text{Splor}_\text{Pm}}) \cdot \mathbf{R}_{\text{Sppm}_X \text{Pmsp}}}{\text{Abst}_{\text{Splor}_\text{Pm}}}, \quad (7)$$

$$F_Y(\text{Abst}_{\text{Splor}_\text{Pm}}) = \frac{F(\text{Abst}_{\text{Splor}_\text{Pm}}) \cdot \mathbf{R}_{\text{Sppm}_Y \text{Pmsp}}}{\text{Abst}_{\text{Splor}_\text{Pm}}}, \quad (8)$$

To determine the resultant force on a mass point, the following equations are used:

$$F_{iZX} = F_X(\text{Abst}_{\text{Pm}_\text{Spl}}) + F_X(\text{Abst}_{\text{Pm}_\text{Spr}}), \quad (9)$$

$$F_{iZY} = F_Y(\text{Abst}_{\text{Pm}_\text{Spl}}) + F_Y(\text{Abst}_{\text{Pm}_\text{Spr}}), \quad (10)$$

$$|\vec{\mathbf{F}}_{\text{Res}}| = \sqrt{\mathbf{F}_{iZX}^2 + \mathbf{F}_{iZY}^2}, \quad (11)$$

The variables  $F_{\text{Res}}$ ,  $F_{iZX}$  and  $F_{iZY}$  are respectively the resultant force, the force for the mass point at one degree "Z" in the X and in the Y direction at the ith period. The indices Pm, Spl and Spr of these variables designate respectively the mass point at "Z" degrees, mass point at "Z-60" and "Z+60" degrees.

In the simulation, the following velocity-dependent friction law is agreed upon:

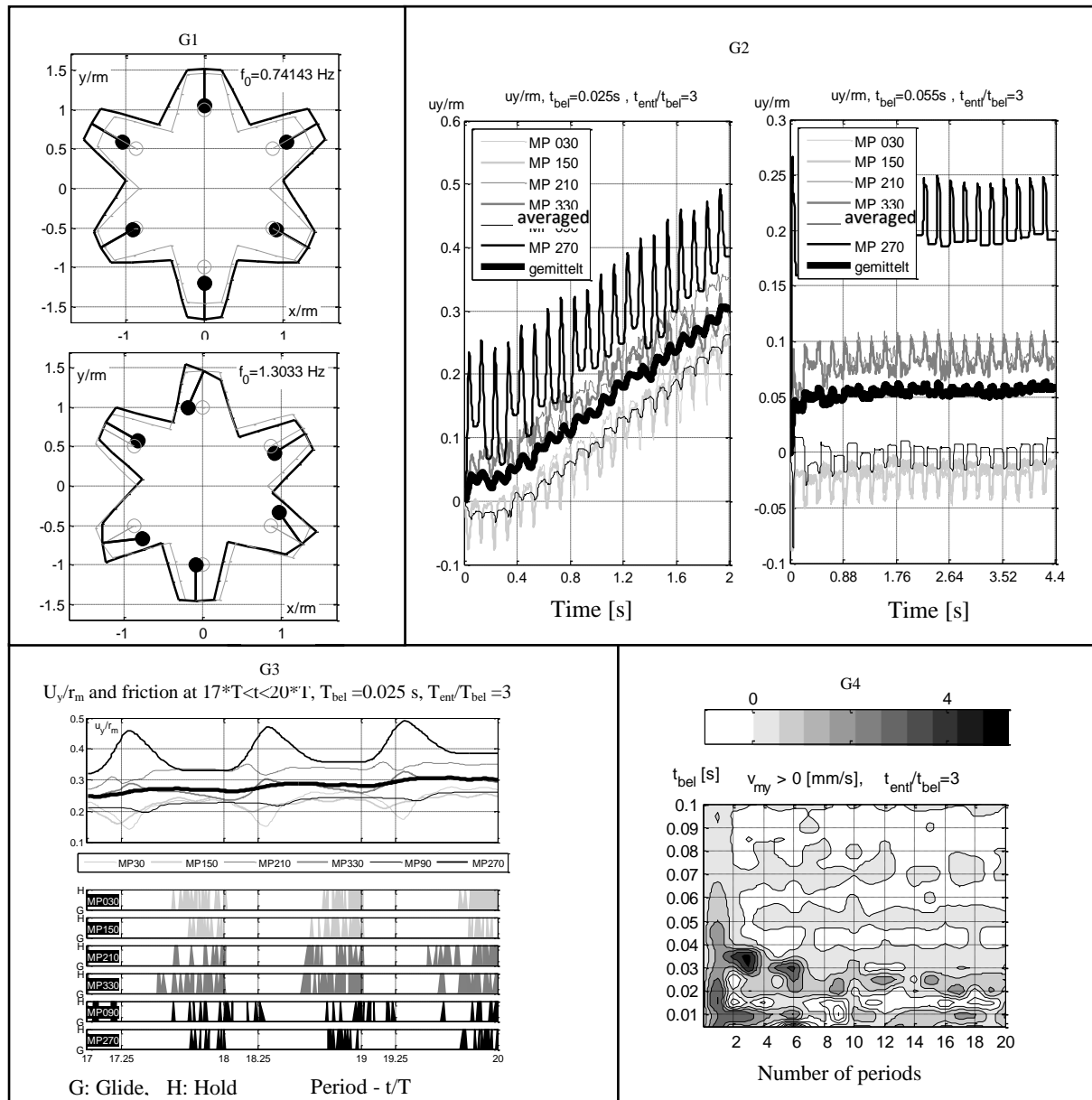
$$\mu(V_{\text{rel}}) = \mu_{\text{gleit}} + (\mu_{\text{haft}} - \mu_{\text{gleit}}) \cdot a^{-\text{DC} \cdot V_{\text{rel}}}, \quad (12)$$

The factors  $\mu_{\text{gleit}}$ ,  $\mu_{\text{haft}}$  and DC are the coefficient of sliding, friction and the decay factor respectively.

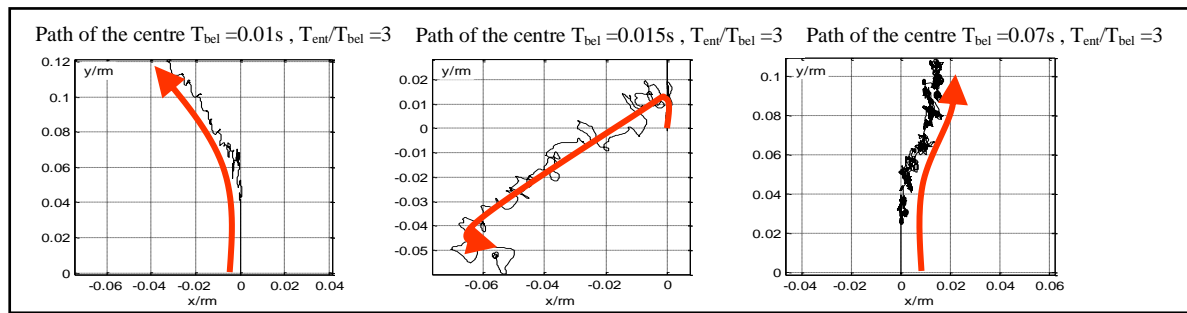
### 3. RESULTS AND DISCUSSION

The modal analysis shows that the system has several eigenfrequencies. By looking some eigenmodes on **Figure 2: G1**, both longitudinal and torsional oscillations of the system can be seen. From the point

of view of locomotion, it proves favourable to take operating frequencies close to those natural frequencies at which longitudinal oscillations occur. With the transient dynamic investigations, the system behaviour can be viewed as close to reality as possible. Due to the symmetrical design, only two different drive variants are examined. In the first variant, two coils (at  $210^\circ$  and at  $330^\circ$ ), in the second one coil (at  $270^\circ$ ) are assumed as the drive. A squared excitation is periodically applied in the form of individual forces at the corresponding mass points. Two different ratios between loading ( $T_{be}$ ) and unloading time ( $T_{ent}$ ) are used in the calculations. In the presentation of the results, the displacements ( $U_y$ ,  $y$  and  $x$ ) of the mass points are standardised over the mean radius  $r_m$  of the coil as a function of time. This means that the figures concerned do not have a unit. **Figure 2:** G2 clearly shows that all mass points of the system move forward. Furthermore, this figure reveals that the system moves quickly forwards with the loading time of 0.025s and the unloading delay equivalent to 3 times the loading time. **Figure 2:** G3, on the other hand, shows the interval during which the mass points MP030, MP090, MP150, MP210, MP270 and MP330 are either in the glide phase or the hold phase. **Figure 2:** G4 indicates the speed of the system when two coils are driven. High speed can be achieved with a loading time of 0.03s to 0.04s. **Figure 3** illustrates the path of the centre point of the system. This varies in different directions depending on the loading and unloading time. The movement is not uniaxial, so further investigation is needed to find the reasons for the numerical inaccuracy.



**Figure 2:** G1: Eigenfrequencies and eigenmodes by modal analysis; G2: influence of period duration and Theoretical proof Functionality G3: Mass point friction; G4: Speed with 2 coils.



**Figure 3:** Direction of motion of the centre point trajectory

#### 4. CONCLUSION

The possible advantages of technical systems with amoeboid motion have been deduced from biological findings. Theoretical proof of the locomotion of the system is provided. It is established that not only the speed of locomotion but also its direction can be influenced by the training frequency. Depending on the drive frequency, a movement from the home position and a subsequent movement in opposite directions can be achieved. High speeds of movement can be achieved in a limited frequency range. This paper highlights that the frequencies found in the present results analysis should be used to move the real system at high speed in the experimental phase. In further work, the locomotion system should be calculated as a coupled electromechanical system. The theoretical results should be verified quantitatively with measurements on the real system.

#### ACKNOWLEDGMENT

The author S. Kouakouo thanks the Department of Engineering Mechanics at the Ilmenau University of Technology. This work would not have been possible without the help of Professor K. Zimmermann, the head of the department, and my supervisor V. Böhm. The author is also grateful to P. Marcon and J. Mikulka for their help in writing this paper.

#### REFERENCES

- [1] K. Zimmermann and V. Böhm. 'Modellbildung, Entwurf Und Simulation amöboider Bewegungssysteme'. Unpublished Worksheet. TU Ilmenau, FG Biomechatronik. 2008
- [2] I-Ming Chen, Hsi-Shang Li, and A. Cathala. 'Design and Simulation of Amoebot-a Metamorphic Underwater Vehicle'. In Proceedings 1999 IEEE International Conference on Robotics and Automation, 1:90–95. Detroit, MI, USA: IEEE, 1999. doi:10.1109/ROBOT.1999.769936
- [3] Murata, Satoshi, and Haruhisa Kurokawa. 'Self-Reconfigurable Robots' IEEE Robotics & Automation Magazine, 2007, 71–78. doi:10.1109/MRA.2007.339607
- [4] Yu, Chih-Han, Kristina Haller, Donald Ingber, and Radhika Nagpal. 'Morpho: A Self-Deformable Modular Robot Inspired by Cellular Structure'. In 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, 3571–78. Nice: IEEE, 2008. doi:10.1109/IROS.2008.4651130
- [5] Voges, D. 'Amöboide Lokomotion – eine Zusammenfassung'. Unpublished Worksheet. TU Ilmenau, FG Biomechatronik. 2008
- [6] Hong, Dennis W., Mark Ingram, and Derek Lahr. 'Whole Skin Locomotion Inspired by Amoeboid Motility Mechanisms'. Journal of Mechanisms and Robotics 1, no. 1 (2008): 1–7. doi:10.1115/1.2976368
- [7] Rogers, Salman S., Thomas A. Waigh, and Jian R. Lu. 'Intracellular microrheology of motile Amoeba proteus' Biophysical journal 94, no. 8, (2008) 3313–3322. doi:10.1529/biophysj.107.123851
- [8] Müller, Günter, and Clemens Groth. FEM für Praktiker, I:: Grundlagen: Basiswissen und Arbeitsbeispiele. expert verlag 2000. ISBN: 3-8169-1857-3



# OBJEV SVOU PŘÍLEŽITOST VE SVĚTĚ TECHNOLOGIÍ



**ŠKODA**  
SIMPLY CLEVER



## Stáže a závěrečné práce

Studuješ a zároveň chceš získat pracovní zkušenosti? Zkus stáž ve ŠKODA AUTO. Pro talentované studenty máme připravených až 200 Kč na hodinu a s našimi odborníky můžeš konzultovat i svoji závěrečnou práci.



## Trainee program

Dokončil jsi studia a přemýšlíš kam dál? Nastup do našeho Trainee programu. V průběhu jednoho roku budeš na plný úvazek pracovat v mezinárodním týmu, poznáš různá oddělení a vycestuješ na zahraniční stáž.



## Doktorandský program

Aplikuj výsledky svého výzkumu v reálném prostředí, využívej nejnovější technologie a dej své disertační práci nový rozměr. Poskytneme ti flexibilní pracovní dobu a zajistíme 250 Kč na hodinu.



## Volné pracovní pozice

Vyber si ze široké nabídky volných pozic, staň se součástí nejprogresivnější české firmy a tvoř s námi budoucnost automobilového průmyslu.



ŠKODA AUTO Kariéra



@WeAreSKODA



ŠKODA AUTO a.s.



ŠKODA AUTO - Career

**SKODA-KARIERA.CZ**



# OBJEV SVOU PŘÍLEŽITOST VE SVĚTĚ TECHNOLOGIÍ



## Stáže a závěrečné práce

Studuješ a zároveň chceš získat pracovní zkušenosti? Zkus stáž ve ŠKODA AUTO. Pro talentované studenty máme připravených až 200 Kč na hodinu a s našimi odborníky můžeš konzultovat i svoji závěrečnou práci.



## Trainee program

Dokončil jsi studia a přemýšlíš kam dál? Nastup do našeho Trainee programu. V průběhu jednoho roku budeš na plný úvazek pracovat v mezinárodním týmu, poznáš různá oddělení a vycestuješ na zahraniční stáž.



## Doktorandský program

Apikuj výsledky svého výzkumu v reálném prostředí, využij nejnovější technologie a dej své disertační práci nový rozměr. Poskytneme ti flexibilní pracovní dobu a zajistíme 250 Kč na hodinu.



## Volné pracovní pozice

Vyber si ze široké nabídky volných pozic, staň se součástí nejprogresivnější české firmy a tvoř s námi budoucnost automobilového průmyslu.



MICHAL VILD  
ZKOUŠKY A PROTOTYPY



**ŠKODA**  
SIMPLY CLEVER



ŠKODA AUTO Kariéra



@WeAreSKODA



ŠKODA AUTO a.s.



ŠKODA AUTO - Career

**SKODA-KARIERA.CZ**



# Design of hybrid computer for modelling of dynamic systems

Miroslav Rujzl

Brno University of Technology, Czech Republic

E-mail: [xrujzl00@vut.cz](mailto:xrujzl00@vut.cz)

**Abstract**—This paper deals with the design of an analog computer with digitally controlled components for the simulation of linear and nonlinear dynamic systems or for the study of transfer functions of electronic filter or oscillator topologies. The device should allow partial adjustment using a standard computer and also record the measured results in digital form.

**Keywords**— analog computer, hybrid computer, PWL function, integrator, dynamic system

## 1. INTRODUCTION

The analog computer is the type of computer which operates with the continuous quantity, most often with voltage. It is based on the analogy between the mathematical description (most often in the form of differential equations) of a physical phenomena and the mathematical description of an electronic circuit [1]. The great advantage of analog computers is their simple implementation of complex problems that can be solved in real time. The solution obtained from an analog computer is not burdened with numerical error and the only problem remains the inherent noise of used components which is very small due to modern technology. The disadvantage of these analog systems is their absence of a method of storing the measured results and their eventual processing. This problem can be solved by a hybrid computer that combines analog parts with digital technology. For example, the measured results in analog form are digitized using an AD converter and these results are processed by available digital methods. However, the price for this method is the numerical error.

Analog and hybrid computers are devices with a long history and at first glance it may seem that this technology is becoming less important with the rise of digital technology. However, research and development departments around the world continue to use these devices. For example, paper [2] deals with the use of analog computer to study the processes between molecules. In recent years, new applications of analog computers in the optical domain have also emerged. Optical analog computer can be used for construction of neural network [3]. Similarly, in paper [4] nano-optical elements are used, whose properties enable them to solve the wave equations. This means that analog computers and analog technology in general have not had their last word.

## 2. BLOCK STRUCTURE

The hybrid computer implementation will contain separate blocks that will be connected by pins and wires. This solution allows a great versatility in which both a simulation of a 4th order differential equation and a set of four differential equations can be solved. The supply voltage is chosen to be symmetrical  $\pm 15$  V for the analog part and 5 V or 3.3 V for the digital part, which allows operation in a large state space. Some blocks are designed as purely analog, others have analog functions supplemented by digital parts that allow either the setting of individual blocks or the sensing of analog quantities for digital processing. In the following sections, the individual components are commented in more details.

### 2.1. Integrators and summaters

Four integrators represent the core of the whole hybrid computer due to the realization of continuous dynamical systems whose basic description are differential equations. Each of the integrators will be built from a commercially available AD844 element. When using the compensation clamp of this element, it behaves as a second generation positive current inverter (CCII+) with a decoupling voltage follower. The advantage of using this element is its easy implementation as an integrator with both positive and negative sign and broadband 20 MHz bandwidth. The change of the sign will be realized by a switch. The switches will also be used to select the desired time constant from four fixed implemented and one selectable which will be accessible by pins. This solution allows using the custom arbitrary value of the

time constant or plug a fractal element [5] into the integration component, which increases the versatility of the computer. The integrators will also include an initial condition setting block (IC block), which will be only one and will be implemented as a switchable block. The whole integrator can be described by the equation

$$u_{out}(t) = \pm \frac{1}{\tau} \int u_{in}(t) dt + u_{IC}(0) \quad (1)$$

where  $\tau$  is a time constant of integrator and  $u_{IC}$  is a initial condition.

The summaters will be realized fully analogue and will be a standard summing amplifier with 6 inputs. The output of the amplifier can be routed through an inverting amplifier so that both positive and negative variants can be implemented. When creating a summing amplifier from an operational amplifier, the equation for the sum voltage can be written as

$$u_{out}(t) = \pm \sum_{n=1}^N u_n(t) \quad (2)$$

where  $N = 6$  (number of inputs) and  $n$  is a index of input.

## 2.2. Multipliers and potentiometers

The proposed analog computer should mainly be used for simulations of nonlinear dynamic systems with chaotic or hyperchaotic behavior. The nonlinearities used in these systems are mostly in a polynomial form, which can be very easily implemented using standard available analog multipliers. For this application, the well-known IC AD633 was selected due to voltage range  $\pm 15$  V which covers the entire state space. This component solved the equation

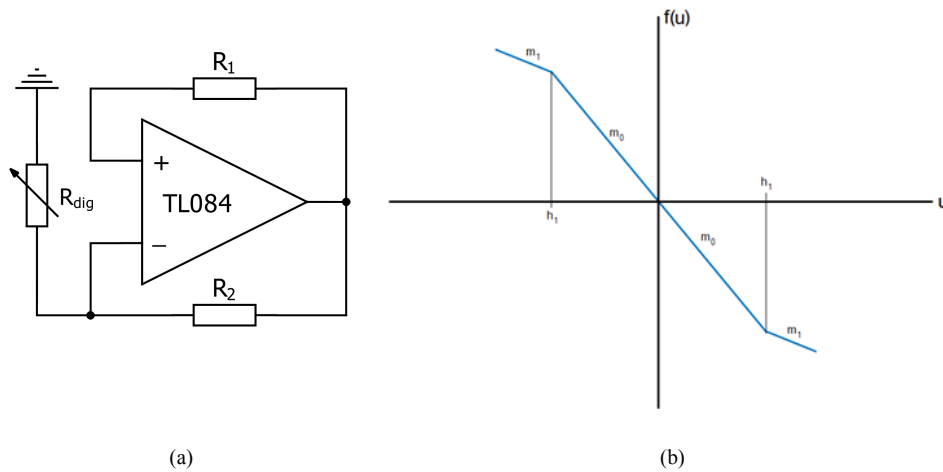
$$W = \frac{(X1 - X2)(Y1 - Y2)}{K} + Z \quad (3)$$

where  $X, Y, Z$  are input voltages ( $X, Y$  are realized as differential inputs) and  $K = 10$  is the internal division constant. This constant is must be taken into account when programming the computer and preparing the programming scheme.

Potentiometers will be implemented as voltage dividers. This configuration allows to generate dimensionless parameter values independent of the current and voltage ratios of the circuit. The range of parameters will be allowed to be 0-1, so the parameters of dynamic systems will have to be normalized to 1. In addition to standard potentiometers, two digital potentiometers will be integrated in the voltage divider configuration, which will be controllable both via internal controls and via application in PC.

## 2.3. PWL block

As mentioned above, it is necessary to create nonlinearity on an analog computer when the nonlinear dynamic systems be simulated. A frequently used type of nonlinearity is the so-called *piecewise linear function* (PWL function). As can be seen in the fig. 1b, there are linear sections with a given slope that changes at certain points. Commonly, the PWL function is formed by a circuit with operational amplifiers, but these configurations have a major disadvantage in the number of points where the slopes change because each point needs at least one operational amplifier. The planned PWL block would advantageously use a digital potentiometer in negative impedance converter circuit and AD converter. Using the control circuit, the resistor value is set to a value that corresponds to the desired slope. AD converters would monitor the voltage value across the PWL block and when the point set by the control circuit is exceeded, the digital potentiometer value changes to the new slope. The PWL block created in this way has the advantage not only of using only one operational amplifier, but also of being able to construct a large number of points and sections with a wide variation of slopes. With a sufficiently small step between points, functions that are not directly labeled as PWL can be approximated very well (e.g. realizing the  $\sin x$  function approximated by the PWL function with a fine step).



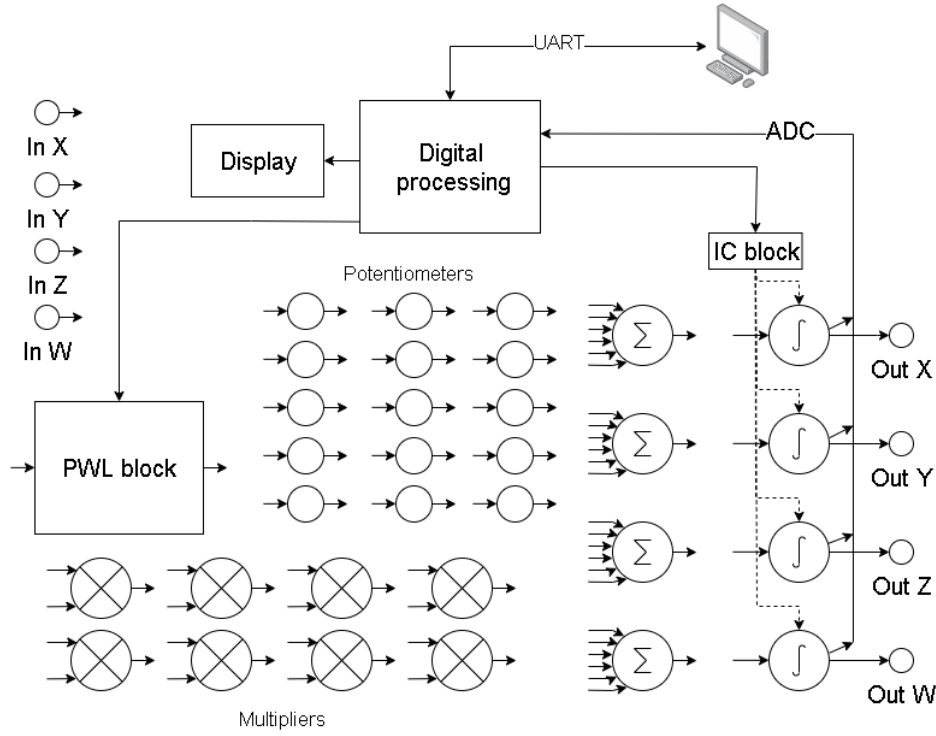
**Figure 1:** Simplified schematic of PWL block and example of PWL function

## 2.4. IC block and digital processing

The IC block will provide the initial conditions settings. This block will contain digitally controllable voltage source that can be realized by digital potentiometer in voltage divider configuration and operational amplifier follower. This structure will precharge the capacitors occurring in integrators to the required voltage level. It is not necessary to construct the IC block for each integrator separately, but it is possible to set all integrators to the available capacitor pin mode using a switch and gradually precharge all capacitors. The IC block will control by digital processing part.

The block labeled as *digital processing* will be the main controlling element of the hybrid computer. The main tasks of the control logic include setting parameters for the PWL block, operating AD converters to digitize the signal from the output nodes, controlling the display, operating the IC block to set initial conditions, controlling parametric digital potentiometers and communicating with a standard desktop PC application. For this purpose, a circuit from the STM32 series will be selected, whose characteristics and quality support from the manufacturer fully suit the application. Manufacturer support offers programming of the device using three different methods: low-level using CMSIS, low-layer (LL) and high-level High Application Layer (HAL). The use of HAL in combination with the STM CubeMX environment greatly simplifies the development of the entire program, because it is not necessary to directly set up the registers to set up the various peripherals, but this system will set up the registers automatically by using predefined functions. This way of programming saves a lot of time and allows you to concentrate more on optimizing the desired program.

The computer mode will be divided into two basic modes: offline mode, where all control of the computer takes place directly on the computer and data is displayed on the screen, and online mode, where the digital part is controlled from the desktop application. The online mode will be functionally identical to the offline mode, but will be complemented by the ability to upload measured results directly to the app. The measurement results will be written to a file (probably a CSV file), which will allow their subsequent processing, for example using Matlab.



**Figure 2:** Block diagram of the designed hybrid computer

### 3. EXAMPLE OF SIMULATION

As an example of working with an analog computer, it is possible to show the solution of a simple ordinary second order differential equation which describe a behavior of electric harmonic oscillator

$$\frac{d^2x}{dt^2} + \omega^2 x = 0 \quad (4)$$

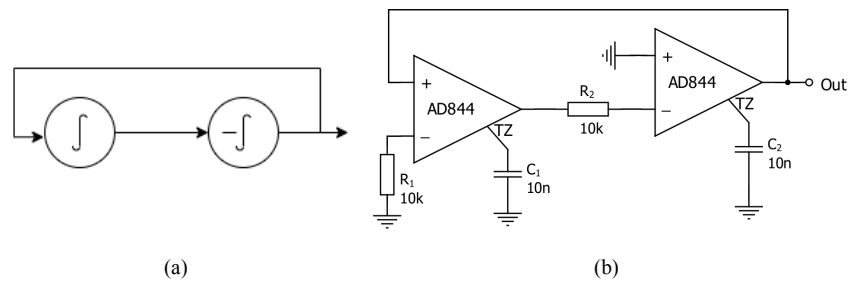
where  $x$  is a state variable which can be replaced by voltage and  $\omega$  is a constant representing the oscillation frequency value. It can be shown [6] that  $\omega$  can be calculated as

$$\omega = \sqrt{\frac{1}{\tau_1 \tau_2}} = \frac{1}{R_1 C_1 R_2 C_2} \quad (5)$$

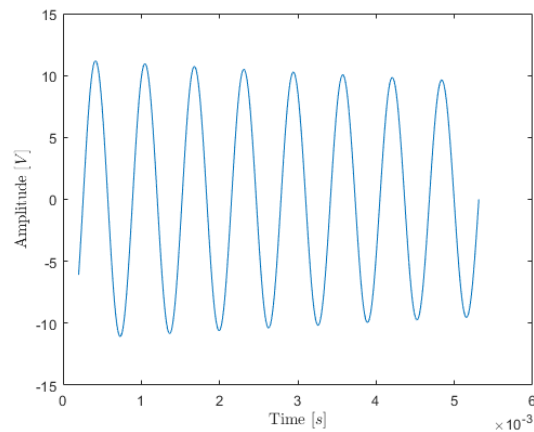
where  $\tau_1, \tau_2$  are time constants, in this case  $\tau = RC$  with elements of integrators. It is possible to achieve harmonic oscillations by introducing positive feedback. The form of the solution of the differential equation under positive feedback can be written as

$$x = e^{\alpha t} \sin \omega t \quad (6)$$

Fig.3b shows a possible implementation using two integrators consisting of AD844 circuits connected as current conveyors. The circuit in the left part implements a positive integrator with time constant  $\tau = RC = 10^4 \cdot 10^{-8} = 100 \mu s$ . It is followed by an inverting integrator with the same time constant. After the integrator, the output signal of periodic oscillations is already taken. This circuit was simulated in PSpice and result of the simulation is shown in fig. 4. Oscillation frequency  $f \approx 1.58 \text{ kHz}$  corresponds with used time constant. The slight deviation from the exact frequency value of 1.5915 kHz is due to the effect of real element used, which affects the value of time constant by its parasitic capacitances. In future use of these elements, the resistors determining the time constant will be replaced by trimmers to ensure accurate setting of the desired time constant. Note that the oscillations do not have a constant amplitude, which corresponds to the case where alpha is negative. This fact is due to the use of realistic macromodels of the integrated circuits, which include parasitic properties and decrease the energy from the dynamic system. This effect could be compensated by adding an amplifier with automatic amplitude stabilization, however, it is not necessary to implement an amplifier to see and understand the problem.



**Figure 3:** Realization of differential equation: a) programming scheme, b) circuit implementation



**Figure 4:** Output voltage waveform (solution of differential equation) simulated in PSpice

#### 4. CONCLUSION

This paper briefly describes the design of hybrid computer which can be used for modelling of linear or nonlinear dynamic system. The presented computer should cover a large number of applications. Thanks to the implementation of a digitally controlled PWL block with a fine step, even functions of a more complex character that would normally be difficult to implement can be approximated very accurately. A great advantage will also be the ability to set initial conditions from the digital part of the circuit and the ability to record results using AD converters. A simple example of a harmonic oscillator was used to show how dynamic systems can be implemented in electronic circuits and used to find solutions.

#### ACKNOWLEDGMENT

This article was created as part of a specific research project FEKT-S-20-6361.

#### REFERENCES

- [1] M. Bobek, J. Haška, I. Serba and M. Lukeš, *Analogové počítače*. Prague: SNTL Praha, 1982.
- [2] S. Köpel, A. Krause and B. Ulmann, Analog Computing for Molecular Dynamics, *arXiv preprint arXiv:2107.06283*, 2021.
- [3] J. Wu, X. Lin, Y. Guo, J. Liu, L. Fang, S. Jiao, and Q. Dai, Analog Optical Computing for Artificial Intelligence, *Engineering*, 2021.
- [4] M. Miscuglio, Y. Gui, X. Ma et. al., Approximate analog computing with metatronic circuits. *Commun Phys* 4, 196, 2021, <https://doi.org/10.1038/s42005-021-00683-4>
- [5] A. Tepljakov, E. Petlenkov and J. Belikov, Efficient analog implementations of fractional-order controllers, *Proceedings of the 14th International Carpathian Control Conference (ICCC)*, 2013, pp. 377-382, doi: 10.1109/CarpathianCC.2013.6560573.
- [6] R. Lániček, *Elektronika (Obvody, součástky, děje)*. Prague: BEN Technická literatura, 2004.

# SIFT and SURF based feature extraction for the anomaly detection

S. Bilik<sup>1</sup> and K. Horak<sup>1</sup>

<sup>1</sup>Brno University of Technology, Czech Republic

E-mail: [bilik@vut.cz](mailto:bilik@vut.cz), [horak@vut.cz](mailto:horak@vut.cz)

**Abstract**—In this paper, we suggest a way to use SIFT and SURF algorithms to extract the image features for anomaly detection. We use those feature vectors to train various classifiers on a real-world dataset in the semi-supervised (with a small number of faulty samples) manner with a large number of classifiers and in the one-class (with no faulty samples) manner using the SVDD and SVM classifier. We prove, that the SIFT and SURF algorithms could be used as feature extractors, that they could be used to train a semi-supervised and one-class classifier with an accuracy around 89% and that the performance of the one-class classifier could be comparable to the semi-supervised one. We also made our dataset and source code publicly available.

**Keywords**—Anomaly detection, Object descriptors, Machine Learning, SIFT, SURF

## 1. INTRODUCTION

The anomaly detection (AD) techniques are state-of-the-art methods for outlier detection in industrial, or medical diagnosis tasks. In comparison with the classification techniques, they are more robust to the unknown samples and they usually require only a small amount of anomalous data to train. This could be advantageous if the anomalous data is difficult to obtain.

Those algorithms usually consist of a pre-trained convolutional neural network (CNN) to extract image features followed by various one-class classifiers (OCC) to make decision on the input data. CNNs are fast and powerful state-of-the-art methods with a high generalization ability of the input samples. On the other hand, their decision analysis is often complicated and not as clear as the traditional computer vision methods. For these reasons, we propose an experiment, where we use SIFT and SURF algorithms to extract image features in order to process them with selected classifiers.

## 2. RELATED RESEARCH

Feature points detectors such as SIFT [1], or the similar and faster SURF [2], are widely used in the area of image matching, or image recognition. These methods search for significant points (usually corners) in the input image and they are invariant to geometric transformations, to light conditions, or image noise. Every feature point is uniquely defined by its high dimensional feature vector, which describes its surroundings.

As stated above, these methods are also used for object recognition. In its simplest form, it consists of comparing the obtained feature points from the input image with the database of descriptors and finding its nearest neighbour with numerous approaches to the efficient search, indexing and suppressing of false matches. Nevertheless, these often fail on obtaining generic image features and therefore are not so suitable for detecting general object categories [1].

An example of the application of the feature point based object detection in the traffic domain is shown for example in [3], or in [4]. In [3], the authors use various feature detectors as Harris, FAST, or SURF to solve the Licence Plate Recognition (LPR) problem. After application of the selected detector to the input image, the licence plate is found based on its known dimensions, known orientation of the car and feature points locations. The latter article [4] uses more similar approach to the one described in [1], where the feature points obtained from the input image are classified as potential corners using a pre-trained model as LDA, SVM, or a decision tree.



A number of articles also use the feature points detectors for the AD task - such as [5], where the authors use SURF algorithm as a part of the processing pipeline, or [6], [7] and [8], which are based on the SIFT algorithm. The authors of [6] introduce a method to recognize epileptic seizures based on the EEG signal processing. The signal is reshaped into 2D space and then is transformed to the frequency domain using FFT. The SURF algorithm is applied to the amplitude spectrum and the classification itself is performed via statistical analysis of the SURF points in time and frequency domains. The results prove a promising potential of this approach, especially in the frequency domain.

In the paper [7], the authors use SIFT descriptors to track people in the video surveillance frames and to estimate their motion in order to recognize potentially dangerous behavior. Unfortunately, the proposed methods are not described very well. Authors of the [8] present a complex algorithm for the vehicle underside inspection based on the SIFT descriptors. Captured images are compared to the frames within the specification from the database using standard feature matching techniques and possible defects are detected using the correlation of those two images after their matching. From the paper is not clear how extensive the dataset was and whether the proposed algorithm was tested in-field.

### 3. MATERIALS AND METHODS

#### 3.1. Dataset description

For this experiment, we used our own *Cookie Dataset* described in the previous paper [9]. This dataset is designed for anomaly detection task and it captures 1225 original samples in four classes of the *Tarallini* biscuits as can be seen in fig. 1. The whole dataset was augmented by three rotations of 90° to the 4900 samples in total and cropped to the size of bounding boxes. Its original structure is as follows:

- No defect (474 captions)
- Defect: not complete (465 captions)
- Defect: foreign object (158 captions)
- Defect: color defect (128 captions)

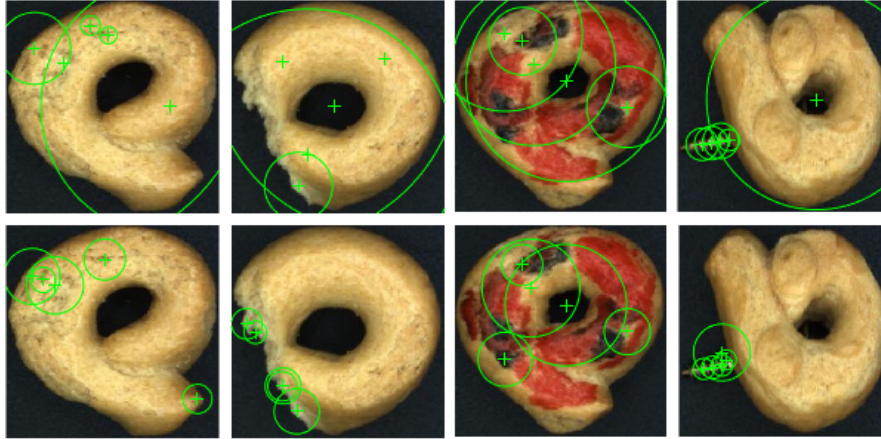
For the purpose of the anomaly detection experiment, we consider only two classes - the samples within specification (OK class) and randomly selected faulty samples from all faulty classes (NOK class). We created two training datasets - a semi-supervised one with 1000 OK samples and 50 NOK samples as the training set for the first experiment and an one-class one with the same number of OK samples and with no NOK samples. The test dataset contains 200 OK samples and 200 NOK samples and it was used for both experiments. The sample ratio of the NOK classes in the test set was set as 0,4 : 0,3 : 0,3 (not complete : foreign object : color defect).



**Figure 1:** *Cookie Dataset* class samples [9]

### 3.2. Experiment description

In this experiment, we use SIFT and SURF algorithms for the feature extraction, as a first part of the AD pipeline, together with semi-supervised and one-class methods for the samples classification. We assume, that the NOK samples have stronger and differently distributed significant points, than the OK ones because of the foreign objects presence, or due to various other defects. An example illustrating this concept is shown at fig. 2, where we can see a significant difference in the point's scale and distribution especially between the OK samples and NOK samples with color defect and strange object.



**Figure 2:** SIFT (upper half) and SURF (lower half) features of the OK (first column) and NOK samples

At first, we tried to directly use the SIFT and SURF descriptors of five strongest feature points to create a descriptor vectors of each sample, but they had not proved as strong descriptors. For this reason, we decided to use the scale and metric properties of five strongest feature points as the descriptor of  $dim(1, 10)$ . When extracting the five strongest features, a smaller number of the feature points was detected on a low number of input samples, which is possibly caused by a not distinct surface of the OK samples. In those cases, the missing features were set to zero. Input data was normalized using the z-score for the OCC experiments.

As a classifier, we compare the performance of the semi-supervised and one-class approach. In the first case, all models were trained using a semi-supervised dataset with 5% of NOK samples without further distinguishing between its classes and various classifiers from the Matlab *Classification Learner* toolbox using the five cross-validation, from these we chose three models with the best test accuracy.

One-class experiment was performed with the Support Vector Data Description (SVDD) classifier with the implementation available from [10] and with the standard Matlab Support Vector Machine (SVM) classifier. For further analysis, we trained these models with both semi-supervised and one-class datasets. In the case of OC-SVM, optimal decision boundary of the anomaly score was determined using the test dataset.

For the evaluation and comparison of the experiment results, we decided to use the test accuracy and the test AUC (area under the ROC curve) metric, which tells us more precisely about the quality of classification together with the false and true positive rates than the test accuracy. The higher the AUC is, the steeper the ROC curve is and the better the classification results are.

The whole experiment was implemented in Matlab 2021b using our scripts and the Matlab *Classification Learner* toolbox. For the dataset parser, a Python script was used, which is available together with the Matlab code at [11].

### 3.3. Experiment results

An overview of the model with the highest test accuracy of the semi-supervised experiment is shown in table I. During the training phase, almost all examined models reached an accuracy over 95% and a high training AUC. Those metrics significantly decreased during the testing, but the results still seem promising and show, that even a small number of a NOK samples is sufficient to train a classifier. The SURF feature extractor outperforms the SIFT one in both inference time and accuracy results, where it has approximately 10% higher accuracy and three-times shorter inference time, than the SIFT. Performance of the best trained models is similar in both cases.

**Table I:** Results of the semi-supervised experiment

Feature extractor	Model	Test accuracy [%]	AUC - Test
SIFT	Logistic Regression	77,2	0.86
	Coarse Tree	77,0	0.83
	One-layered NN	75,5	0.83
SURF	Naive Bayes (Gaussian)	89,0	0.92
	Subspace kNN	88,8	0.90
	RUSBoosted Trees	88,0	0.91

Three models with the best test accuracy were selected for each feature extractor.

A comparison of the models trained during the one-class experiment is shown in table II. Similarly as in the previous case, the metrics showed very good results during the training, but they significantly decreased during the testing phase. Both classifiers show only slightly better results for the semi-supervised training and the SURF feature extractor clearly outperforms the SIFT in both cases. The SVDD classifier didn't prove to perform well in this case and the results of the OC-SVM algorithm with the SURF feature extractor are comparable to the best results achieved with the semi-supervised experiment.

**Table II:** Results of the one-class experiment

Feature extractor	Model	Test accuracy [%]	AUC - Test
SIFT	One-class SVDD	60.0	0.60
	One-class SVM	75,7	0.56
	Semi-supervised SVDD	63,7	0.63
	Semi-supervised SVM	79,2	0.60
SURF	One-class SVDD	56.0	0.80
	One-class SVM	87,5	0.92
	Semi-supervised SVDD	60,0	0.79
	Semi-supervised SVM	89,5	0.90

SVM anomaly score was calculated using both OK and NOK samples.

## 4. DISCUSSION

The results of both experiments show, that SIFT and SURF algorithms could be used as a feature extractors for anomaly detection in both semi-supervised and one-class manner. In the semi-supervised experiment, the SIFT algorithm was clearly outperformed by the SURF in test accuracy, AUC and inference time. Further experiments could be performed to evaluate the minimal number of the NOK samples required in the training dataset to successfully train the model. The best trained models were the Logistic Regression for the SIFT algorithm with test accuracy of 77,2% and the Naive Bayes for the SURF algorithm with test accuracy of 89,0%.

In the one-class experiment, the best results were achieved with the SURF algorithm together with the SVM classifiers reaching the test accuracy of 87,5%. For the SVDD classifiers, the SIFT algorithm proved to be a better feature extractor than SURF. The results between the training on the semi-supervised and one-class dataset are not such significant and we can assert, that one-class learning with the SIFT and SURF features is possible. The OCC results achieved with the SURF feature extractor and SVM classifier are comparable to the results of the semi-supervised experiment.

## 5. CONCLUSION

In this paper, we used the SIFT and SURF algorithms as feature extractors for anomaly detection. For the evaluation of the experiment, we compared a semi-supervised learning approach with a small number of a faulty samples and a one-class approach with no faulty samples used during the classifier training. The results clearly show the potential of this approach, when we reached a test accuracy of around 89% and AUC around 0.90 with the both semi-supervised and one-class learning. The results of the semi-supervised and one-class learning are comparable, so we can say that the semi-supervised learning has not brought a significant improvement in the second part of experiment, although it could be used to train a conventional classifier as was shown in the first part of the experiment.

In this paper, we prove, that the SIFT and SURF algorithms designed for detection of the significant points in an image could be successfully used instead of the convolutional neural networks for the feature extraction and we suggest a way to build a descriptor vector from the obtained parameters. To the best of our knowledge, we are the first to use the SIFT and SURF feature extractors to directly create a descriptor vectors and to perform the anomaly detection task in this way.

As a future development, we plan to test our method on another anomaly detection datasets and to test it with the preprocessing algorithms based on the convolutional and variational autoencoders presented in our previous work [9]. We also plan on further developing the proposed feature extraction method in order to use more parameters given by the SIFT and SURF algorithms and also to test it with more one-class classifiers.

We made our code and dataset publicly available at [11] to allow other researchers to quickly reproduce the achieved results and to test our proposed method on their own dataset.

## ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-20-6205 -"Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0" financially supported by the Internal science fund of Brno University of Technology.

## REFERENCES

- [1] David G. Lowe, "Distinctive image features from scale-invariant keypoints", *International journal of computer vision*, vol. 60, no. 2, p. 91-110, 2004, doi: <https://doi.org/10.1023/B:VISI.0000029664.99615.94>.
- [2] H. Bay, T. Tuytelaars, L.V. Gool, "Surf: Speeded up robust features", *European conference on computer vision*, p. 404-417, 2006, doi: [https://doi.org/10.1007/11744023\\_32](https://doi.org/10.1007/11744023_32).

- [3] K. Horak, J. Klecka, P. Novacek, "License plate detection using point of interest detectors and descriptors", *39th International Conference on Telecommunications and Signal Processing (TSP)*, p. 484-488, 2016, doi: [10.1109/TSP.2016.7760926](https://doi.org/10.1109/TSP.2016.7760926).
- [4] K. Horak, J. Klecka, O. Bostik, D. Davidek, "Classification of SURF image features by selected machine learning algorithms", *2017 40th International Conference on Telecommunications and Signal Processing (TSP)*, p. 636-641, 2017, doi: [10.1109/TSP.2017.8076064](https://doi.org/10.1109/TSP.2017.8076064).
- [5] L. Gao, Y. Jiu, X. Wei, Z. Wang, W. Xing, "Anomaly Detection of Trackside Equipment Based on GPS and Image Matching", *IEEE Access*, vol. 8, p. 17346 - 17355, 2020, doi: [10.1109/ACCESS.2020.2966783](https://doi.org/10.1109/ACCESS.2020.2966783).
- [6] A. Sedik, H. M. Emara, A. Hamad, et al., "Efficient anomaly detection from medical signals and images", *Int J Speech Technol*, vol. 22, p. 739-767, 2019, doi: <https://doi.org/10.1007/s10772-019-09610-z>.
- [7] N. Ojha, A. Vaish, "Spatio-temporal anomaly detection in crowd movement using SIFT", *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, p. 646-654, 2018, doi: [10.1109/ICISC.2018.8398878](https://doi.org/10.1109/ICISC.2018.8398878).
- [8] X. Gao, Y. Wu, K. Yang, J. Li, "Vehicle bottom anomaly detection algorithm based on SIFT", *Optik*, vol. 126, no. 23, p. 3562-3566, 2015, doi: <https://doi.org/10.1016/j.ijleo.2015.08.268>.
- [9] S. Bilik, "Feature space reduction as data preprocessing for the anomaly detection", *Proceedings I of the 27th Conference STUDENT EEICT 2021*, p. 415-419, 2021, ISBN: 978-80-214-5942-7.
- [10] Kepeng Qiu, "Support Vector Data Description (SVDD)", 2022 [online]. [cit. 2022-03-12]. Available from: <https://github.com/iqiukp/SVDD-MATLAB/releases/tag/v2.1.5>
- [11] S. Bilik, Source code for the "SIFT and SURF based feature extraction for the anomaly detection" paper [online]. [cit. 2022-03-11]. Available from: <https://github.com/boortel/SIFT-and-SURF-based-AD>

# Performance comparison of a signal processing pipeline execution using CPU and GPU

A. Tomašov<sup>1</sup> and T. Horváth<sup>1</sup>

<sup>1</sup>Brno University of Technology, FEEC, Dept. of Telecommunications, Technická 12, 616 00 Brno, Czech Republic

E-mail: [tomasov@vutbr.cz](mailto:tomasov@vutbr.cz), [horvath@vut.cz](mailto:horvath@vut.cz)

**Abstract**—The paper compares the execution performance of NumPy and PyTorch mathematical libraries in embedded systems with graphics processing unit (GPU) acceleration. Both frameworks execute a signal processing pipeline from a fiber manipulation detection system, which inspects a signal from a state of polarization analyzer to enhance the security of optical fiber. The performance comparison is evaluated in the NVIDIA Jetson Nano system with 128-core Maxwell GPU. Based on the measured results, the PyTorch library executed on the GPU has performance improvement from 59 % to 84 % on different batch sizes. The results prove the real-time analysis capabilities of such a system with GPU acceleration.

**Keywords**—Fiber Optics Security, GPU, NVIDIA Jetson Nano, PyTorch, Signal Processing Acceleration

## 1. INTRODUCTION

Information security is currently a key topic in many science fields, including fiber optics as well. Protecting and authenticating transmitted data on more layers of the Open System interconnection (OSI) model significantly reduces the risk of data leak or manipulation. It also makes it difficult to perform undetected wire tapping on modern communication paths. The encryption and authentication in higher layers of the OSI model are performed using Media Access Control security (MACsec), Internet Protocol security (IPsec), and transport layer security (TLS) [1, 2, 3]. This ensures data privacy and authenticity, but it does not detect eavesdropping.

An optical fiber path provides new possibilities of eavesdropping detection by analyzing side-channel information of the path, such as reflected light, phase shifts, or polarization changes. These optic attributes vary in response to changes in a fiber optics network. One of the eavesdropping detection methods is to analyze paths using Optical Time Domain Reflectometer (OTDR) [4], which is able to evaluate path attenuation and detect optical connectors. This information can be analyzed using a deep convolutional neural network, which can detect changes caused by wiretapping in OTDR output [5]. Another detection method using a deep convolutional neural network analyzes a transmission eye diagram from an optical transmitter. Sudden unexpected changes of a signal-to-noise ratio or jitter in the diagram indicate a potential eavesdropping attack in the network [6].

Potentially wire-taped path has to be avoided until is proven secure, which creates new requirements for routing algorithms in optical networks. Optical routing algorithms are already complex and consider many network factors. The infected path can be labeled as inactive and then the network traffic has to be rerouted through the rest of the network [7, 8, 9, 10].

The Optolab research group developed a fiber protection system detecting unauthorized manipulation with transmission media. The protection system consists of a state of a polarization analyzer and a signal processing pipeline. A fiber causes polarization changes after applying pressure due to an optical effect called birefringence [11]. The pipeline is responsible for signal pre-processing and applying the anomaly detection method in the last step. This paper focuses on the performance comparison of signal pre-processing pipelines in an embedded system implemented using two distinct math libraries in Python3 [12]. Furthermore, one of the tests utilizes a graphics processing unit (GPU) to offload difficult computations and execute them in parallel. The first library is NumPy [13], which is well supported and used by many programmers. It is mostly used to perform multi-dimensional array operations and its core is written and optimized in C language with a Python3 application programming interface (API).



However, the NumPy focuses only on math operations with arrays and does not implement workload distribution. The second investigated library is PyTorch [14], which is primarily developed as a machine learning library for experimenting with deep neural networks. Consequently, it implements many math operations required by neural networks with possible acceleration in peripheral devices (e.g. GPU).

## 2. ACCELERATED SIGNAL PROCESSING PIPELINE

The signal pre-processing pipeline used in the fiber security-enhancing device consists of a Fast Fourier transform (FFT) block and a normalization block. A pipeline input is spawned by a generator, which reads data from a hard drive or receives a data through user datagram protocol (UDP) connection. The generator creates a new frame for processing when it has enough data for defined window length. Generated windows overlap with a defined portion of samples, so the generator stores the last window in an internal buffer. Hence, the windows overlap with some parts, each window is shaped by the hamming window [15]. The anomaly detection system<sup>1</sup> focuses only on single-sided magnitude spectrogram, so each window is transformed using `rfft` function a then absolute value and common logarithmic functions applied. The most normalization function for anomaly detection consists of subtracting a mean value of a frequency vector from itself. All these steps of data processing using typical functions of the NumPy [13] library are shown in Listings 1.

This paper focuses on the comparison of a CPU-based math library such as NumPy against a library with a GPU acceleration. The most suitable GPU-based option for this comparison is a library called PyTorch [14]. An alternative to this library is TensorFlow [16], but PyTorch is easy to install in Linux systems with a custom NVidia driver and has fewer package dependencies. Another possible alternative is the Caffe framework [17], but according to the GitHub repository<sup>2</sup>, it is no longer supported. The whole data processing pipeline developed using the PyTorch framework is shown in Listing 2. The main difference compared to NumPy is that you need to reference a device being used for operation execution, which for GPU-accelerated operations is `cuda:0`. The next difference is that it is necessary to move the data array to the device prior to any operation, which is shown in tensor object constructor using `device=gpu` argument.

GPU-accelerated functions benefit from executing on matrices due to parallel execution on GPU cores. Thus, the generator implements `get_batch(N)` method, which returns  $N$  time windows for processing at once. However, the process of gathering sufficient data for the whole batch adds a delay to the whole pipeline.

**Listing 1:** Implementation of data processing pipeline using NumPy.

```
1 import numpy
2 gen = ReadFile(WINLEN, OVERLAP, DATA)
3 fft_win = np.hamming(WINLEN)
4
5 for d in gen.get_batch(b):
6     # convert to frequency vector
7     d = d * fft_win
8     d = np.fft.rfft(d)
9     d = np.abs(d)
10    d = np.log(d)
11    # normalize
12    m = np.mean(d)
13    d = d - m
```

**Listing 2:** Implementation of data processing pipeline using PyTorch.

```
1 import torch, numpy as np
2 gen = ReadFile(WINLEN, OVERLAP, DATA)
3 gpu = torch.device("cuda:0" \
4     if torch.cuda.is_available() else "cpu")
5 fft_win = torch.tensor(
6     np.hamming(WINLEN), device=gpu)
7 for d in gen.get_batch(b):
8     d = torch.tensor(d, device=gpu)
9     d = torch.mul(d, fft_win)
10    d = torch.fft.rfft(d)
11    d = torch.abs(d)
12    d = torch.log(d)
13    m = torch.mean(d)
14    d = torch.sub(d, m)
```

## 3. EXPERIMENT

The experiment focuses on the comparison of execution time duration of both signal processing approaches describe in Section 2. The PyTorch pipeline is executed twice to investigate the time duration of execution on both central processing unit (CPU) and GPU. The PyTorch running on CPU is capable of distributing workload on every CPU in a system and runs computational tasks in parallel. The NumPy is not capable of this by default and parallel execution is done using other libraries (e.g. `thread` [12], `multiprocessing` [18], `numba` [19]).

<sup>1</sup>The anomaly detection system is out of the scope of this paper.

<sup>2</sup>GitHub repository can be found at <https://github.com/BVLC/caffe>.

### 3.1. Methodology

An embedded system used in this experiment is an NVIDIA Jetson Nano board with Quad-core ARM A57 of aarch64 architecture and 128-core Maxwell GPU chip<sup>3</sup>. The system has 4 GiB of RAM, which is used as a data cache by the Linux operating system.

To avoid measurement inconsistency, all signal windows are generated prior to tests and measurements. The system tries to store the newest results and deletes the oldest items from memory<sup>4</sup>. Considering the size of the input dataset, it is clear this step clear all data from memory and the measurement does not favor and disadvantage any method. There is also a program running during the whole experiment called `jetson_clocks` maintaining the maximum available frequency on both GPU and CPU cores, which increases the measurement stability and consistency as well.

After the clear cache phase, the experimental measurements come into place. The experiment evaluates the performance of NumPy functions, PyTorch functions running on CPU, and PyTorch functions running in GPU. The experiment tests batch sizes of 1, 4, 16, 64, and 256. All these tests are executed 5 times for each batch size to evaluate the variance of measured data. All information regarding the experiment together with data sizes and pipeline parameters are shown in Table I.

**Table I:** Experiment parameters and pipeline data sizes.

Parameter name	Value	Description
Window length	32768	Number of signal elements in a single window
Overlap	0.5	Size of overlapping portion of two following windows
Number of windows	116260	Number of generated windows from dataset
Data type	float32	Each signal element is represented by this type in memory
Accumulated size	14.2 GiB	Memory size of all generated windows
Test repetition	5 times	Number the same test execution

### 3.2. Measured results

Each measured result represents how long does it take to process each signal window by processing pipeline. Thus, the lower measured time means better performance. Mean measured value from tests for each library and batch size is shown in Table II. The table also contains processing time values relative to NumPy. All measurements are shown in Figure 1 using boxplots, which illustrate important statistical information, particularly the minimal value, the first quartile, mean value, the third quartile, and the maximal value respectively.

**Table II:** Mean data processing time for each batch over the computational framework and relative time duration compared to NumPy.

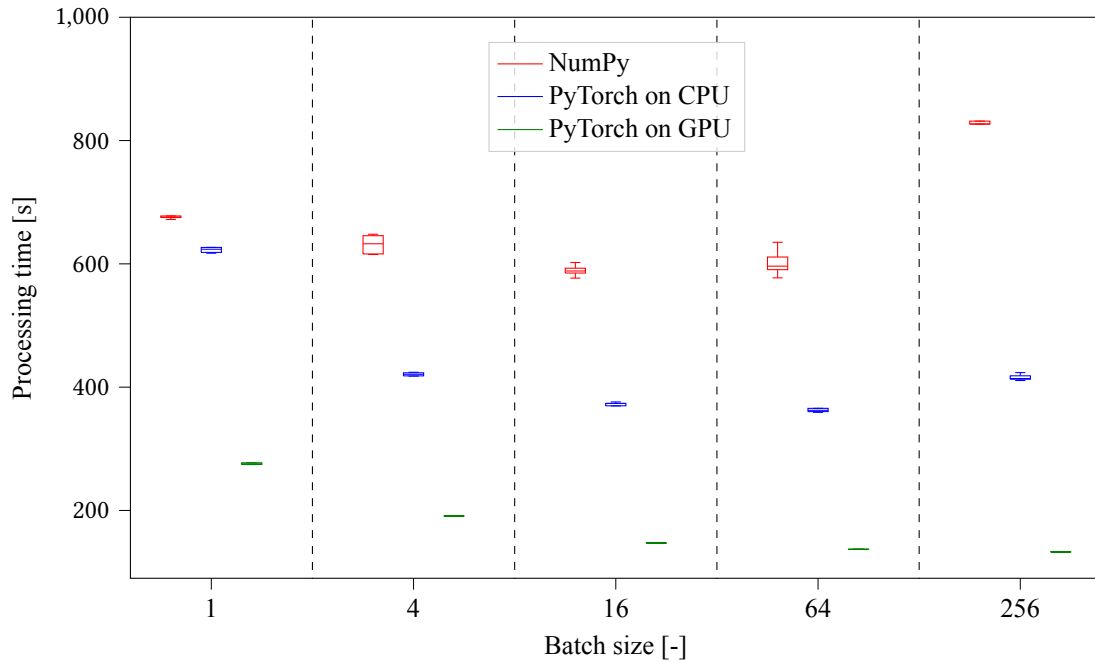
Batch size	1	4	16	64	256
Numpy [s]	677.87	631.52	589.10	602.06	826.65
CPU Torch [s]	622.59	420.90	372.51	369.81	415.93
GPU Torch [s]	275.03	190.95	147.15	137.05	132.68
<b>Relative to Numpy</b>					
CPU Torch vs NumPy [%]	91.84	66.65	63.23	61.42	50.32
GPU Torch vs NumPy [%]	40.57	30.24	24.98	22.76	16.05
<b>Relative to PyTorch</b>					
GPU vs CPU Torch [%]	44.18	45.37	39.50	37.06	31.90

<sup>3</sup>Full specification of the used system can be found here: <https://developer.nvidia.com/embedded/jetson-nano-developer-kit>.

<sup>4</sup>This algorithm is called least recently used (LRU).

PyTorch running on CPU outperforms NumPy in each test and the effectiveness rises proportionally with batch size. PyTorch utilizes all CPUs, which gives him more computational power.

Based on the measured processing time of PyTorch running on GPU, it is clear that utilizing GPU for data processing is faster with all batch sizes, even with a penalty for transferring data from the CPU into the GPU memory, which is also included in the measured time. In the best case, the PyTorch running on GPU is up to 84 % faster compared to NumPy, and 50 % compared to PyTorch on CPU.



**Figure 1:** Graphical representation of measured pipeline processing time duration over various batch sizes using boxplots, which show mean valued and variance in the measured data.

#### 4. CONCLUSION

The paper evaluates the data processing performance of NumPy as PyTorch math functions libraries in Python3. The measured performance task was a data pre-processing pipeline consisting of a magnitude spectrum of signal and data normalization function.

From all tested methods, the PyTorch running on GPU achieves the best performance, which was up to 84 % better compared to the NumPy baseline. This means that signal processing with GPU acceleration is much faster and more suitable for real-time data processing. The experiment with PyTorch running on the CPU also achieves better performance results, but not as good as the GPU experiment. The advantage of PyTorch is it can utilize all CPU cores and executes evaluated tasks in parallel.

The paper also proves that implementing the whole fiber protection system with real-time analysis in embedded devices is doable. The NVidia Jetson Nano with GPU has enough computational power to process incoming data in real-time and protect data transmitted through the fiber.

#### ACKNOWLEDGMENT

Research described in this paper was financed by the grant of the Ministry of the Interior of the Czech Republic, Program of Security Research, VI20192022146 (Distributed fiber optic sensing system for use in perimeter and line structures protection).

#### REFERENCES

- [1] "Ieee standard for local and metropolitan area networks-media access control (mac) security," *IEEE Std 802.1AE-2018 (Revision of IEEE Std 802.1AE-2006)*, pp. 1–239, 2018.
- [2] A. Keromytis, J. Ioannidis, and J. Smith, "Implementing ipsec," in *GLOBECOM 97. IEEE Global Telecommunications Conference. Conference Record*, vol. 3, pp. 1948–1952 vol.3, 1997.
- [3] S. Turner, "Transport layer security," *IEEE Internet Computing*, vol. 18, no. 6, pp. 60–63, 2014.

- [4] M. Barnoski, M. Rourke, S. Jensen, and R. Melville, "Optical time domain reflectometer," *Applied optics*, vol. 16, no. 9, pp. 2375–2379, 1977.
- [5] H. Song, Y. Li, M. Liu, K. Wang, J. Li, M. Zhang, Y. Zhao, and J. Zhang, "Experimental study of machine-learning-based detection and location of eavesdropping in end-to-end optical fiber communications," *Optical Fiber Technology*, vol. 68, p. 102669, 2022.
- [6] M. Liu, Y. Li, H. Song, Z. Tu, Y. Zhao, and J. Zhang, "Experimental demonstration of optical fiber eavesdropping detection based on deep learning," in *Asia Communications and Photonics Conference*, pp. T1H–4, Optical Society of America, 2019.
- [7] S. Yuan and D. Stewart, "Protection of optical networks against interchannel eavesdropping and jamming attacks," in *2014 International Conference on Computational Science and Computational Intelligence*, vol. 1, pp. 34–38, 2014.
- [8] Z. Zhang, J. Fu, D. Guo, and L. Zhang, "Lightpath routing for intelligent optical networks," *IEEE Network*, vol. 15, no. 4, pp. 28–35, 2001.
- [9] C. Assi, Y. Ye, A. Shami, S. Dixit, and M. Ali, "A hybrid distributed fault-management protocol for combating single-fiber failures in mesh-based dwdm optical networks," in *Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE*, vol. 3, pp. 2676–2680 vol.3, 2002.
- [10] Y. G. Huang, W. Wen, J. P. Heritage, and B. Mukherjee, "A generalized protection framework using a new link-state availability model for reliable optical networks," *J. Lightwave Technol.*, vol. 22, p. 2536, Nov 2004.
- [11] A. Smith, "Birefringence induced by bends and twists in single-mode optical fiber," *Applied optics*, vol. 19, no. 15, pp. 2606–2611, 1980.
- [12] M. Summerfield, *Programming in Python* Bearson Education Incorporated, 1900.
- [13] T. E. Oliphant, *A guide to NumPy*, vol. Trelgol Publishing USA, 2006.
- [14] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, eds.), pp. 8024–8035, Curran Associates, Inc., 2019.
- [15] P. Podder, T. Z. Khan, M. H. Khan, and M. M. Rahman, "Comparative performance analysis of hamming, hanning and blackman window," *International Journal of Computer Applications*, vol. 96, no. 18, 2014.
- [16] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. Software and documentation available from [www.tensorflow.org](http://www.tensorflow.org).
- [17] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," *arXiv preprint arXiv:1408.5093*, 2014.
- [18] J. Hunt, "Multiprocessing," in *Advanced Guide to Python 3 Programming*, pp. 363–376, Springer, 2019.
- [19] S. K. Lam, A. Pitrou, and S. Seibert, "Numba: A llvm-based python jit compiler," in *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*, pp. 1–6, 2015.

# Network Supervision via Protocol Identification in the Network

E. Holasova, K. Kuchar, and R. Fujdiak

Brno University of Technology, Faculty of Electrical Engineering and Communications,  
Dept. of Telecommunication, Technicka 12, 616 00 Brno, Czech Republic

E-mail: [eva.holasova@vut.cz](mailto:eva.holasova@vut.cz), [karel.kuchar@vut.cz](mailto:karel.kuchar@vut.cz), [fujdiak@vut.cz](mailto:fujdiak@vut.cz)

**Abstract**—This paper is focused on a comparison of ML (Machine Learning) and DNN (Deep Neural Network) techniques in protocol recognition to support network supervision for further proper handling, e.g., detection of a security incident. The DNN approach uses 11 layers and the ML approach is consisting of 28 mutually different predictive models. Both techniques were performed/compared on a freely accessible dataset containing browsing pcap files for further comparison, e.g., with other approaches. The predictive multiclass models were trained (fitted) to be capable of detecting five network protocols. Both approaches were compared by the achieved accuracy (based on testing and validating data), learning time, and predicting the time point of view. Using the ML approach, we were able to recognize the protocol with an accuracy of 1 and using DNN with an accuracy of 0.97.

**Keywords**— IT protocols, neural networks, machine learning, protocol recognition

## 1. INTRODUCTION

This paper deals with IT application protocol recognition using ML (Machine Learning) and DNN (Deep Neural Networks) to provide information for network supervision. The main aim of this paper is to provide an overview of relevant techniques for protocol distinguishing. IT field still grows and new protocols are released, but on the other hand, many of them are outdated and their use is not recommended. Many of these protocols may lack security features or contain security vulnerabilities. These vulnerabilities can be further exploited by attackers to carry out malicious activity. The supervision using ML and DNN is able to provide an overview of the used protocols, their version, the percentage of their usage, etc. These techniques can be used to distinguish network traffic so that it can be processed and further handled as needed. The protocol identification process can be also crucial when known or unknown traffic anomalies/abnormalities are detected from the security point of view. Each protocol uses a specific port number, but these values might be changed to different values to hide unwanted behavior. Moreover, protocol identification might help to provide better variability for other applications to adjust their settings to specific values based on the used protocol in the monitored network.

## 2. RELATED WORK/STATE OF THE ART

The detection area of the protocol used in the network is slightly increasing, but many existing papers use their own generated/recorded dataset. Thus, the comparison of other techniques is not possible. The papers also often do not provide sufficient information about the NN's or ML's settings. This makes the data even less comparable to other approaches. Wenbo Feng et al. [1] implemented CNN (Convolutional NN) that is possible to distinguish four network protocols (namely, HTTP, FTP, SMTP, and DNS) with an accuracy of 98 %. However, our paper deals with a publicly available dataset, and is possible to distinguish five network protocols. Ruolong Ma and Sujuan Qin presented in their paper [2] their approach based on the DNN consisting of 9 hidden layers and with an input and output layer. The paper introduces a solution able to distinguish 13 protocol types with an overall accuracy of 86 %. In comparison to [2], we present a solution/overview of possible solutions capable to distinguish a lower number of protocol types but with higher accuracy, with determined input parameters, and presented on a freely accessible dataset. Hamidreza Anvari and Paul Lu presented in their paper [3] a ML method using the KNN (k-nearest neighbors) algorithm capable of 2 protocol identification with an accuracy of 96 %. However, they do not provide input parameters specification and provide testing and evaluation of the used method on a not publicly available dataset. Abdurrahman Pektas published paper [4] focused

on the ML method with an accuracy of 99 % able to distinguish 20 protocols. However, paper [4] does not evaluate the accuracy of the proposed method on a publicly available dataset, so it is not possible to compare it with other approaches including our paper. Tian Feng et al. [5] reached using the ML method accuracy of 89 % (in IT protocol identification) using 20 input parameters. Our paper, in comparison to [5], is capable to distinguish protocol types using a lower number of input parameters and reaches an even higher value of accuracy on validation data. Wu Amei et al. [6] presented a paper focused on the protocol identification using the ML approach and reached an accuracy of 90 %. However, paper [6] does not mention the number of protocols that are able to be detected (identified), the number of input parameters is also not clearly described, and comparison is not possible due to the used private dataset. Table I compares the relevant papers with the paper we present in terms of the accuracy achieved and selected aspects. As the table shows, none of the selected papers used a publicly available dataset, which makes them difficult to compare with each other. Our paper mutually compares ML and DNN techniques to make an overview and comparison of possible methods from many aspects. Moreover, our paper uses a publicly available dataset [7] to enable further work on this article and its evaluation.

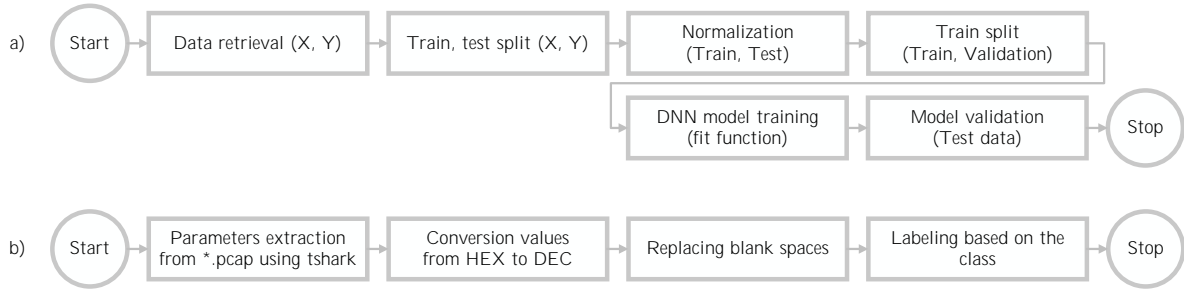
**Table I:** State of the art comparison from the selected point of view

Paper	Year	Approach	Protocol types	Input param.	Layers	Epochs	Accuracy	Dataset
[1]	2020	CNN	4	Window	6 + 2	20	0.98	Private
[2]	2018	DNN	13	?	9 + 2	20	0.86	Private
[3]	2021	ML - KNN	2	?	-	-	0.96	Private
[4]	2018	ML	20	21	-	-	0.99	Private
[5]	2020	ML - K-Means	8	20	-	-	0.89	Private
[6]	2011	ML - KNN	?	?	-	-	0.90	Private
Our	2022	ML	5	17	-	-	0.34–1.0	[7]
		DNN	5	17	9 + 2	500	0.76–0.97	

### 3. DATA PREPROCESSING

As the input data for our comparison of possible ML and DNN techniques, the data from [7] was chosen. The *browsing.pcap* data file was selected because it contains only legitimate traffic (no attacks occur), a large number of used protocols, and it is created by browsing the Internet. So the simulation of the legitimate client was recorded in this file. From this file, we have chosen (filtered from the pcap file) 5 protocols, namely, DNS (Domain Name System), HTTP (Hypertext Transfer Protocol), HTTP-jpeg, where the jpeg image was loaded using HTTP, TCP (Transmission Control Protocol), and TLS (Transport Layer Security). These created files (filtered protocols from the pcap file) do not contain each other's data. It is thus possible to perform mutual data identification over them. These are also the most used protocols in a user's normal browsing on the Internet. Figure 1 shows the process of data preprocessing and extracting from the \*.pcap file into \*.csv file. Some of the data are stored in the hexadecimal value and need to be converted into the decimal format. Moreover, the empty values need to be replaced by a placeholder value (-10). All this process was done using our created Python script. To be able to handle traffic and provide relevant IT protocol distinguishing, classification DNN and, respectively, ML was chosen. The 5 protocols (classes) that have been selected are marked 0 or 1 for each class (where 1 for a specific class indicates the occurrence of the protocol). Each class represents one protocol and only one of all classes can be labeled 1 in one record/prediction. This approach is referred to as multi-class classification. The browsing pcap taken from the dataset is consisting of 219 535 records (captured packets), from which the filtered groups were: DNS - 10 332, HTTP - 26 529, HTTP-jpeg - 24 030, TCP - 168 713, and TLS - 46 854 packets. The dataset for ML or DNN model creation and evaluation contained a total of 276 458 labeled packets (records). This created data file (dataset) was divided into two groups according to the set ratio of training and testing data. Training data are later divided into training and validation data. The training data are used for model training and validation for model validation in the process of learning (fitting). The testing data are used for final model testing on data that were not used during the training and validation phases. This process of data handling is visualized in Figure 1. Each record is consisting of 17 input parameters, taken from frame: frame.len, from IP layer: ip.len, ip.hdr\_len, ip.id, ip.flags, ip.dsfield, ip.frag\_offset, ip.ttl, ip.checksum, ip.proto, from TCP layer: tcp.len, tcp.srcport, tcp.flags, tcp.dstport, tcp.window\_size\_value, tcp.window\_size\_scalefactor, and tcp.urgent\_pointer + DNS, HTTP, HTTP-JPEG, TCP, and TLS labels.

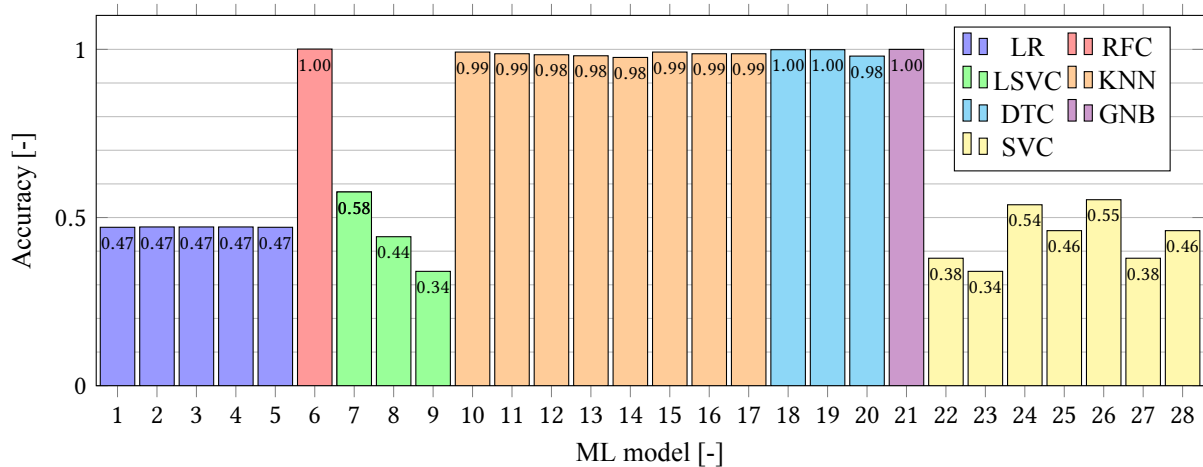




**Figure 1:** a) Data preprocessing from csv file and b) Data usage for model training and testing

#### 4. MACHINE LEARNING APPROACH

The machine learning method used the prepared multilabel dataset. After the dataset was loaded, the *min max scaler* was performed to perform input data normalization. After that, 28 predictive machine learning models were created based on 7 selected ML models, namely LR (Logistic Regression), RFC (Random Forest Classifier), LSVC (Linear Support Vector Classification), KNN, DTC (Decision Tree Classifier), GNB (Gaussian Naive Bayes), and SVC (Support Vector Classification). These models were constructed using sklearn software [8] library. Figure 2 compares the reached model accuracy based on the testing data (not previously seen while the model was trained). Each model was marked with a number in the range 1-28 to be able to distinguish individual models. The LR 1 model used the *lbfgs* solver, LR 2 model used the *liblinear* solver, LR 3 model used the *sag* solver, and LR 4 used the solver *saga*. The multiclass was set to *ovr* in LR model 1-4. The LR 5 model was set to default values. The RFC model was also set to default values. The LSVC 7 had penalty set to *l1*, loss to *squared\_hinge*, random\_state to 0, and dual parameter was set to *False*. LSVC 8 had penalty set to *l2* and random\_state to 0. LSVC 9 had penalty set also to *l2*, random\_state to 0 and loss was set to *hinge*. KNN 10, 11, 12, 13 and 14 had n\_neighbors parameter equal to 3, 5, 7, 9, and 11. The KNN 15 had weights set to *distance*, KNN 16 had algorithm set to *ball\_tree* and KNN 17 had algorithm set to *kd\_tree*. The DTC 18 had criterion set to *gini*, DTC 19 had criterion set to *entropy* and DTC 20 had splitter set to random. The GNB was implemented without changes. The SVC 22 had gamma set to *auto*, SVC 23 had kernel set to *linear*, SVC 24 had kernel set to *poly*, SVC 25 had kernel set to *rbf*, SVC 26 had kernel set to *sigmoid*, SVC 27 had kernel set to *auto*, and SVC 28 had decision\_function\_shape parameter set to *ovo*. The highest accuracy was achieved by the GNB 21 model, where the accuracy on testing data was equal to 1, followed by the DTC model where the accuracy was in the range from 0.98 to 0.99, and further by the KNN model where the reached accuracy was in the range from 0.976 to 0.992 and the RFC model with accuracy equal to 0.997 is also possible as a model capable of high protocol labeling. Each ML model gives five binary outputs (one for each class) to determine the predicted protocol based on the input data.



**Figure 2:** Machine learning models comparison from the accuracy point of view

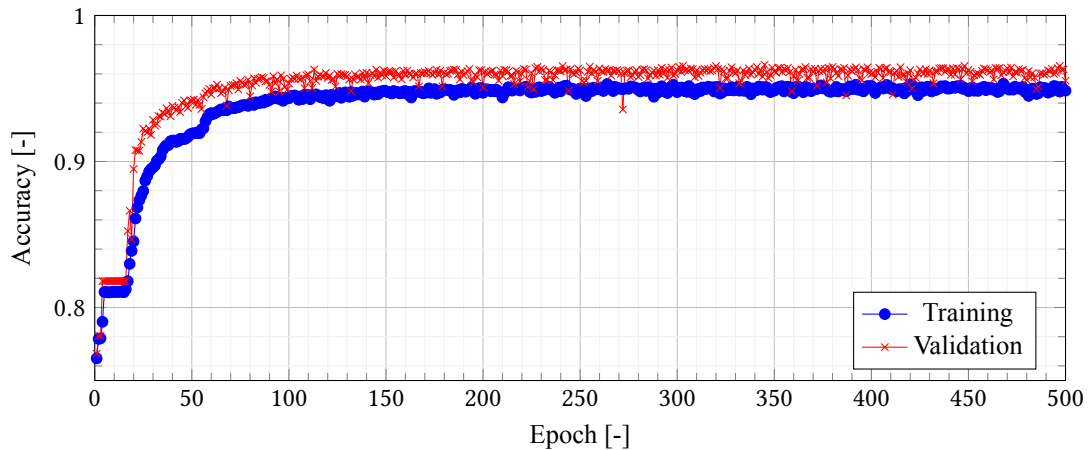
## 5. DEEP NEURAL NETWORK APPROACH

The neural network approach was implemented using Tensorflow, especially the Keras software [9] library. The loaded data from the created dataset are firstly divided into training and testing sets, where the training set is used after the end of the learning (fitting) phase. The training data are further divided into training and validation sets (in the same way as in the case of ML). Secondly, the data are normalized using *standard scaler* to ensure faster convergence against not normalized input data. These training and validation data enter into building a sequential model. The model is consisting of 9 dense hidden layers and input and output layers where 17 parameters are taken by the input layer and 5 binary outputs are provided (one for each class). Three dropout layers were implemented to prevent overfitting with a dropout with a rate equal to 0.3. Due to the binary multiclass approach, the *binary\_crossentropy* was used as the loss function, batch size was set to 10, and epoch parameter was set to 500. As the activation function the *sigmoid*, *tanh* and *relu* were selected. The structure and the setting based on experimental settings with different parameters of each layer is shown in Listing 1.

**Listing 1:** The chosen structure and settings of the deep neural network model

```
1 model.add(Dense(70, input_dim=17, activation='relu'))
2 model.add(Dense(60, activation='tanh'))
3 model.add(Dense(55, activation='sigmoid'))
4 model.add(Dense(50, activation='sigmoid'))
5 model.add(Dropout(0.3))
6 model.add(Dense(40, activation='sigmoid'))
7 model.add(Dropout(0.3))
8 model.add(Dense(10, activation='sigmoid'))
9 model.add(Dropout(0.3))
10 model.add(Dense(5, activation='sigmoid'))
```

To be able to provide all relevant information during the training (fitting) phase, the callback had to be implemented. Using this callback it is possible to take data at the end of each epoch. This was also used to store the provisional model with the highest accuracy. The taken data are visualized in Figure 3, accuracy based on the training data is marked blue, and the validation results are marked red. The highest accuracy was equal to 0.9533 in epoch 468 based on the training data and the accuracy was equal to 0.9662 in epoch 345 based on the validation data. The accuracy based on the last epoch based on the testing data was equal to 0.955.



**Figure 3:** Impact of the number of epochs on the achieved accuracy of the DNN model

## 6. MUTUAL COMPARISON OF ML AND DNN

Table II provides a comparison of ML and DNN from the reached accuracy and time requirements point of view. From the table, it is visible that DNN requires a longer time for model training (fitting), on the other hand, it requires a lower time for providing predictions. Reached accuracy is comparable for both ML and DNN approaches. As mentioned above, the accuracy based on the testing data is in the case of the DNN approach provided only for the last model (500th epoch), due to the created callback, each model with higher accuracy is stored and it is possible to use a specific DNN model.

**Table II:** Comparison of Machine Learning and Deep Neural Network

	Machine Learning			Deep Neural Network		
	min.	max.	avg.	min.	max.	avg.
No. of trained (fitted) models	28			1		
No. of predicted values	91231			91231		
Accuracy (val. data)	0.4740	1.0000	0.8313	0.7683	0.966	0.9518
Accuracy (test. data)	0.3340	1.0000	0.7152	0.9955		
Predicting time (total) [s]	0.1362	978.0690	292.9850	2.7070		
Model learning time [s]	0.4410	3398.2620	419.6270	12574.9498		
Epoch time [s]	-			22.5978	31.0769	25.1497

## 7. CONCLUSION

This article was focused mainly on the comparison of the selected ML and DNN approaches protocol identification to support network supervision. The multiclass approach was chosen to give five (five selected protocols were distinguished) binary outputs (one for each protocol) where only one class might be set to 1. As input data, the freely available dataset was chosen to ensure further work and easy comparison with other approaches. Based on the evaluation, the ML method is capable to recognize the protocol based on input data with an accuracy of 1 and DNN with an accuracy of 0.97. The results are mutually comparable. From the time needed for prediction-making point of view, the DNN provides a slightly better result, but the training time is highly demanding. Both approaches are capable of high accuracy prediction, but the result might change based on the input data and on the number of protocols.

## ACKNOWLEDGMENT

The described research is part of the grant project registered under no. TK02030013 and funded by the Technology Agency of the Czech Republic.

## REFERENCES

- [1] W. Feng, Z. Hong, L. Wu, M. Fu, Y. Li, and P. Lin, "Network protocol recognition based on convolutional neural network," *China Communications*, vol. 17, no. 4, pp. 125–139, 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/9089183/>
- [2] R. Ma and S. Qin, "Identification of unknown protocol traffic based on deep learning," *2017 3rd IEEE International Conference on Computer and Communications (ICCC)*, pp. 1195–1198, 2017. [Online]. Available: <http://ieeexplore.ieee.org/document/8322732/>
- [3] H. Anvari and P. Lu, "Machine-learned recognition of network traffic for optimization through protocol selection," *Computers*, vol. 10, no. 6, 2021. [Online]. Available: <https://www.mdpi.com/2073-431X/10/6/76>
- [4] A. Pektas, "Proposal of machine learning approach for identification of instant messaging," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 2, no. 6, pp. 97–102, 2018-06-29. [Online]. Available: <https://www.ijisae.org/IJISAE/article/view/551>
- [5] T. Feng, D. Man, and H. Fu, "Unknown protocol identification based on improved k-means algorithm," *Journal of Physics: Conference Series*, vol. 1646, no. 1, 2020-09-01. [Online]. Available: <https://iopscience.iop.org/article/10.1088/1742-6596/1646/1/012023>
- [6] W. Amei, D. Huailin, W. Qingfeng, and L. Ling, "A survey of application-level protocol identification based on machine learning," *2011 International Conference on Information Management, Innovation Management and Industrial Engineering*, pp. 201–204, 2011. [Online]. Available: <http://ieeexplore.ieee.org/document/6114688/>
- [7] A. Habibi Lashkari, G. Draper Gil, M. S. I. Mamun, and A. A. Ghorbani, "Characterization of tor traffic using time based features," in *Proceedings of the 3rd International Conference on Information Systems Security and Privacy - ICISSP, INSTICC*. SciTePress, 2017, pp. 253–262.
- [8] "Scikit-learn." [Online]. Available: <https://scikit-learn.org/stable/>
- [9] "Keras." [Online]. Available: <https://keras.io/>

# SVM Algorithm Training for DDoS on SDN Networks

Shujairi Murtadha<sup>1</sup>, and Vladislav Škorpil<sup>2</sup>

Brno University of Technology, FEEC, Czechia

E-mail: [xshuja00@vutbr.cz](mailto:xshuja00@vutbr.cz), [skorpil@feec.vutbr.cz](mailto:skorpil@feec.vutbr.cz)

**Abstract**— Despite the flexibility provided by SDN technology is also vulnerable to attacks such as DDoS attacks, Network DDoS attack is a serious threat to the Internet today because internet traffic is increasing day by day, it is difficult to distinguish between legitimate and malicious traffic. To alleviate the DDoS attack in the campus network, to mitigate this attack, propose in this paper to classify benign traffic from DDoS attack traffic by SVM of the classification algorithms based on machine learning. As the contribution of this paper is to train the SVM algorithm which has been used in the approach for the training process. Due to the complexity of the dataset, using a type of kernel called a polynomial kernel to accomplish non-linearity discriminative. The results showed that the traffic classification was with the highest accuracy 96 %.

**Keywords**— SDN, ML, SVM, RYU, DDoS

## 1. INTRODUCTION

The great development in computers and information systems coincided with a similar development in communication networks and the speed of information transmission, which increased the need to protect this information and protect the networks through which information is transmitted. There have always been attempts to find appropriate security solutions to protect networks, these solutions varied between intrusion detection systems, firewalls, and anti-malware programs, but the problem is that the rapid development of information systems and networks and the emergence of new concepts such as software-defined networks and the adoption of major companies [1]. The data centres of these concepts were accompanied by developments and a great diversity in the types of threats and risks so that the traditional methods of network protection became insufficient and unsatisfactory for the network administrators. Despite much research in this field, its workers are still looking for better results to prevent any intrusion, attack, or penetration attempt[2]. Therefore, it was necessary to find new methods or develop old methods so that firewalls and intrusion detection systems become more capable of detecting and stopping any attacks. It is no longer sufficient to detect the attack in its early stages and take the necessary measures to stop it. Rather, it has become necessary to have the ability to study the behaviour of network users and anticipate the attack before it begins. Machine learning algorithms are the best methods in this field, due to their ability to extract new information at any User interaction with the network, so that user behavior can be studied, which facilitates the process of discovering at-tacks before they start and increases the ability to discover and combat new types of attacks. But the problem is that using these types of algorithms with large networks requires an additional cost in terms of buying new devices and adding them to the network, and in terms of overloading the network to extract information about users' behavior [3].

## 2. SUPPORT VECTOR MACHINE(SVM)

Machine learning is one of the branches of artificial intelligence, which is responsible for making the computer system able to find a solution to a problem by studying a group of data (training data). Moreover, finding a certain patterns within that group make the system able to build a model that expresses the problem that has been studied [4]. Further, the high accuracy that provided by the SVM algorithm also the ability to deal with data from several dimensions, make this algorithm able to classy among the functions of the kernel algorithms, and it aims to find the best geometric intervals between rows of data.

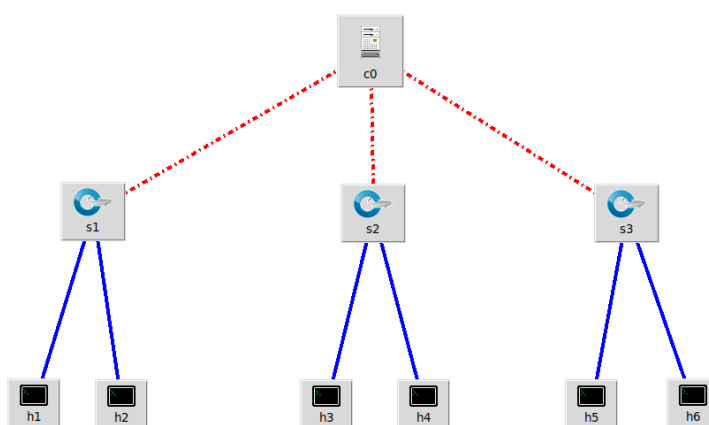
This was done by measuring the distance between each pair of data [5]. It also features the SVM algorithm, including ease of presentation and the ability to represent the results as a set of rules, ease of handling large data also, the possibility of a problem over-fitting occurring less than other algorithms.

### 3. DDOS ON SDN

on the Control Plane, the attacker sends several requests from spoofed IP addresses, which leads to the processing of many packets in the Control Plan delays or rejects the legitimate user's request [6]. The communication channels between the control plane and the data plane: The attacker may attack the connection link between the control plane and the data plane, and it may exhaust the bandwidth of the link. On the Data plan, an attacker can attack the data plane by flooding the existing flow table by flooding the existing flow table in switches and routers, this leads to an overflow attack on the flow table [7].

### 4. METHODOLOGY

Working on a computer environment HP EliteBook with RAM 8 Giga, 64-bit processor, and Linux Ubuntu operating system using open-source controller Ryu. Ryu supports various protocols. We simulate a DDoS attack on SDN using the Mininet emulator 3hping tool to generate malicious IP spoofing traffic with the rate of 450 packets per second and generates the combination TCP, UDP and ICMP. Our topology consists of some hosts, three switches, and one RYU controller (see Fig. 1). Mininet creates SDN elements such as controllers, switches, and hosts and can share them with the other networks. Ryu provides software components with well-defined API. It is easy for developers to create new network management and control applications. Benign traffic is generated from random hosts with the help of the Mgen tool with the rate of 450 packets per second and generates the combination of TCP-SYN attack, UDP flood attack and ICMP attack. Both benign and malicious traffic is counted in a table. It is written in a CSV file.

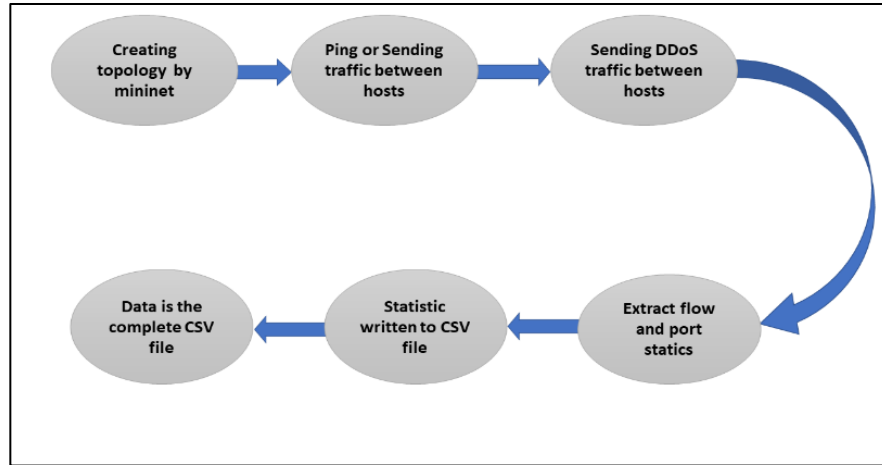


**Figure 1:** SDN Topology

### 5. IMPLEMENTATION

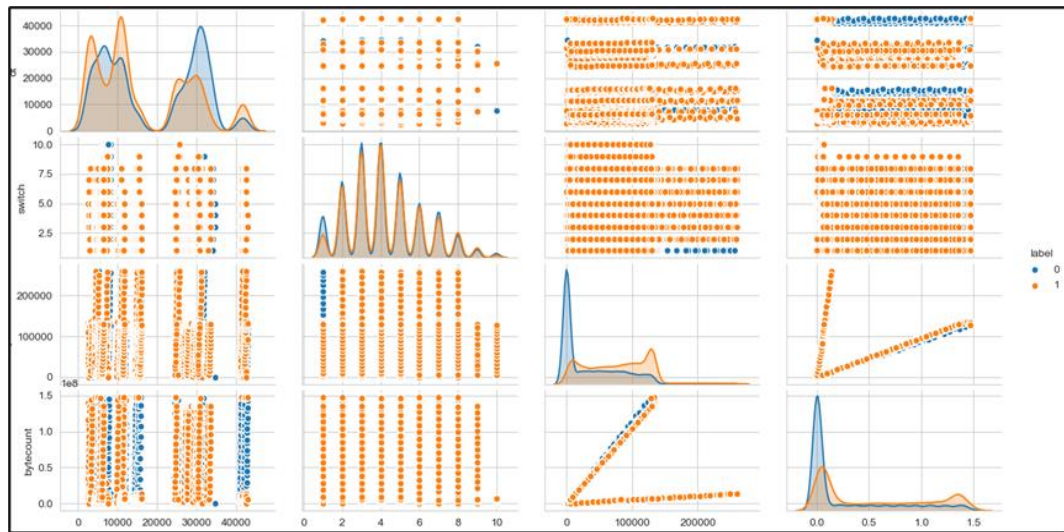
In this work was used a complex dataset consisting of 10 topologies with 1 controller RYU, 9 switches and 20 hosts, the simulation process was 250 minutes [8]. A dataset is created through a console extension of a Python application that was created with the help of the RYU API. It monitors the switches in a hierarchy and collects the stats of the various flows and ports in a regular monitoring period. It also writes the combined stats into a CSV file.

Figure 3 shows the process set for creating a set of data. The gradual stages of creating a set of data are shown. After the data set is created and commented on, the machine learning algorithm is trained and tested on a data set.



**Figure 2:** Create the dataset

Figure 3 shows the annotations on the data set automatically by applying our code, The encoding is done in such a way that the label column of the dataset is set to (0) when non-malicious traffic is turned on and for malicious traffic to (1). After annotations to the data, an algorithm is applied, the implement SVM algorithm, one of the machine learning algorithms for classification of traffic which corresponds to 1 and 0



**Figure 3:** the visualizing of the dataset used in the training process which contains the features.



Traffic category	Number of cases
Benign	63561
Malicious	40784
TCP	29436
UDP	33588
ICMP	41321

**Table 1:** Traffic category of each traffic instance

Traffic category	Benign	Malicious
ICMP	24957	16364
TCP	18897	10539
UDP	22772	10816

**Table 2:** Number of instances in the dataset

```

Train Result:n=====
Accuracy Score: 96.29%

CLASSIFICATION REPORT:n
precision    0.977700    0.940759    0.962879    0.959229    0.963264
recall      0.960987    0.965830    0.962879    0.963408    0.962879
f1-score     0.969271    0.953129    0.962879    0.961200    0.962963
support     47676.000000    30582.000000    0.962879    78258.000000    78258.000000

Confusion Matrix: n [[45816  1860]
 [ 1045 29537]]n
Test Result:n=====
Accuracy Score: 96.14%

CLASSIFICATION REPORT:n
precision    0.976919    0.938227    0.96136    0.957573    0.961787
recall      0.959207    0.964713    0.96136    0.961960    0.961360
f1-score     0.967982    0.951286    0.96136    0.959634    0.961452
support     15885.000000    10202.000000    0.96136    26087.000000    26087.000000

Confusion Matrix: n [[15237   648]
 [   360 9842]]n

```

**Figure 4:** Train and Test Result for SVM

## 6. CONCLUSION

In this work implemented a Training of machine learning algorithm SVM for traffic classification based on SDN network topology built by Mininet simulation Trafficking as table 1 shows, traffic category of each traffic instance and table 2 represents the number of instances in the dataset has been classified as malicious and benign, the performance of the system is evaluated by applying the SVM algorithm to the dataset and obtained high accuracy of Train Accuracy 96.29% and Test Accuracy 96.14% illustrated in figure 4. In the future, the plan is to extend the work and use a new algorithm to counteract cyber-attacks on the ONOS Controller.

## REFERENCES

- [1] A. N. Alhaj and N. Dutta, “Analysis of Security Attacks in SDN Network: A Comprehensive Survey,” *Contemp. Issues Commun. Cloud Big Data Anal.*, pp. 27–37, 2022.
- [2] G. S. Aujla, S. Garg, K. Kaur, and B. Sikdar, “Software Defined Internet of Everything.” Springer, 2022.
- [3] S. Prabakaran *et al.*, “Predicting Attack Pattern via Machine Learning by Exploiting Stateful Firewall as Virtual Network Function in an SDN Network,” *Sensors*, vol. 22, no. 3, p. 709, 2022.
- [4] J. P. D. Shubhabrata Datta, *Machine Learning in Industry*. Springer Cham, 2022.
- [5] Z. Gao, S.-C. Fang, J. Luo, and N. Medhin, “A kernel-free double well potential support vector machine with applications,” *Eur. J. Oper. Res.*, vol. 290, no. 1, pp. 248–262, 2021
- [6] R. C. Meena, K. Meena, A. Kumar, M. Gupta, and S. Kumar, “Status of address spoofing attack prevention techniques in software-defined networking (SDN),” *J. Discrete Math. Sci. Cryptogr.*, vol. 24, no. 8, pp. 2341–2351, 2021.
- [7] A. Pradhan and R. Mathew, “Solutions to vulnerabilities and threats in software defined networking (SDN),” *Procedia Comput. Sci.*, vol. 171, pp. 2581–2589, 2020.
- [8] “Ahuja, Nisha; Singal, Gaurav; Mukhopadhyay, Debajyoti (2020), ‘DDOS attack SDN Dataset’, Mendeley Data, V1, doi: 10.1763.”

# Multi-purpose PLC gateway for management, control and visualization

M. Ruzs<sup>1</sup>

<sup>1</sup>Brno University of Technology, Czech Republic

E-mail: [xruszm00@vut.cz](mailto:xruszm00@vut.cz)

**Abstract**—This paper presents the concept of a smart gateway that uses PLC communication technology. The smart gateway will enable monitoring and control of electricity consumption at the point of consumption behind the electricity meter. The primary purpose is the aggregation of data from measuring units, performing more computationally intensive operations with measured values, responding to non-standard events on the grid and forwarding data from end devices to the superior system. A part of this concept is also a use interface, which will enable real-time monitoring of measured quantities.

**Keywords**—PLC, gateway, energy consumption, monitoring application, non-standard events, data forwarding

## 1. INTRODUCTION

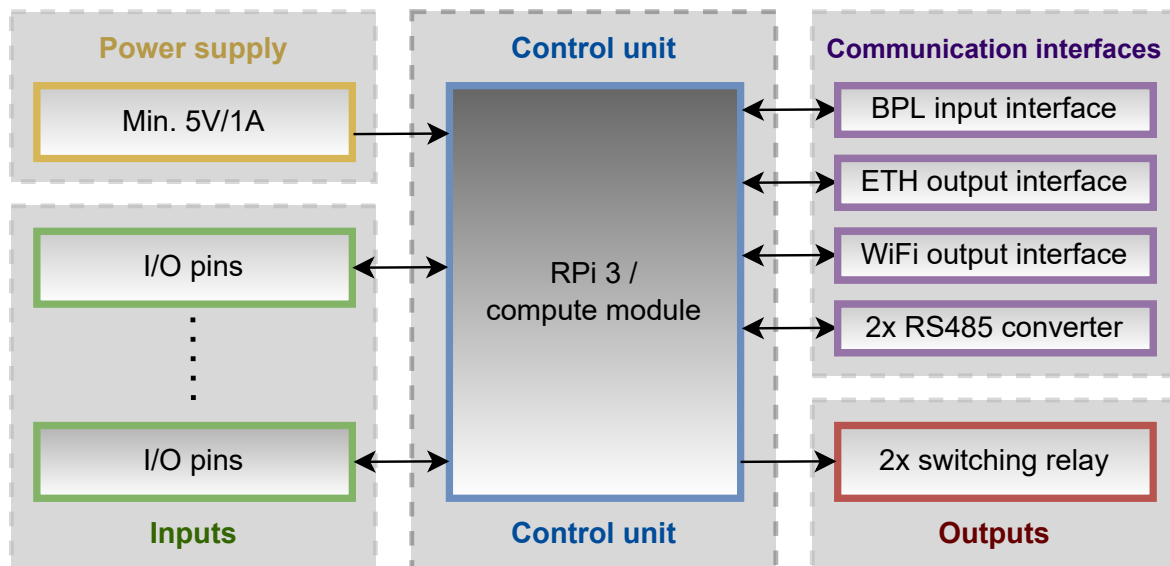
Production of electricity is inherently linked to the use of fossil fuel sources, which have an undeniable negative impact on the environment. By using renewable sources of electricity, intelligent metering of electricity and the associated higher energy efficiency, we are able to significantly reduce these negative effects [1]. The concepts that aim at these goals are undoubtedly Smart Home, Smart Buildings and Smart City. All the mentioned concepts focus on control, management, signaling and response to non-standard events [2]. All these functions result in a better distribution of energy load, prediction of possible non-standard events and increased energy efficiency [3], [4].

One of the possible technologies for this type of application is PLC (Power Line Communication). For data transmission, it uses the existing electrical wiring on which the useful signal is modulated using appropriate modulation depending on the standard and the quality of the transmission channel. Its main advantage is its easy implementation in a wide range of applications, fast deployment, reliability and affordability. To ensure a high transmission speed, low delay and sufficient transmission distance, broadband PLC (BPL) technology is considered [5]. At present, this technology is used mainly in the small local networks, remote meter reading and also as a communication channel between the EV (electric vehicle) and the EVSE (electric vehicle supply equipment) [6].

This paper deals with a detailed description of a multi-purpose gateway that processes measured data from measuring units using broadband PLC technology [7]. The possibility of measuring electrical and non-electrical quantities in real time, automatic response of the system to network events and control of several external devices provides the prerequisites for a wide range of uses of this system.

## 2. SMART GATEWAY CONCEPT DESCRIPTION

The smart gateway forms the communication interface between the integrated socket module (measuring units) and the superior monitoring system. Up to 16 socket modules (using another PLC module even more) can be connected to each gateway via a BPL interface. The primary purpose of the smart gateway is the aggregation of measured data from integrated socket modules, forwarding data to the superior system, switching of two external devices and the possibility of reading data from external devices using RS485 serial communication. As the gateway has a higher computing power, its purpose is also to perform more complex computing operations with measured data. The smart gateway also implements the monitoring and control function, described in more detail in chapter 4. The gateway's task is also to transmit measured and calculated data via wireless and wired interfaces to the superior monitoring system. The design of the smart gateway and all sub-parts is shown in the block diagram in the figure 1.



**Figure 1:** Block diagram of the broadband PLC gateway.

The smart gateway will be primarily designed in transparent mode. The input data will be forwarded to the output without any modification. From the point of view of communication between individual system entities, measuring units or other power quality monitor of electrical quantities and superior SW the communication will take place transparently. In practice, this means that at the application layer, for example the DLMS/COSEM or IEC 60870-5-104 protocol, the superior SW will communicate directly with the terminal device. The gateway works here as a transparent node, which forwards but does not modify the data (from the point of view of the application layer).

The main element of the design is the control part, which was implemented during the design using a minicomputer Raspberry Pi3 or RPi compute module. RPi 3 has sufficient computing power, which is needed to implement all the functionalities described in this chapter. In the final solution, an expansion board will be added to the RPi 3, which will contain all the sub-parts described below, or a completely new board with connectors for connecting the RPi computing module will be designed. The DIN rail of the switchboard is the considered place for mounting the device.

The list of individual smart gateway functionalities is as follows:

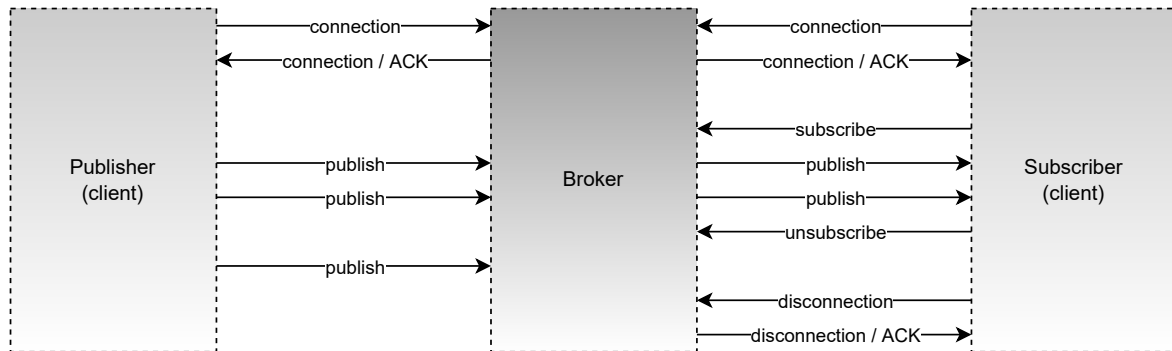
- Possibility of switching the external appliance 2x (relay box function)
- Possibility of communication with other devices via serial communication RS-485 2x
- Implementation of more complex computational operations with measured data
- Defining thresholds for automatic system responses to emergencies/conditions
- Signaling of non-standard events or conditions to the superior system
- Aggregation of measured data
- SW applications for supervision, management and control
- Possibility of various type of communication with the superior system (monitoring center)

### 3. MQTT COMMUNICATION PROTOCOL

Communication and data transfer between the individual integrated socket module and the master gateway is implemented via the MQTT (Message Queue Telemetry Transport) protocol. It is a lightweight and open communication protocol that is built on top of the TCP protocol and uses port 1833 for transmission. If secure transmission is required, then port 8883 via TLS / SSL.

In the MQTT protocol environment, messages are exchanged between the client and the broker. In the case of the presented system, the client is measuring unit and the broker is a smart gateway, which can

also be a client or a publisher/subscriber. The client can be a publisher or a subscriber. The client sends the data to the central point (MQTT broker) in publisher mode. The content of the data has a certain topic contained within it, thanks to which the subscriber is able to receive data according to the set topic. Recipients and news providers can publish and receive data from a variety of topics [8]. The principle of communication is shown in figure 2.



**Figure 2:** MQTT communication diagram.

Basic types of MQTT messages:

- Connect - establishing communication with the broker
- Disconnect - interruption of communication with the broker
- Publish - publishing a message to a specific topic
- Subscribe - subscribe to a message from a specific topic
- Unsubscribe - unsubscribe from a topic

The basic principles of data transmission security in the MQTT protocol are implemented by the following protection methods [9]:

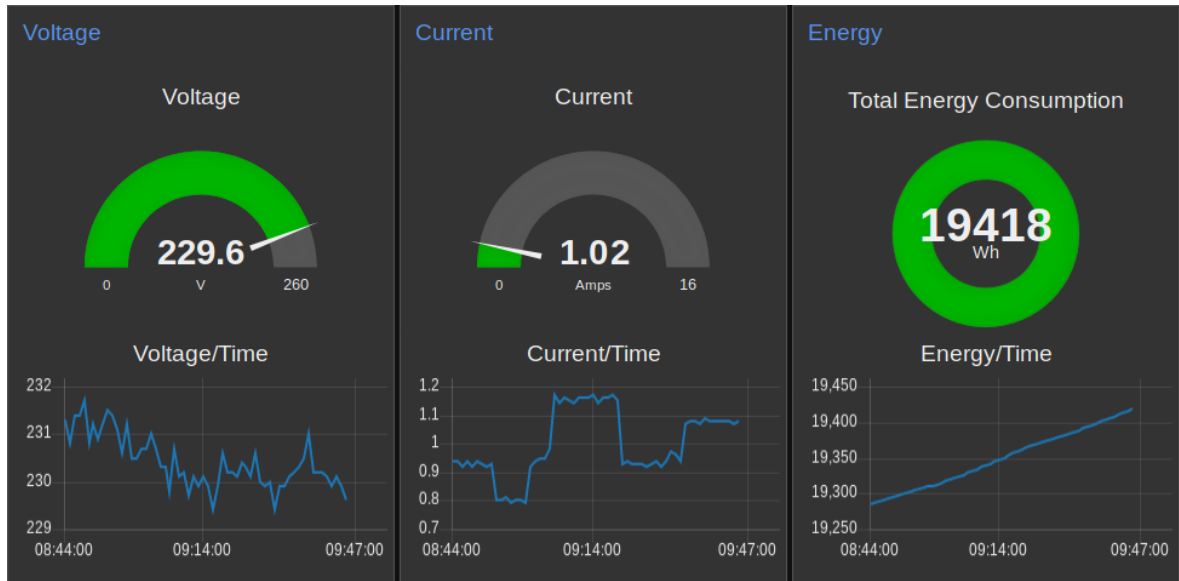
- Control client access to the broker using a specific client ID
- Client access to the broker via secure TLS/SSL protocol

Client authentication, where the "connect" message contains the "username" and "password" fields. The broker is then able to use these fields to authenticate the client.

#### 4. SOFTWARE FOR MANAGEMENT AND MONITORING ENERGY CONSUMPTION

The software application for control and monitoring is implemented using the Node-RED tool. It is a low-level programming language that allows fast and efficient design of applications and control functions through nodes with a defined function or with the ability to define a specific function. Node-RED is built on the Node.js software system, which is suitable for easily scalable applications and is written in JavaScript. It is characterized by low overhead and maximizing performance. Another advantage of Node-RED is the possibility of simple implementation of the entire monitoring system on the cloud [10].

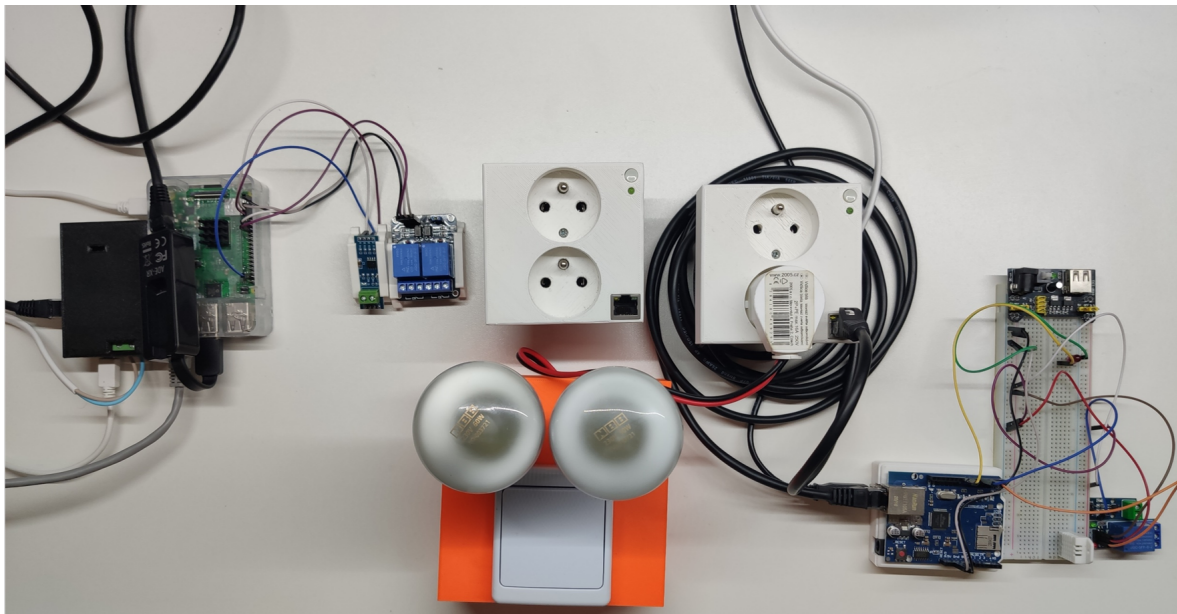
Node-RED also contains the necessary libraries for implementing the MQTT protocol, described in chapter 3. The MQTT protocol will be used to forward the measured values from the measuring units to the SW application, where the data will then be used for subsequent post-processing. In order to be able to use the MQTT protocol on the selected HW for the smart gateway, it is necessary to install an MQTT broker on this device, which in principle acts as a server. The open source software Eclipse Mosquitto was selected to install the MQTT broker, which supports the latest version of protocol 5.0, but is also backward compatible with previous versions. The smart gateway with all the necessary tools, FW and SW applications forms a full-fledged system for monitoring and necessary control of the entire system. Figure 3 shows the example dashboard of a simple application in the Node-RED environment.



**Figure 3:** Example of user interface in Node-RED dashboard.

## 5. TESTBED FOR SYSTEM TESTING

Figure 4 shows realized testbed. The testbed was designed to verify the functionality of the entire system, which contains all the described sub-parts of the system and is still in the process of development. The smart gateway described in this article is implemented by an RPi 3B+ microcomputer with an external broadband module meeting the HomePlugAV standard. Through PLC technology, communication is realized with measuring units, the measuring part of the integrated socket is realized in a breadboard using the Arduino prototyping platform. The electronic load consists of a pair of light bulbs with the possibility of switching.



**Figure 4:** Testbed for evaluation of PLC monitoring system.



## 6. CONCLUSION

This article describes in more detail the concept of the design of an intelligent gateway for monitoring the energy network behind the point of consumption. The gateway concept is based on the RPi unit using a PLC module for the possibility of communication with measuring units via PLC technology. The main advantage of this solution is the possibility of fast deployment and scalability of the entire system. Thanks to sufficient throughput and delay, it is possible to monitor measuring units in real time and perform more complex operations on the basis of measured data, which require higher computing power. The gateway also has monitoring software for displayed measured values, and thanks to the implementation of switching relays and an RS485 converter, it is also possible to switch or control additional external devices using the gateway. In the following phase, a more detailed measurement of all partial parameters and optimization of the entire system will be performed.

## REFERENCES

- [1] C. V. Ionescu and G. Darie, "Energy Effectiveness-New Energy Performance Indicator to optimize the Industrial Energy Consumptions," International Conference and Exposition on Electrical And Power Engineering (EPE), 2020, pp. 462-465, doi: [10.1109/EPE50722.2020.9305663](https://doi.org/10.1109/EPE50722.2020.9305663)
- [2] A. Aigerim, B. Omarov and Y. I. Cho, "Multi-objective Optimization for Smart Building Energy and Comfort Management as a Case Study of Smart City Platform", In: 2017 IEEE 19th International Conference on High Performance Computing and Communications, IEEE, 2017, pp. 627-628, ISBN 1538625881, doi:[10.1109/HPCC-SmartCity-DSS.2017.86](https://doi.org/10.1109/HPCC-SmartCity-DSS.2017.86)
- [3] X. Zhong, H. Ma, T. K. Saha and R. Zhang. "Consumption Scenario-Based Probabilistic Load Forecasting of Single Household", IEEE transactions on smart grid, IEEE, 2022, pp. 1075-1087, ISSN 1949-3053, doi:[10.1109/TSG.2021.3132039](https://doi.org/10.1109/TSG.2021.3132039)
- [4] SEGARRA, E. Lucas, G. R. Ruiz and C. F. Bandera, "Probabilistic load forecasting for building energy models", *Sensors*, MDPI, 2020, pp. 1-20, ISSN 1424-8220, doi:[10.3390/s20226525](https://doi.org/10.3390/s20226525)
- [5] P. Mlýnek, M. Ruzs, L. Beneš, J. Sláček and P. Musil. "Possibilities of broadband power line communications for smart home and smart building applications", *Sensors*, 2021, pp. 1-17, ISSN 1424-8220, doi:[10.3390/s21010240](https://doi.org/10.3390/s21010240).
- [6] V. Blazek, M. Petruzela, J. Vysocky, L. Prokop, S. Misak and D. Seidl, "Concept of Real-Time Communication in Off-Grid System with Vehicle-to-Home Technology," 21st International Scientific Conference on Electric Power Engineering (EPE), 2020, pp. 1-6, doi: [10.1109/EPE51172.2020.9269236](https://doi.org/10.1109/EPE51172.2020.9269236)
- [7] W. d. S. Costa, W. G. V. d. Santos, H. R. d. O. Rocha, M. E. V. Segatto and J. A. L. Silva, "Power Line Communication based SmartPlug Prototype for Power Consumption Monitoring in Smart Homes," in IEEE Latin America Transactions, pp. 1849-1857, 2021, doi: [10.1109/TLA.2021.9475618](https://doi.org/10.1109/TLA.2021.9475618)
- [8] LIMA, Diana Bezerra Correia, Rubens Matheus Brasil DA SILVA LIMA, Douglas DE FARIAS MEDEIROS, Renata Imaculada Soares PEREIRA, Cleonilson Protasio DE SOUZA and Orlando BAIOCCHI. "A Performance Evaluation of Raspberry Pi Zero W Based Gateway Running MQTT Broker for IoT," In: 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). IEEE, 2019, pp. 76-81, doi:[10.1109/IEMCON.2019.8936206](https://doi.org/10.1109/IEMCON.2019.8936206)
- [9] P. Chang-seop and H. Nam, "Security Architecture and Protocols for secure MQTT-SN", *IEEE access*, Piscataway, 2020, ISSN 2169-3536, doi:[10.1109/ACCESS.2020.3045441](https://doi.org/10.1109/ACCESS.2020.3045441).
- [10] S. Chanthakit and C. Rattanapoka, "MQTT Based Air Quality Monitoring System using Node MCU and Node-RED," 2018 Seventh ICT International Student Project Conference (ICT-ISPC), 2018, pp. 1-5, doi: [10.1109/ICT-ISPC.2018.8523891](https://doi.org/10.1109/ICT-ISPC.2018.8523891)

# Radio spectrum analysis with a focus on WiFi bands using software-defined radio

J. Pospisil<sup>1</sup>, R. Fujdiak<sup>1</sup>

<sup>1</sup>Department of Telecommunications, Brno University of Technology, Technická 12, 616 00 Brno, Czech Republic

E-mail: [xpospi90@vutbr.cz](mailto:xpospi90@vutbr.cz)

**Abstract**—This paper deals with the issue of radio spectrum analysis focused on WiFi (Wireless Fidelity) bands. The aim of the article is to approach the issue of capturing the radio spectrum and its visualization. Ways of approaching the solution of the problem are discussed, and on the basis of this comparison, attention is further focused on the use of SDR (Software Defined Radio). This device creates digitalized samples of captured radio signals and forwards them via USB to a computer, where the samples are further processed and visualized. The result is a functional program capable of visualizing this data in the form of a waterfall spectrogram. These results can also be additionally used for advanced machine analysis.

**Keywords**—Wireless spectrum, analysis, SDR, spectrogram, wireless probe, WiFi spectrum

## 1. INTRODUCTION

At a time when most people in the modern world use the Internet on a daily basis, and there is an ever-increasing number of end devices connected to the Internet, there is also the aim to address reliability and optimize connection quality. Therefore, it is necessary to define a way to obtain data on the basis of which the analysis and subsequent optimization of wireless data transmission can be performed. This article deals with getting the data for analysis of the radio spectrum in the bands in which WiFi (Wireless Fidelity) communication (2.4 GHz) is operated, i.e., in the ISM (Industrial, Scientific and Medical) bands that can be used without a license. The result is a state-of-the-art software that will be able to log and display current events in the radio spectrum. In case of an unreliable connection, it is then possible to quickly verify whether, for example, external interference by another device is occurring. Multiple types of devices can generally be considered to implement the capture of the physical layer of WiFi communication. The following options are available:

- *Wi-Fi card* – The hardware of the WiFi transceiver is intended for this communication, so it would certainly be appropriate to use this HW for analysis as well. However, from the implementation point of view, the manufacturers mostly don't offer support for the provided functions of individual integrated circuits. The WiFi cards in monitoring mode are generally able to capture only physical frames but are not able to analyze the physical layer in detail from the point of view of interference. Furthermore, the WiFi cards are limited by the amount of bandwidth for which they are intended. Therefore, to scan multiple channels, it is necessary to switch between channels.
- *Wireless spectrum probe* – Directly designated HW from a specialized manufacturer is another option. These devices generally provide more comprehensive information about the physical layer than Wi-Fi cards. The problem, however, is that every manufacturer provides a proprietary solution. Each company provides its analytical tool (stand-alone solution), so it is necessary to individually address the software modification of the manufacturer for the ability to redirect the program output to its own database/client system — furthermore, possible issues of support from the manufacturer for these specific solutions may arise.
- *Software defined radio (SDR)* – The last option is a custom solution for the implementation of specific requirements. In this case, it is possible to use universal hardware in the form of SDR. The SDR device is able to receive the ambient radio signals at the set parameters in the form of the center frequency (Frequency - F), bandwidth (Bandwidth - BW), sampling frequency (Sampling Rate - SR), and gain (Gain - G). After setting these parameters, the device starts capturing (digitizing) the received signal. It is then possible to implement signal processing by software upon the captured signal samples.

Following the list of possibilities for implementing the physical layer probe, we decided to completely implement our own solution based on the SDR platform, given the opportunity of any modification in the future. However, the subsequent task was to select a suitable device.

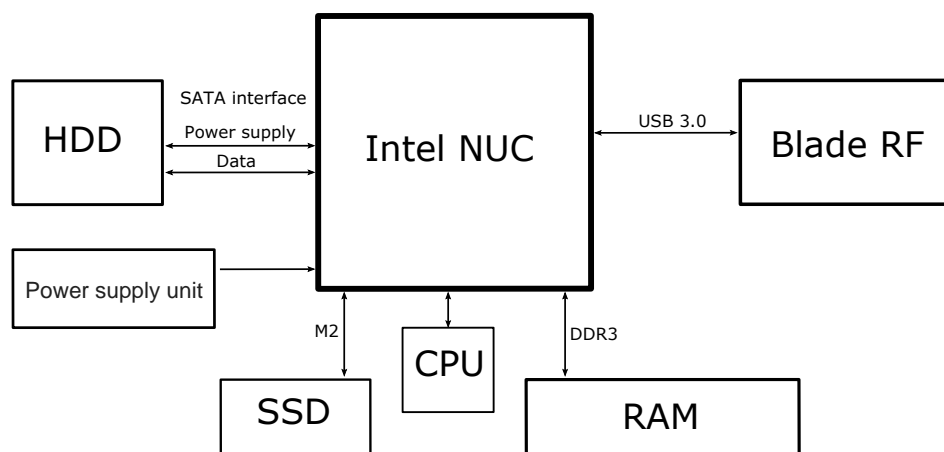
**Table I:** Parameters comparison of selected SDRs (2021).

	Frequency range	Bandwidth	Sampling frequency	Approximate price
AirSpy R2 [1]	24 – 1700 MHz.	10 MHz	10 Msps	€150
FreeSRP [2]	70 – 6000 MHz	61.44 MHz	61.44 Msps	€345
LimeSDR [3]	100 kHz – 3.8 GHz	61.44 MHz	61.44 Msps	€260
USRP B200 [4]	70 MHz – 6 GHz	56 MHz	56 Msps	€580
easySDR [5]	64 – 1700 MHz	-	96 ksps	€95
BladeRF 2.0 [6]	47 MHz – 6 GHz	56 MHz	61.44 Msps	€415
HackRF One [7]	1 MHz – 6 GHz	20 MHz	20 Msps	€260

There are already studies dealing with radio spectrum capture using SDR. Most focus is on RTL-SDR devices [8]–[11], also sold as AirSpy R2. They can receive signals only in the frequency range from 24–1700 MHz and bandwidth only 10 MHz (Table I). Because the probe is focused on the analysis of the physical layer for Wi-Fi technology, the ability to capture not only in the 2.4 GHz band (Wi-Fi b/g/n/ax) but also in the 5 GHz band is required. (Wi-Fi a/n/ac). Thus, according to Table I, we narrowed down the selection to FreeSRP, BladeRF, and HackRF One devices only. Due to the technical parameters, but especially the availability and support of individual manufacturers, we chose the BladeRF 2.0 device. Of course, there are already programs that work with SDRs and are able to graphically display what is happening in the spectrum, mentioning only a two GQRX [12] and CubicSDR [13]. However, our goal was to create our own program that could be modified according to future needs.

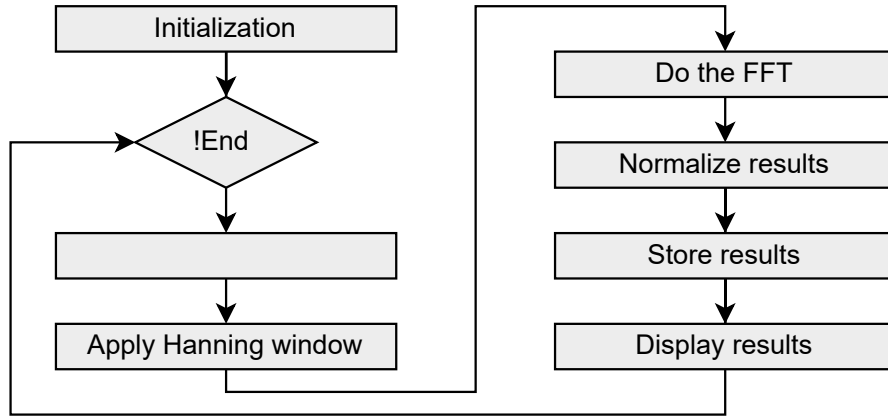
## 2. DESIGN OF WIRELESS PROBE

The computing module uses the Intel NUC platform and consists of a dual-core Intel Core i5-7300U processor with a frequency of up to 3.5 GHz and HyperThreading support. The board is equipped with two Hynix SODIMM DDR4 RAM modules with a frequency of 2666 MHz and a capacity of 16GB. The Samsung 860 EVO 250GB SSD drive with M.2 slot is also connected. The performance of the computing unit is dimensioned so that it would be sufficient even in the case of expanding the application with more complex functionalities. The SDR device is connected to the computing unit via a USB (Universal Serial Bus) version 3.0 port. We used the SDR device BladeRF 2.0 micro xA9 to receive the radio signal for the resulting analysis. The Table I lists the basic parameters. In addition to the parameters in the table, the advantage of the device is the support of 2x2 MIMO (Multiple Input, Multiple Output) and also the presence of the Altera Cyclone V FPGA circuit. Figure 1 shows the complete diagram.



**Figure 1:** Complete wireless probe hardware.

The procedure for processing the received data from SDR is described as follows. For an overview of the flowchart, see Figure 2. In the first step, the sequence of radio signal samples is received in the software. The Hanning window is applied to the received samples to create a signal's gradual rising and falling edge due to sidewall reduction constraints. Subsequently, a fast Fourier transform is applied, the output of which are the weight coefficients of the individual frequency components of the analyzed signal. However, the Fourier transform output is not directly suitable for displaying due to its large dynamic range. In the next step, the individual result coefficients are converted to the logarithmic unit dB and further quantized to the range 0–255 (one byte). The results processed in this way are stored in a file and then sent to the main process, which takes care of displaying the results.



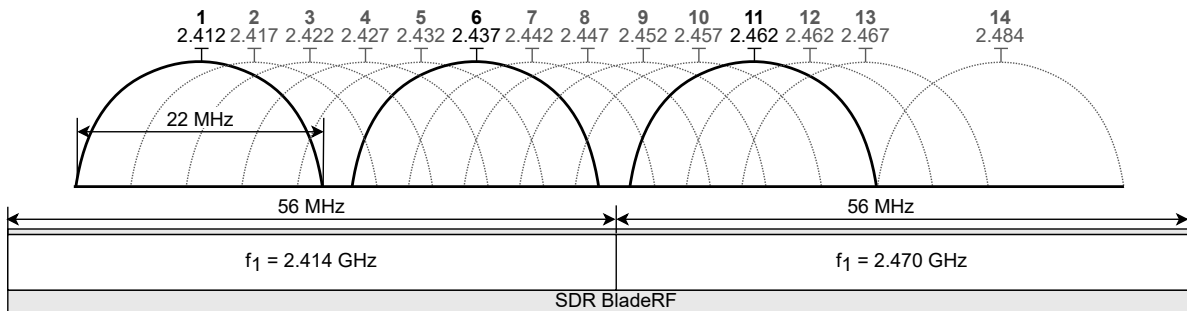
**Figure 2:** Flowchart for SpectrumAnalyzer software.

For signal capture, the center frequency depending on the channel on which the capture is to take place must be set at first. Because BladeRF provides a bandwidth of up to 56 MHz, 2430 MHz can be set to capture the first 7 WiFi channels in the 2.4 GHz band. It is advisable to set the sampling frequency to the maximum value to reduce the distortion of the results. Gain value can be arbitrarily changed over time to ensure optimal results. The FFT size is defined by default at 2048. This number indicates the number of frequency coefficients, in which case it can be converted to the number of pixels in one line on the screen. For clarity, the default program settings are defined as follows: center frequency is 2428 MHz, bandwidth is 56 MHz, sampling frequency is 61 MHz, gain is 25 dB, and FFT size is 2048.

Since the radio signal also attenuates with increasing distance, it is important to have the probe close enough (units up to tens of meters) to observe the ongoing communication in the spectrum. The basic attenuation model FSPL (Free Space Path Loss) is then given[14] by:

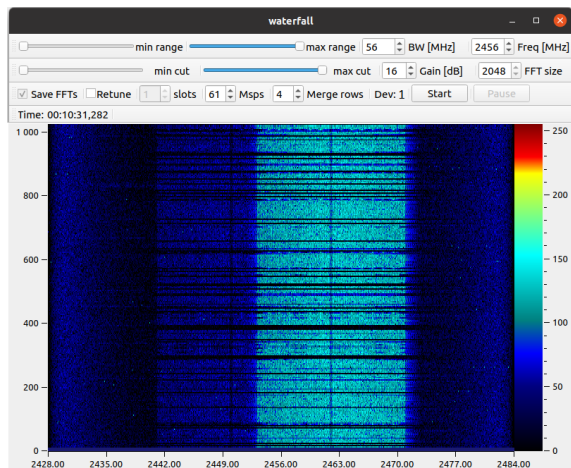
$$FSPL(dB) = 20 \log_{10}(d) + 20 \log_{10}(f) - 147.55, \quad (1)$$

where  $d$  is the distance from the transmitter given in meters and  $f$  is the frequency of the electromagnetic wave given in Hertz units. As part of the probe optimization, the problem with the bandwidth limiting

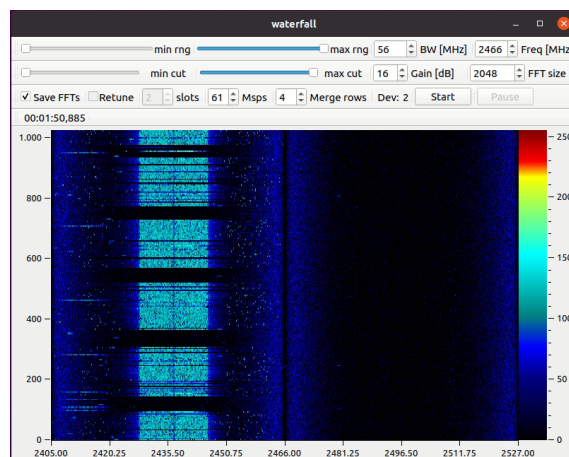


**Figure 3:** Distribution of channels (with a bandwidth of 22 MHz) of the 802.11 standard in the 2.4 GHz band [15].

factor of the captured signal was solved. To cover the entire 2.4 GHz band (all 13 channels) of the 802.11 protocol, the use of two frequencies (2.414 and 2.470 GHz) was proposed, in the mode of two SDR devices (see Figure 3). For this purpose, an extension of the program was made, which allows increasing the bandwidth. This was achieved by connecting a second SDR device. Each SDR then captures a different part of the radio spectrum, which is then displayed as one uniform part of the radio spectrum in the waterfall (see Figures 4 and 5).

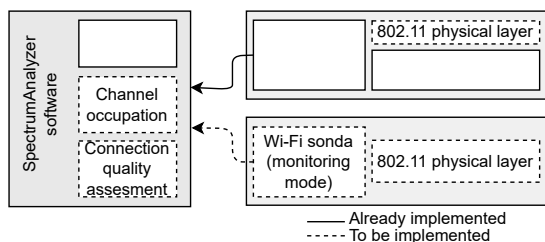


**Figure 4:** Capturing the radio spectrum with a native maximum bandwidth of 56 MHz.

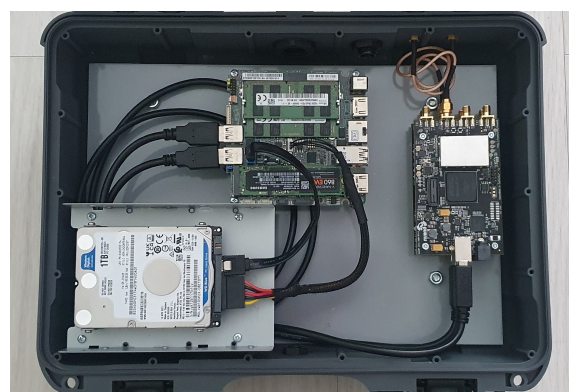


**Figure 5:** Capturing the radio spectrum using two SDRs, the whole bandwidth of 112 MHz.

Due to the own development of the SpectrumAnalyzer program, it is possible to further expand the functionality according to our needs. Figure 6 illustrates in more detail the current functionalities of the program (blocks bounded by a solid line) as well as possible extensions in the future (blocks bounded by a dashed line).



**Figure 6:** Modular scheme of a wireless probe for 802.11 standard.



**Figure 7:** Finished probe device in a briefcase.

The portable wireless probe is intended for measurements at specific locations in various buildings, so it is advisable to choose a box that can be conveniently carried and is resistant to shocks and water (see Figure 7). It was selected from a series of durable NANUK suitcases made in Canada. These cases meet the described requirements - they have a durable transmission handle, are suitable for protecting sensitive objects (electronics, weapons), and are sufficiently waterproof (IP67). The model NANUK Case 910 with internal dimensions 363 mm (length), 234 mm height, 104 mm depth was selected.

To bring more detailed information and traffic statistics, we consider the following. The current capability of the program is the reception of raw samples of the radio spectrum and, based on the conversion of samples from the time domain to the frequency, the ability to observe the allocation of the spectrum in real-time graphically. Other possible extensions include adding 802.11 signal demodulation



support directly to the SDR for possible analysis at higher layers of communication. Due to the complexity of this solution, it is also possible to create a software module for an external Wi-Fi card instead of demodulating the signal in SDR, which will take care of the analysis of 802.11 data communication and will be able to add more detailed information in SpectrumAnalyzer to data from the radio spectrum.

### 3. CONCLUSION

Using SDR, we managed to create and put into operation a complete device for radio spectrum analysis focused on WiFi communication. In its current state, the device is able to graphically display the course of events in the spectrum in the form of a real-time spectrogram. In this form, the program's output can be used for advanced software, which, based on the input data, can detect anomalies on the physical layer and help solve the cause of problems that may occur during wireless communication. Our next goal is to continue the development of the wireless probe and expand it in terms of adding 802.11 frame analysis in order to obtain more detailed information for better traffic analysis.

### ACKNOWLEDGMENT

The research of this paper is a part of the project no. TK02030013 funded by Technology Agency of the Czech Republic.

### REFERENCES

- [1] AispySDR. High Performance Software Defined Radios [online]. 2013. [Accessed 22 March 2022]. Available from: <https://airspy.com>
- [2] FreeSRP Alpha. Open source software-defined radio [online]. 2017. [Accessed 22 March 2022]. Available from: <http://freesrp.org>
- [3] Lime microsystems, LimeSDR [online]. 2005. [Accessed 22 March 2022]. Available from: <https://limemicro.com/products/boards/limesdr/>
- [4] Ettus Research, USRP B210 [online]. 2013. [Accessed 22 March 2022]. Available from: <https://www.ettus.com/all-products/UB200-KIT/>
- [5] Microsat, EasySDR USB Dongle [online]. 2011. [Accessed 22 March 2022]. Available from: [http://microsat.com.pl/product\\_info.php?products\\_id=35](http://microsat.com.pl/product_info.php?products_id=35)
- [6] Nuand LLC, BladeRF 2.0 [online]. 2018. [Accessed 22 March 2022]. Available from: <https://www.nuand.com/bladerf-2-0-micro/>
- [7] Great Scott Gadget, HackRF [online]. 2012. [Accessed 22 March 2022]. Available from: <https://greatscottgadgets.com/hackrf/>
- [8] RAHMAN, Md Habibur; ISLAM, Md Mamunoor. A practical approach to spectrum analyzing unit using rtl-sdr. *Rajshahi University Journal of Science and Engineering*, 2016, 44: 151-159.
- [9] SRUTHI, M. B., et al. Low cost digital transceiver design for software defined radio using RTL-SDR. In: *2013 international mutli-conference on automation, computing, communication, control and compressed sensing (iMac4s)*. IEEE, 2013. p. 852-855.
- [10] WICKERT, Mark A.; LOVEJOY, McKenna R. Hands-on software defined radio experiments with the low-cost RTL-SDR dongle. In: *2015 IEEE signal processing and signal processing education workshop (SP/SPE)*. IEEE, 2015. p. 65-70.
- [11] SIERRA, E. G.; ARROYAVE, GA Ramirez. Low cost SDR spectrum analyzer and analog radio receiver using GNU radio, raspberry Pi2 and SDR-RTL dongle. In: *2015 7th IEEE Latin-American conference on communications (LATINCOM)*. IEEE, 2015. p. 1-6.
- [12] CSETE, A. Gqrx SDR-open source software defined radio by Alexandru Csete OZ9AEC [online]. Denmark, 2017. [Accessed 22 March 2022]. Available form: <https://gqrx.dk/>
- [13] CHARLES, Cliffe J. CubicSDR, Cross-Platform and Open-Source Software-Defined Radio Application [online]. Canada, 2013. [Accessed 22 March 2022]. Available from: <https://cubicsdr.com/>
- [14] ISLAM, Syed Kamrul; HAIDER, Mohammad Rafiqul. *Sensors and low power signal processing*. Springer Science & Business Media, 2009.
- [15] FLICKENGER, Rob. *Wireless Networking in the Developing World: A practical guide to planning and building low-cost telecommunications infrastructure*. Hacker Friendly LLC, Seattle, WA, US, 2007.



# Získej víc s IEEE



[www.ieee.org](http://www.ieee.org)

## Co je to IEEE?

IEEE (Institute of Electrical and Electronics Engineers) je celosvětová nezisková organizace, která pomáhá v rozvoji technologií, inovací a excelentnosti ve prospěch lidstva. Ač vám to může znít až příliš honosně, níže najdete některé z aktivit, které tuto misi potvrzují.

## Kde všude IEEE působí a kolik má členů?

IEEE vznikla v USA už v roce 1884 pod názvem AIEE. Hlavním cílem bylo spojit odborníky v elektrotechnice a sdílet nejnovější technologie ku prospěchu všech. Za jejím založením stály jména jako A.G. Bell, T.A. Edison nebo společnosti jako General Electric nebo Siemens. Značka IEEE se potom objevila až v roce 1963 po transformaci AIEE, ale to už bylo registrovaných více než 150 tisíc členů. V současnosti působí IEEE celosvětově a hlásí se k ní více než 400 tisíc členů.

Mezi hlavní aktivity IEEE patří organizování mezinárodních vědeckých konferencí, vydávání odborných publikací a mezinárodních norem (např. IEEE 802.11), podpora členů, studentské aktivity či humanitární akce.

## Výhody členství v IEEE

### Aktuální informace z oboru elektrotechniky

IEEE spravuje jednu z největších databází odborných publikací na světě IEEE Xplore. Tato databáze je celosvětově uznávaná a obsahuje více než 3 milióny publikací, převážně článků z konferencí a odborných časopisů.

### Nástroje pro vaši kariéru

IEEE poskytuje například nástroje pro vytvoření cíleného životopisu (ResumeLab), databáze mentorů, celosvětovou databázi volných pracovních míst, atd.

### Podpora odborných aktivit

IEEE vypisuje množství soutěží a grantových programů, do kterých se můžete zapojit. Váš příspěvek z EEICT můžete například přihlásit do soutěže Student Paper Contest Regionu 8. Taktéž podporuje publikační činnost a svým členům umožňuje volný přístup k licenci Overleaf PRO.

# Application for Testing Smart Meters

David Kohout

Brno University of Technology, FEEC, Czech Republic

E-mail: [david.kohout@vut.cz](mailto:david.kohout@vut.cz)

**Abstract**—This paper presents our developed application for smart metering, which we are using to test the interoperability of different smart meters, compatibility with transmission technologies, to measure data volumes, test data security, and also ensure data security through the whole life-cycle of smart meters. The paper concludes with future development and planned features of our application, which is already being used for testing smart metering technologies at the Smart Grid Lab at BUT.

**Keywords**— DLMS, electricity measuring, smart metering, testing application

## 1. INTRODUCTION

In today's modern age, more and more elements of everyday life are being connected to the Internet and this brings us to terms like IoT (Internet of Things), smart grids and 5G (5th generation networks). This trend has also reached energy measuring technologies and it is being referred to as smart metering. For this purpose there exist multiple protocols and standards, but the most promising one is the standard DLMS (Device Language Message Specification). One of the biggest advantages of using DLMS is the interoperability between smart meters (in short SMs) from different manufacturers. A disadvantage can be in higher data complexity and also larger overhead data.

According to the Czech Republic legislation, we have a new law (act no. 359/2020 Sb. [1]), it comes into force in 2024, and this law requires from electricity distributors to mount SMs to all customers, whose energy consumption is higher than 6 MWh per year. This means more than 1 million new SMs in the Czech Republic. It also specifies a level of security on SMs. The specified security level is high/secure enough by today's standards, but we need to count with longevity (15+ years) of electricity meters to ensure, that they will be capable to secure all the data in the future with higher security requirements.

### 1.1. State of the Art

Before deploying new SMs into everyday usage, we need to test all of it's functionalities and also compatibility between meters from multiple manufacturers in the controlled environment. First real world deployment will be mostly selective, and that means using transmission technologies like NB-IoT (NarrowBand) [2], LTE Cat. M1 (Long Term Evolution) or GPRS (General Packet Radio Service). With regards to the high amount of SMs that will be deployed and with these transmission technologies, we need to consider a data volume that will be needed for communication to the SMs. Data volume is directly related to the price for data transfer in relation to the data plans from the mobile network operators and also to the setup of the mobile network to manage high amount of SMs connected to one Base Transceiver Stations. We need to test all desired parameters (e.g., reading data volume, latency, data security, interoperability), but there is no complete solution on the market that could satisfy all requirements. At the moment, there exist only a small number of applications capable of DLMS communication, some are proprietary from SM manufacturers and they can't be used with SMs from different manufacturers. Some applications are free and some can cost up to 60,000 €. Example applications are:

- Gurux – GXDLMSDirector [3],
- OpenMUC – jDLMS [4],
- DLMS Conformance Testing Tool [5].

### 1.2. Contribution

This paper is focused on introduction of our application (for testing DLMS and related protocols) called DATEL (DLMS Application for Testing Energy Labs). Our solution can be used to test all the mentioned requirements and it can also be expanded to fulfill any new needs in smart metering testing in the future. All measured data can be used for further research in machine learning, threat detection, and other related areas. We can also use it to test new transmission technologies (e.g., NB-IoT, LTE Cat. M1).

## 2. USED LIBRARIES

DATEL uses the programming language Java 11 (specifically JDK-11.0.5). It looks that the application is quite small, but appearances are deceptive. It uses a handful of libraries, namely:

- Gurux DLMS – DLMS library,
- JavaFX (OpenJFX) – graphical interface,
- Jmetro – modern graphic look,
- JAXB – XML framework for saving configurations.

The most important library is the Gurux DLMS framework. Gurux LTD is a Finnish company which primarily deals with the DLMS specification. Their library is available for most common programming languages like Python, C++, C# and Java. Until recently the Java version of the Gurux library was their most updated version and it contained most elements that DLMS standard specifies. That is why this application is written in Java. Today Gurux puts most of its effort into C# (.Net) version of the library, because that is used in their client application called GXDLMSDirector. Gurux provides all their products under dual licensing [6]. One license is paid and the second one is open-source using GNU GPL v2.

For testing purposes we are using other programs and devices. To have full control over communication we are using the packet capturing software Wireshark. Because most SMs use the RS485 port for communication, we need to use converters to be able to communicate with them directly with DATEL. There are some USB converters, but most of the time, we are using Ethernet to RS485 converters.

## 3. DLMS TESTING APPLICATION – DATEL

We developed our own application to be able to test DLMS meters from multiple manufacturers to satisfy all requirements from electricity distributors and the Czech Republic legislation. The main application window is shown in Figure 1. The left side represents emulated meters, bridges and added meters. Emulated section can be used with our own server application, that can emulate multiple meters with different configurations on a single device. In this case, you are only specifying the IP address and the port, rest will be configured automatically. The bridges section is used to bridge different type of communication. The last meters section represents real meters, which can be used for testing and all parameters need to be configured manually. That means setting up all DLMS parameters (e.g., client and server address, interface, referencing, authentication, passwords, IP address or serial port, etc.).

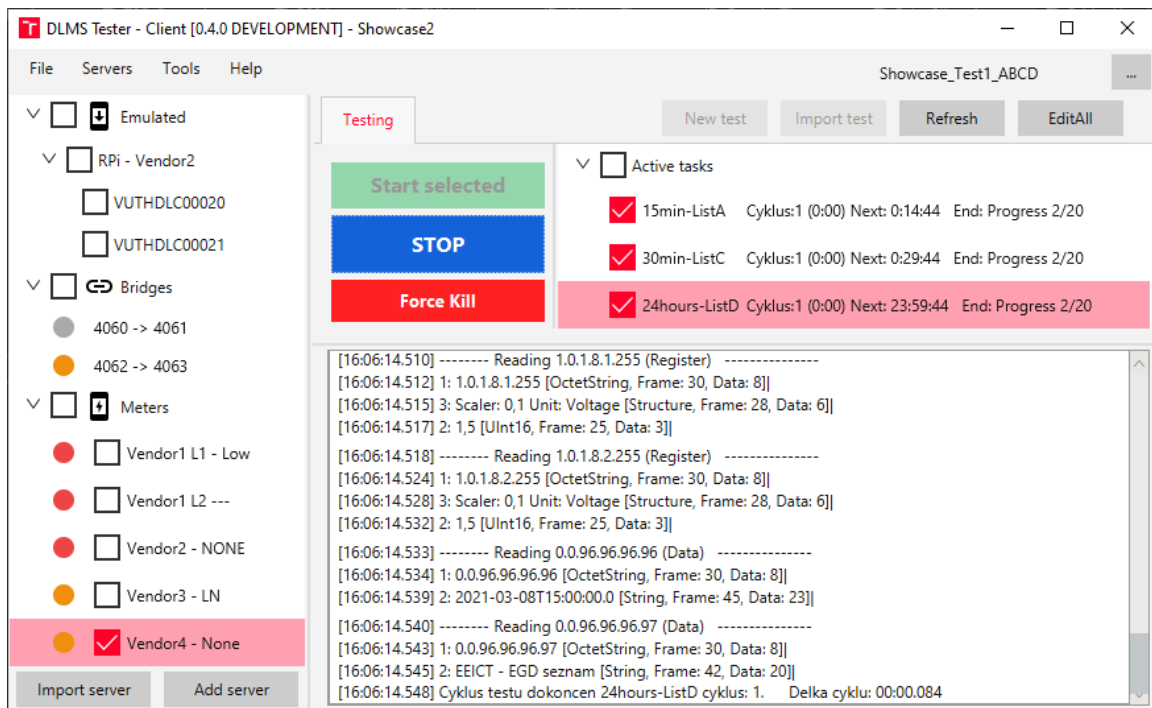


Figure 1: Main application window

The top right part is for testing configuration. The largest portion of this window is filled with the text terminal output. This output contains read values from the test '24hours-listD' during testing a meter marked as 'Vendor4 - None'. All important functionality is described in the next sections.

#### 4. DLMS DATA READING

Manual data reading (DLMS client) enables you to read values directly from the meter. Our manual reading is slightly improved, in comparison to other DLMS applications, in a way, that we can select which attributes we want to read and we can also specify, whether we want to read an object in multiple messages (attribute by attribute) or to read all selected attributes in a list. Another difference is the ability to search in all associated objects. This enables you to find desired objects more easily.

Figure 2 shows the window for manual reading. The object list (1) can be used for selecting objects for reading. All read values are outputted into the console view (4) and to the text fields section (3). This section can be used to write new values to the meter, but it's implementation is very time consuming, because it is necessary to create a new interface for each object type. It can be also used to call object methods (e.g., reconnect on a disconnect control object). Lastly, the top part of this window (2) contains the keep alive function and a stopwatch to keep track on how long the connection to the meter is open.

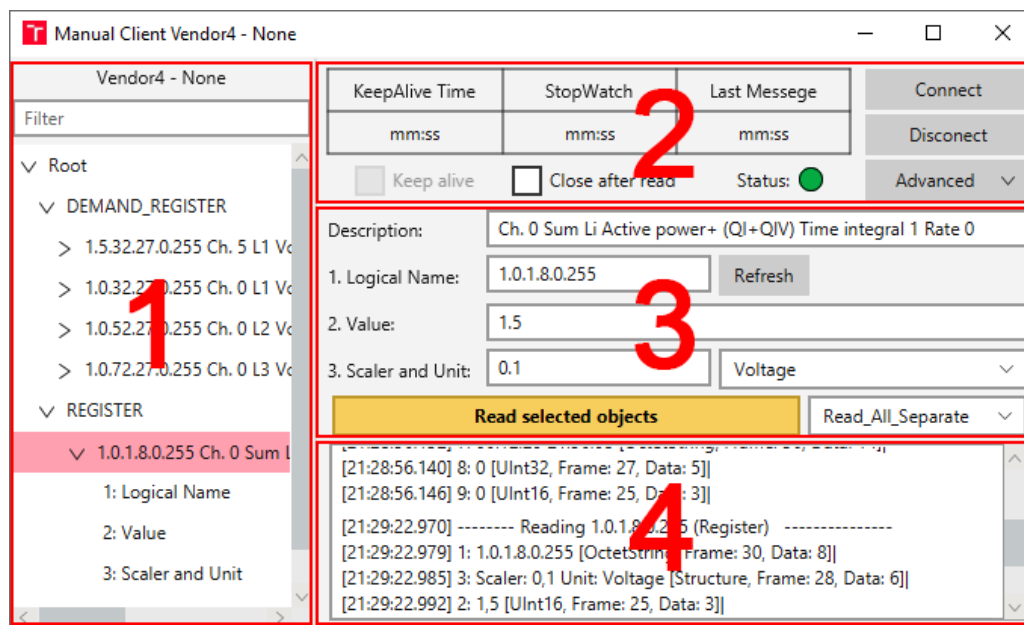


Figure 2: Manual client window

#### 5. SCHEDULED TESTING

The ability to schedule tasks to test meters is the most important part in DATEL. At the moment, we are mostly testing data volumes, that will be needed for remote reading of the meters in real deployment on various transmission technologies, or long-term testing of a technology .

The tester can add multiple tasks and specify all needed parameters like starting time, ending point (specified by time or cycles count), period (time to repeat cycle), wait times, reading type and the most important parameter is selection of all objects, that will be read during one task cycle. Objects can be searched in associated files from a meter or can be written manually into the text input field. After all parameters are filled and confirmed, this new task is added into the main window to the testing list (top right in Figure 1). The tester can change the order and also enable or disable any task. After selecting all desired tasks and meters, the test can be started in the main window by clicking on the 'Start selected' button. During testing, all specified information is saved into logs. If all running tasks had set ending points, testing will be automatically terminated, otherwise the test needs to be stopped manually.

Test evaluation is the next step after testing. Completed tests logs can be open in our log parser, and this parser will go through all log entries and after that it will print the results (e.g., number of transmitted and received messages, error messages, number of connection losses or it can count transmission delays between transmitted requests and received responses). It can also count the success rate of received



messages. Together with raw data logs we are using software for capturing data packets (e.g. Wireshark) for better error recognition and thorough results. Final results and even raw data can be used for further research e.g. designing wireless technologies specifically designated for smart metering. All captured data and knowledge can be also used for testing new implementations for meter manufacturers or finding deviations in their implementation from the standard.

DATTEL is already being used for testing real smart meters in combination with new communication units developed at BUT. In Figure 3 DATTEL is reading values from multiple SMs connected to communication units. This setup is used for testing SMs and also for verification of new technologies (e.g. LTE Cat. M1) before deployment to a production environment.

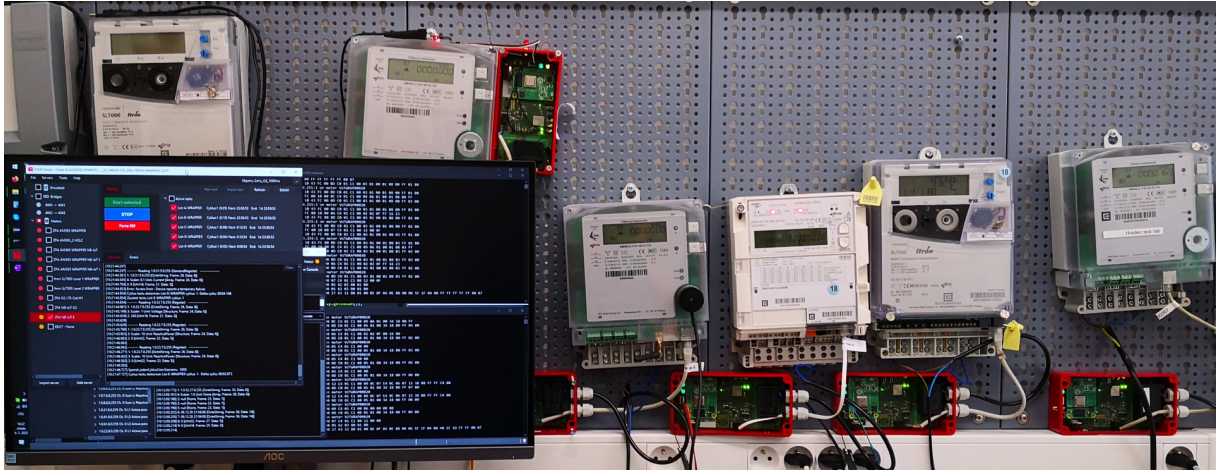


Figure 3: Smart Grid Lab at BUT

## 6. SECURITY FEATURES

DLMS offers multiple security levels called Security Suites. All suites are using for encryption AES-128 (Advanced Encryption Standard). The Suite 0 uses a pre-shared key and Suites 1 and 2 use key exchange for establishing symmetric keys. Multiple key exchange models (specified in [7]) are supported by the DLMS standard and all are using the ECDH (Elliptic Curve Diffie-Hellman) key exchange algorithm.

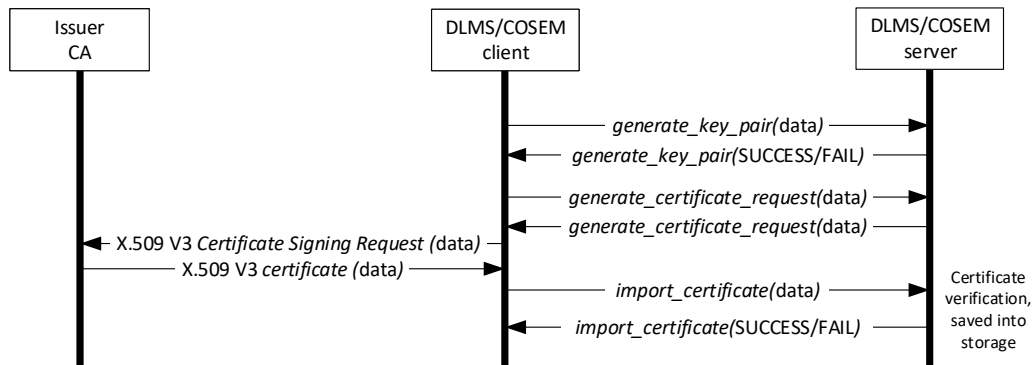


Figure 4: Certificate request and exchange [8]

DLMS Security Suite uses multiple digital certificates for digital signature and key agreement schemes. For this reason, we implemented controls for new certificate generation on client and also on server side. The standardized principle for creating new certificates on the server side is shown in Figure 4. This principle can be controlled in DATTEL with only one difference: for testing purposes it is not necessary to have a real CA (Certificate Authority). This part is directly implemented in our application. In fact, our client side also works as a certificate authority and it can be used to sign new certificates. We want to fully implement DLMS Security Suite 1 and 2, because with securing data in this way, there is a possibility to lower overhead data capacity in comparison to tunneling DLMS data in TLS tunnels. Also with working Suite 1 and 2 we can possibly help and test security implementation on new SMs from manufacturers to ensure their solution will work with other SMs and reading centrals in production environment.

## 7. CONCLUSION

This paper presents the development of our DLMS testing application – DATEL, that can be used to test all required parameters from electricity distributors and also from the Czech Republic legislation. DATEL uses Java 11 and the Gurux DLMS library. Compared to other applications, we can specify custom testing (periodically read values from SM), we can read all attributes separately, etc.

DATEL contains small adjustments in comparison to Gurux's flag software. It is possible to open multiple windows to edit multiple meters at once. This can help with configuration of multiple meters, connecting to the same meter with different access levels, or when you are debugging why this communication media is not working when the same meter was working on different media. DATEL is already being used for testing SMs and other related technologies, and it can help with future development of technologies more suited for energy measurement or for analyzing new smart meters to ensure interoperability.

As a future work, we work on the security side of DLMS with Security Suites and even with tunneling this communication inside other more secure data channels.

## 8. FUTURE WORK

DATEL already implements some important functionalities, but we can still add new functions. At the moment, the most important future task is the implementation of key exchange mechanisms into the DLMS library. In that way, we can finally have fully working Security Suites with **key exchange** mechanisms. The most difficult part is to have working key exchange inside DLMS messaging. We are currently preparing other alternatives for protecting DLMS data, and the most obvious solution is using tunneling of DLMS communication in more secure and flexible technologies e.g. **TLS tunnels** (Transport Layer Security). There are also some requests from electricity distribution companies to look into TLS tunneling. With this it is possible to setup security more precisely for this use case and also be more flexible than waiting for standardization of a new Security Suite.

Alongside, we are planning to **rework logging** part of the application in a way, that we can use raw logs directly in other programs (mainly table processing) to visualise data more easily in charts and tables. To be able to do that, the main log needs to be formatted in a slightly different way or we can make more logs each containing various important data. Together with new logging, we plan to use **TShark** and be able to launch packet capturing automatically together with testing.

Next planned feature is small modification of scheduling tasks in testing. Now all tasks are executed together at specified times and cycle periods. In some cases we need to execute task B after task A has completed all of its cycles. This can be configured, but it is very difficult as of now. Another small tweak is the ability to add some tasks to only some meters. Last change in testing is the capability of connecting to multiple meters at the same time.

Another planned feature is a **DLMS server**, which can be used directly from DATEL and it can be configured only by cloning data from real meters. This requires some verification process of downloaded data, because every meter usually has some customized part, which does not behave according to the specification or some values are not accessible.

## REFERENCES

- [1] "Electricity metering decree, Act no. 359/2020 Sb." 2020.
- [2] A. A. Amarsingh, H. A. Latchman, and D. Yang, "Narrowband power line communications: Enabling the smart grid," *IEEE Potentials*, vol. 33, no. 1, pp. 16–21, 2014.
- [3] "GXDLMSDirector." [Online]. Available: <https://www.gurux.fi/GXDLMSDirector>
- [4] "jDLMS." [Online]. Available: <https://www.openmuc.org/dlms-cosem/>
- [5] "DLMS – The Conformance Test Tool." [Online]. Available: <https://www.dlms.com/certification-scheme/conformance-test-tool>
- [6] "Dual licensing of Gurux products." [Online]. Available: <https://www.gurux.fi/duallicense>
- [7] E. Barker, L. Chen, S. Keller, A. Roginsky, A. Vassilev, and R. Davis, "Recommendation for pair-wise key-establishment schemes using discrete logarithm cryptography," National Institute of Standards and Technology, Tech. Rep., 2017.
- [8] *Green Book: DLMS/COSEM Architecture and Protocols*, 10th ed., Switzerland, 2020.



# Network probe: Network monitoring and management tool

A. Boháčik, R. Fujdiak, and J. Mišurec

Brno University of Technology, FEEC, Czech Republic

E-mail: [Antonin.Bohacik@vut.cz](mailto:Antonin.Bohacik@vut.cz)

**Abstract**—Nowadays, there are many risks associated with computer networks, some of them can be eliminated with network probes. This paper is focused on the developed BUT network probe as a tool representing a hardware protection element of the network. Furthermore, the basics of IDS and IPS systems are described, including their possible applications. The basic concept of the network probe, the description of its basic parts and the created user interface are discussed. The last part is focused on the testing of hardware components that directly affect the proper functioning. The test results showed that the BUT network probe is able to perform network traffic analysis even at its maximum load.

**Keywords**— IDS, IPS, network analysis, network probe, Suricata

## 1. INTRODUCTION

With the development of computer technology and Internet technologies, the threat of remote attacks also began to develop [1]. This threat has resulted in the creation of specific organizations dealing with protection against computer attacks, such as anti-virus companies or state security authorities. Today, most attacks are carried out remotely using the Internet [2]. Therefore, it is often difficult to find the real attacker behind the attack. For these reasons, we use IDS (Intrusion Detection System) and IPS (Intrusion Prevention System) systems [2, 3, 4]. As the names suggest, these are systems that are used to detect or block specific network traffic [4]. Currently, there are many solutions to prevent a remote attack, but most solutions, such as the firewall, have a negative effect on the performance of the device itself [1, 2]. On the contrary, there are network devices that provide detection (or blocking) directly during data transmission, but most of these devices are expensive and mostly used only by large companies or organizations [4, 5].

Network probes can be a compromise [6]. These are devices that communicate with network devices (e.g. routers) and inform an authorized user (usually a network administrator) about certain events at the real time. This is the so-called real-time traffic analysis, which can prevent more serious damage to the monitored network [2, 4, 7]. Network probes can also combine various features of other systems, such as recording of transmitted data, traffic filtering, or analysis of specific protocols. Another advantage is the affordability for smaller companies or organizations [6]. Table I compares some existing solutions with comparable performance for analyzing, recording, or filtering traffic, including their pricing.

Device	Throughput	Data recording	Storage capacity	IDS	IPS	Price [€]
BUT network probe	1 Gb/s	✓	64 GB + external	✓	✓*	1 600
TAP Profishark 1G	1 Gb/s	✓	only external	✗	✗	2 800
Cisco FirePOWER 2110 NGFW	2 Gb/s	✓	100 GB	✓	✓	5 200
IPCopper models	1-12 Gb/s	✓	1-48 TB	✓	✓	5 200-23 400
Flowmon	SaaS**	✗	✗	✓	✓	9 000/year

\*Not fully implemented.

\*\*Software as a service – Baud rate depends on the infrastructure and the agreed price.

**Table I:** Comparison of some solutions for network analysis

## 2. NETWORK PROBE

A network probe (see Figure 1) was created at the Department of Telecommunications for teaching and testing purposes. This probe is a tool for detecting cyber and security threats in real time. It can be used to protect the internal network terminal equipment or to analyze network traffic at a given point. The detection of defined attack signatures is used for the analysis of network communication, but the software solution enables the analysis of suspicious activities, so the probe is able to recognize even unknown types of attacks. The probe also includes an external hard disk for storing recorded data using the implemented T-shark module.

A single-board computer called 4x4-4500U from ASrock-Industrial was chosen for the network probe. The device has two Realtek R8111FPV (enp1s0) and Realtek RTL8125BG (enp2s0f0) network cards used as the primary interface for traffic analysis. Among other things, it also contains two USB connectors of type C, which support the connection of powerful external network adapters for possible needs of processing traffic from multiple inputs. Due to the great flexibility of settings, the operating system Ubuntu 20.04 was chosen. Traffic filtering is implemented by using IPtables and its successor NFtables. These tools allow you to create specific rules that determine which data a device should receive, discard, or further process. Signatures are detected using the implemented NIDS system [3, 4, 7].

The internal block diagram of the probe is shown in Figure 2. All data traffic intended for processing is sent to the Suricata module. It analyzes individual messages and compares them with the internal signature database that is a part of the web application. The continuous module stores action logs, which are further operated by the web server.



Figure 1: BUT network probe

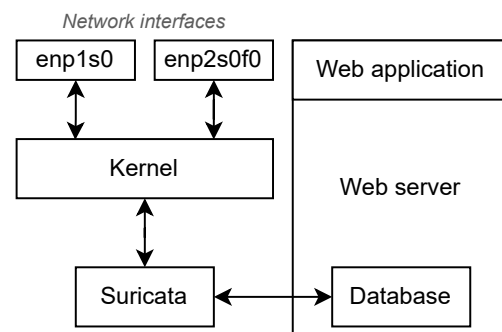


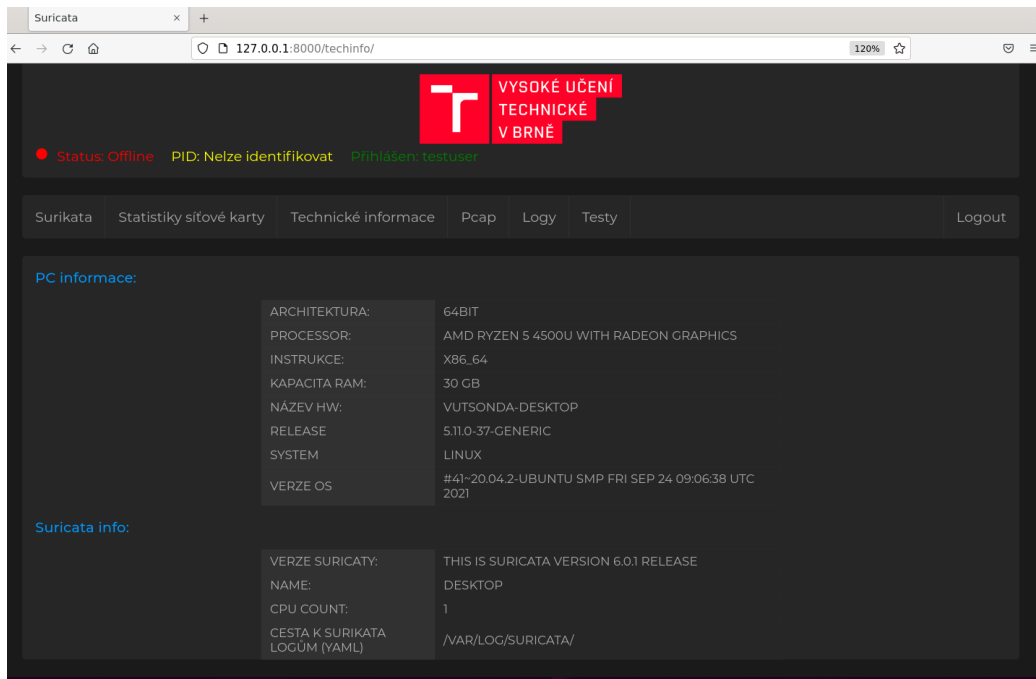
Figure 2: Block diagram of the network probe

The network probe has a multi-platform web interface (see Figure 3), which allows the user to control and configure the device (e.g. add rules, delete logs, start traffic logging, etc.). The basis of the web interface is the open-source web framework Django. A characteristic feature of this framework is the Model-View-Controller architecture. The model represents a database of user data, the view represents a graphical representation of the control elements and the controller is used to control individual elements. The interface consists of six basic parts and contains modules that provide the required functionality. The interface is used to communicate with the Ubuntu OS, create individual processes, or rewrite Suricata configuration files, which is used for network traffic analysis.

In addition to the basic ability of the network probe to process network traffic according to the set rules, the Scappy library was implemented in the probe. It is used for detection and partly analysis of various industrial protocols. Supported protocols that can be detected by the probe include IEC 60870-5-104, Device Language Message Specification (DLMS), protocols from IEC 61850 standard (GOOSE and Sample Value). These protocols were chosen because of their frequent deployment in practice.

## 3. PERFORMANCE TESTING

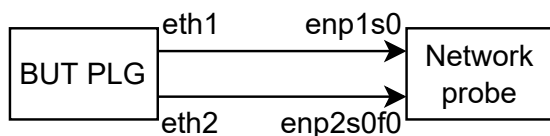
The functionality and usability of the probe itself depends on the individual network cards. For these purposes, the RFC 1242 and RFC 2544 standards were used, which define the methodology for testing



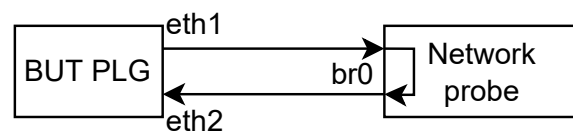
**Figure 3:** Web application – Information page

network parameters (e.g. throughput, latency, etc.). The aim of the tests was primarily to verify hardware limits of the device, which include: maximum processing speed of input/output data, load of operational memory or processor.

The internal BUT payload generator (PLG) was used to test the network probe. This generator contains an Intel Xeon E5-2650 v4 @2.2 GHz CPU with 128 GB of RAM and an NVIDIA GeForce GT 730 VGA graphics card. It also includes 5 network interfaces, 2 x 1 Gb/s (on the motherboard), 2 x 10Gb/s (Intel X540-AT2) and 2 x 10 Gb/s (Intel 82599ES SFI/SFP+). The whole device runs on CentOS7 operating system. Due to the nature of testing, this generator was set for all tests to generate TCP packets at a rate of 200 000 packets/s (so-called TCP Flood). The size of one packet was set to 1 460 B. Such a setting generates a load of approximately 2.336 Gb/s for one network interface. Two connection schemas were used for the testing. The schema shown in Figure 4 was used to test the hardware components. The schema in Figure 5 shows the so-called bridge mode, which was used to test the probe's ability to detect the set Suricata rules. This mode software combines *enp1s0* and *enp2s0f0* interfaces into one virtual *br0*.



**Figure 4:** Performance testing schema



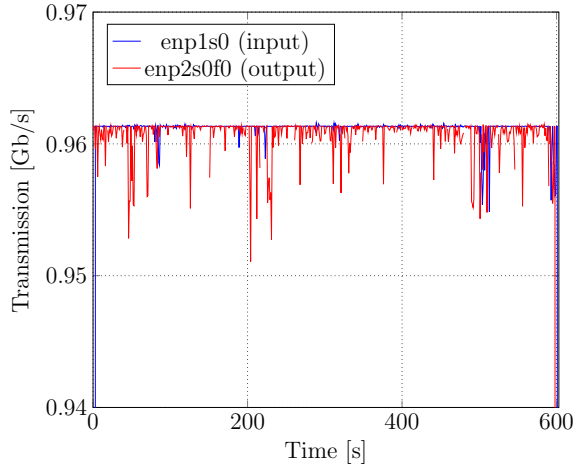
**Figure 5:** Bridge mode schema

### 3.1. Hardware components test

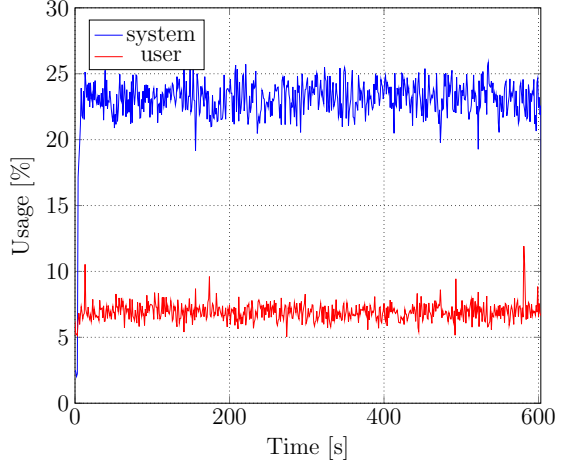
The first set of tests was focused on the performance of the individual components of the network probe. These were mainly the CPU (Central Processing Unit), RAM (Random Access Memory) and network interfaces. These components directly affect the usability of the network probe. Testing was performed on the schema shown in Figure 4. In order not to affect the results, all unnecessary processes in the network probe were switched off. The test was performed 3 times in 10 minute intervals. The results for all 3 intervals were very similar. For clarity, Figure 6 and Figure 7 show data from the third 10-minute interval of the test.

As expected, the maximum processing speed of incoming data using network interfaces was approx. 0.96 Gb/s ( $\approx 120$  MB/s), as these are Gigabit interfaces. However, the CPU load throughout the testing

anged from 20 to 25 %. As for the use of RAM, the load here was in the order of percent. It is therefore clear that the limiting hardware component is the network card itself, as the other components did not reach even half of their maximum load.



**Figure 6:** Usage of Ethernet interfaces during performance test

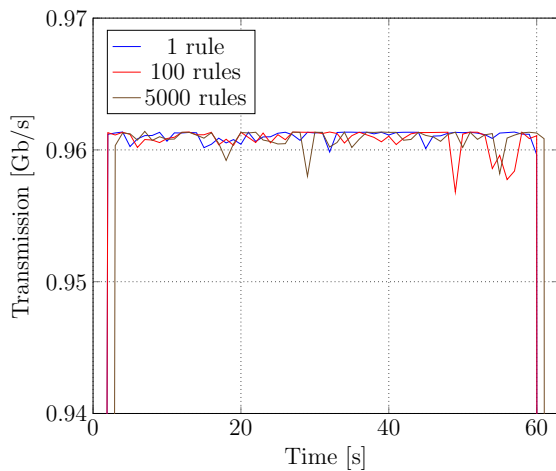


**Figure 7:** CPU usage during performance test

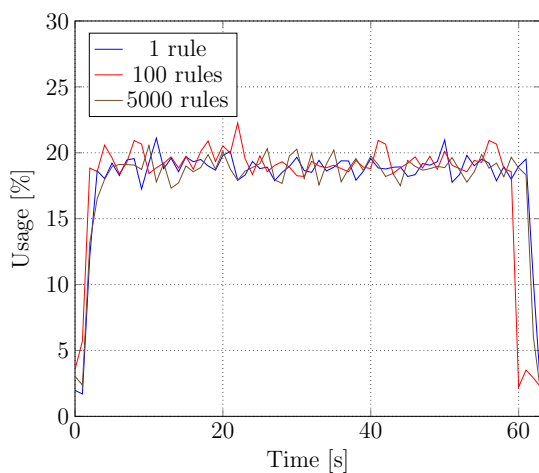
### 3.2. Rule detection test

The second set of tests focused on the ability of the network probe to evaluate the rules set in the Suricata module. As in the first set of tests, they focused on the usage of hardware components (CPU, RAM and network interfaces). Testing was performed on the schema shown in Figure 5. The network probe was set to the so-called bridge mode, which represents a transport connection, where traffic from one network interface is forwarded to another one.

The test set consisted of 3 times 6 measurements, where each measurement contained a different number of Suricata rules (1, 5, 100, 500, 1 000 and 5 000). Due to the larger number of tests, the time interval for this testing was reduced to 1 minute. Again, the most important parameters were the load on the network interfaces and the CPU load caused by the system. Due to the consistency of all recorded results and for clarity, only the data for the third set of tests for 1, 100 and 5 000 rules are shown in Figure 8 and 9. The results show that even for 5 000 rules, there is no noticeable increase in CPU usage. The average value ranged from 17 to 20 %, with a maximum of 22.5 %. As in the previous set of tests, the RAM load is in the order of percent. It is clear that the number of rules does not significantly affect the functionality of the probe.



**Figure 8:** Usage of Ethernet interface *br0* during rule detection test



**Figure 9:** CPU usage during rule detection test

The delay and latency (so-called jitter) of the transmitted data passing through the network probe was also measured. The *Iperf* tool and the schema in Figure 5 were used for this testing. The results are displayed in Table II. The average delay caused by the network probe is around 0.007 ms. This value does not represent a significant deterioration in the connection properties. In the case of jitter, the values were 6 times higher when the mains probe was fitted. This is due to the different congestion densities of the network adapter when received packets are processed. A jitter value of 0.038 ms is not a problem for common networks.

Path	Minimum [ms]	Average [ms]	Maximum [ms]	Jitter [ms]
Not over network probe	0.167	0.184	0.251	0.007
Over network probe	0.217	0.257	0.410	0.038

**Table II:** Delay and jitter caused by the network probe

#### 4. CONCLUSION

Nowadays, the issue of network security is a constantly discussed topic. On one hand, there is an effort to develop new technologies, but on the other hand, new technologies contain vulnerabilities that represent potentially unprotected areas of the system. To improve the security, for example in a corporate network, we can use the so-called network probes. This paper discusses the developed BUT network probe, which combines features of devices designed for analysis and filtering of network traffic. For easier user control, a web interface has been implemented in the network probe. This allows, among other things, setting up the network probe as IDS/IPS. For better analysis and recording of transmitted data, the T-shark module was implemented in the network probe, which creates a record of traffic in PCAP format. The last special feature of the network probe is the ability to analyze special protocols that are used in the industry, such as DLMS, IEC 60870-5-104, GOOSE and Sample Value protocols from the IEC 61850 standard. The paper also describes the tests used in verifying the functional characteristics and capabilities of the network probe. These tests have shown that the BUT network probe is capable of handling network traffic up to 1 Gb/s without significant degradation of traffic properties. Future work on the network probe will focus on expanding the supported industrial protocols, deploying external network cards for higher data rates using the USB-C interface, and adding new capabilities within the web interface.

#### REFERENCES

- [1] H. Wang, B. Han, J. Su, and X. Wang, "A high-performance intrusion detection method based on combining supervised and unsupervised learning," in *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)*. IEEE, 2018, pp. 1803–1810.
- [2] I. Mukhopadhyay, M. Chakraborty, and S. Chakrabarti, "Hawkeye solutions: a network intrusion detection system," in *Proceedings of the International Conference & Workshop on Emerging Trends in Technology*, 2011, pp. 252–257.
- [3] M. M. Aravind and V. Kalaiselvi, "Design of an intrusion detection system based on distance feature using ensemble classifier," in *2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*. IEEE, 2017, pp. 1–6.
- [4] S. F. Diyar and A.-K. Samir, *Automatic Intrusion Prevention Technique to Improve Network Security: Intrusion Detection System (IDS), Intrusion Prevention System (IPS) and Firewall to secure the network*, 1st ed. LAP LAMBERT Academic Publishing, 2017.
- [5] E. D. Knapp and J. Langill, *Industrial Network Security: Securing critical infrastructure networks for smart grid, SCADA, and other Industrial Control Systems*. Syngress, 2014.
- [6] D. Hirš, "Proposal of cyber threat detector using raspberry pi," in *Proceedings I of the 27th Conference STUDENT EEICT 2021: General papers*. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2021, pp. 192–195. [Online]. Available: <http://hdl.handle.net/11012/200741>
- [7] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Practical tools for attackers and defenders," in *Network Traffic Anomaly Detection and Prevention*. Springer, 2017, pp. 201–242.

# Pilot testing of LoRaWAN technology in disconnecting fuse boxes

L. Beneš<sup>1</sup>, P. Mlýnek<sup>1</sup> and D. Harman<sup>2</sup>

<sup>1</sup>Brno University of Technology, Technická 12, Brno, Czech Republic

<sup>2</sup>EG.D Distribuce, F.A. Gerstnera 2151/6, České Budějovice, Czech Republic

E-mail: [xbenes44@vut.cz](mailto:xbenes44@vut.cz), [mlynek@vut.cz](mailto:mlynek@vut.cz) [dominik.harman@egd.cz](mailto:dominik.harman@egd.cz)

**Abstract**—This article deals with the concept of Smart Cities concerning Smart Grids. As part of Smart Grids and secondary metering, there is also the lifespan of the devices that are deployed in distribution companies. This also applies to switch cabinets, disconnecting fuse boxes (DFB), and other elements that are made of a material that can degrade due to adverse weather conditions. For this reason, secondary temperature and humidity measurements have been introduced. During the measurements, temperature extremes were recorded, which could reach more than -10 °C in winter months and almost 45 °C in the warm months. Humidity could reach up to 100 % in some cases. Another monitored parameter was the successfully received messages, where LoRaWAN technology reached almost 95 %.

**Keywords**— Smart Grids, LoRaWAN, temperature and humidity measurement, weather conditions.

## 1. INTRODUCTION

The Smart Cities concept is a way to sustainable urban development, with the help of modern technologies, it is possible to improve the quality of life and streamline the management of the trust. The greatest application of the concept is mainly in the field of transport and energy, but partial use in the areas of waste management, water management, e-government, and crisis management. For the population, it means a better quality of life, energy savings, faster and safer transport, more nature in cities, smart infrastructure solutions, fast communication with the authorities, an ecological environment, and space for future generations to live.

However, this concept must follow a lot of partial steps. Especially in the field of energy, in the Smart Grid, which is already followed by the concept of Smart Metering [1].

Smart Grid networks should make it possible to regulate the production and consumption of electricity in real-time. The basic principle of the Smart Grid is mutual two-way communication between electricity generation sources and appliances or consumers about the immediate possibilities of energy production and consumption. Smart Grid has three basic features:

- Full automation,
- full integration of customers,
- adaptation to different ways of generating electricity.

Full automation means the connection of a digital control and management system together with sensors that control the behavior of the network and automatic resumption of operation after a possible failure. Thanks to full automation, information on network load, quality of electricity supply, supply interruption, etc. is available in real-time [2].

The essence of full customer integration is to equip customers with digital meters with a two-way flow of information in real-time, which allows the creation of price tariffs (ie. price lists and conditions of their use) according to the current situation in the network [3]. In this context, there is usually talk of the Smart Metering system. Full integration allows customers to efficiently manage consumption, such as water heating, laundry, or battery recharging at times of free production capacity.

Adaptation to different ways of generating electricity enables the connection of, for example, solar and wind power plants, gas microturbines, and other decentralized production technologies, which allows customers to produce electricity from their resources and sell its surplus to the grid.



Communication between production and consumption usually takes place over separate data networks. This requires, on the one hand, a high degree of standardization for the tools and formats of the data transmitted, and, on the other hand, ensuring data security, both against the consequences of data network failures and against possible unauthorized use.

It is possible to match the supply of electricity in the grid with immediate demand in a short time, Smart Grids are a prerequisite for mass renewable energy sources where, due to geographical and climatic conditions, production from these sources is irregular or difficult to predict or does not overlap with requirements for consumption. This applies, for example, to solar or wind energy in Europe.

Smart Grids are therefore considered an important tool for achieving the EU's environmental goals by 2020 compared to 1990, the so-called goals (20/20/20):

- 20 % reduction in CO<sub>2</sub> emissions,
- 20 % increase in energy efficiency,
- 20 % use of renewable energy sources.

By 2020, 80 % of households in the EU [4] are expected to be equipped with Smart Meters for this purpose. In addition, Smart Grids offer overall energy efficiency by enabling efficient matching of production and consumption with the lowest possible operating costs. Thus, for example, they allow appliances that can be switched on at any time of the day (for example heating) to be switched on at the very moment when unused electricity sources are available.

At the same time, the Smart Grids system also places demands on technology in terms of complexity and reliability in the interaction of individual elements, ie. devices or data inputs and outputs. If in a "classical" energy network, system integration (ie. harmonization of flawless operation) involves less than a thousand such elements, there are many millions of elements in Smart Grids.

Today, the practical application of Smart Grids is limited to selected projects in defined regions, where the technical and operational characteristics of their operation are tested.

The Smart Metering system is needed for the Smart Buildings concept to exist, as well as Smart Grids and of course Smart Cities. You could say that Smart Metering is the cornerstone of other concepts. The idea of Smart Metering is in principle very simple. This is a reading of consumption points with the transmission of consumption information to the data center in real-time or at short time intervals. Based on the evaluation of the obtained data, there will be feedback from the data center back to the Smart Meter with the appropriate command. After receiving the command, the Smart Meter performs the required action. The actions and commands that the Smart Meter can perform depend on its equipment or agreement with the relevant consumer. The consumer is informed about the current situation, commands, and needs of the network via a home display module, on a special website, or in a mobile application. The historical predecessor was the technology of bulk remote reading (HDO), which after a certain command activated water heating or a command to change the tariff. A large number of new Smart Meters will have to be deployed in the future roll-out, which is set for 2024 [5].

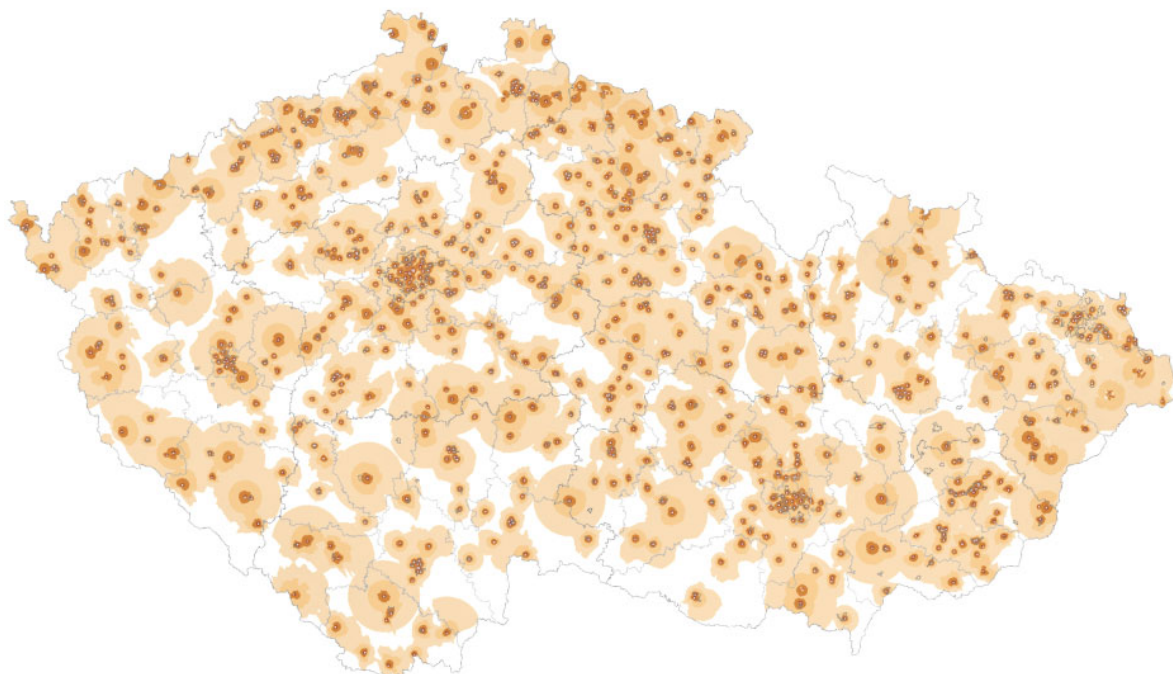
## **2. DESCRIPTION OF TECHNOLOGY**

Czech distribution companies manage a huge number of disconnecting fuse boxes, which are used for disconnection, branching, continuous connection, or protection of low voltage distribution systems. These DFB are often exposed to bad weather conditions, especially in high or low temperatures and humidity. Manufacturers often do not state for which range of ambient temperatures they should be used and also what maximum humidity they can be exposed to. Over time, UV radiation will also affect the technical condition of DFB. Very often, these DFB are made of a thermoset composite material based on unsaturated polyester resins reinforced with glass fibers. These fibers crack over time and there are leaks in the casing and nothing prevents the moisture. Distribution companies want warranty repair of the DFB or to change problematic piece for a new one, but they cannot prove that the weather conditions will be responsible for this, which will reduce the stated lifespan by several years. It was for this purpose that temperature and humidity measurements were made.

Continuous measurement was not necessary and once every four hours would be sufficient. In DFB, it was not desirable for the devices to be powered directly in the DFB. Therefore, battery power was chosen. LoRaWAN (Long Range Wide Area Network) technology best suited these conditions.

LoRAWAN is a reliable and secure network for wireless connection to the Internet of Things. And at long distances up to 15 km [3]. It can securely provide two-way communication at data rates from

0.3 kbps to 50 kbps. Another advantage is that LoRaWAN networks in Europe operate on the 868 MHz frequency, which falls within the unlicensed ISM (Industrial, Scientific, and Medical) band. There is also an extensive public network operated by České radiokomunikace a.s. As can be seen in the Figure 1, the coverage of LoRaWAN networks is quite extensive and it is sufficient for testing the technology. The localities Brno and Brno-venkov were selected for testing.



**Figure 1:** Map of the coverage of the Czech Republic by LoRaWAN networks [6]

LoRaWAN network architecture consist of [7]:

- End nodes - sensors/actuators that read data or control other devices
- gateways (GW) - devices receiving and forwarding data from sensors to the network server
- network server (NS) - ensures coordination of the entire LoRaWAN network
- application server - hosts an application that processes the measured data from sensors, or sends commands to actuators.

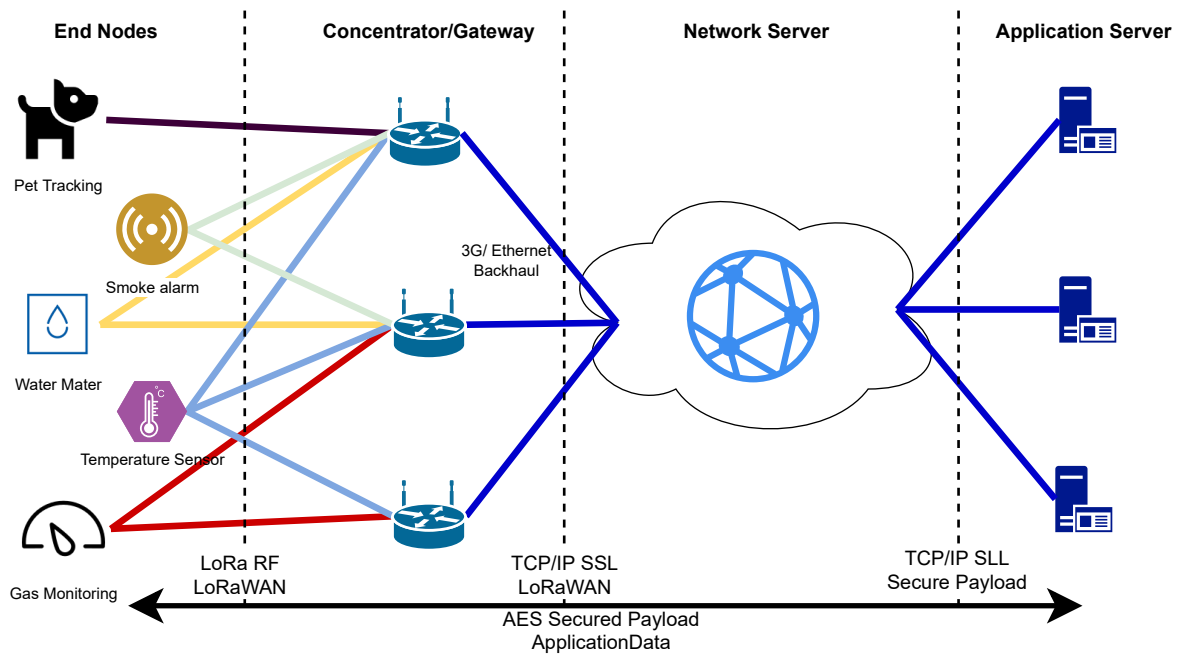
End devices, such as sensors, connect to gateways, which are connected via Ethernet or another communication medium to a network server, which provides coordination throughout the network (see Figure 2). This communication is already controlled by TCP/IP using Secure Sockets Layer (SSL). The network server then passes the acquired data to the application server, which processes and visualizes the data from the sensors. It can also send commands back to sensors. All data flow is secured by Advanced Encryption Standard (AES) in both directions of communication [8, 9] .

**How LoRaWAN works:** The LoRa physical layer uses Spread Spectrum Modulation (CSS), or its variant over CSS (Chirp Spread Spectrum).

When using CSS modulation, a so-called chirp signal is generated, which increases its frequency linearly with time from the lower limit of the band to the upper limit of the band (chirp-up), or decreases its frequency (chirp-down). When the upper limit is reached, the frequency returns to the lower limit and the process is repeated. The transmitted information is then modulated on this signal by chirp [7].

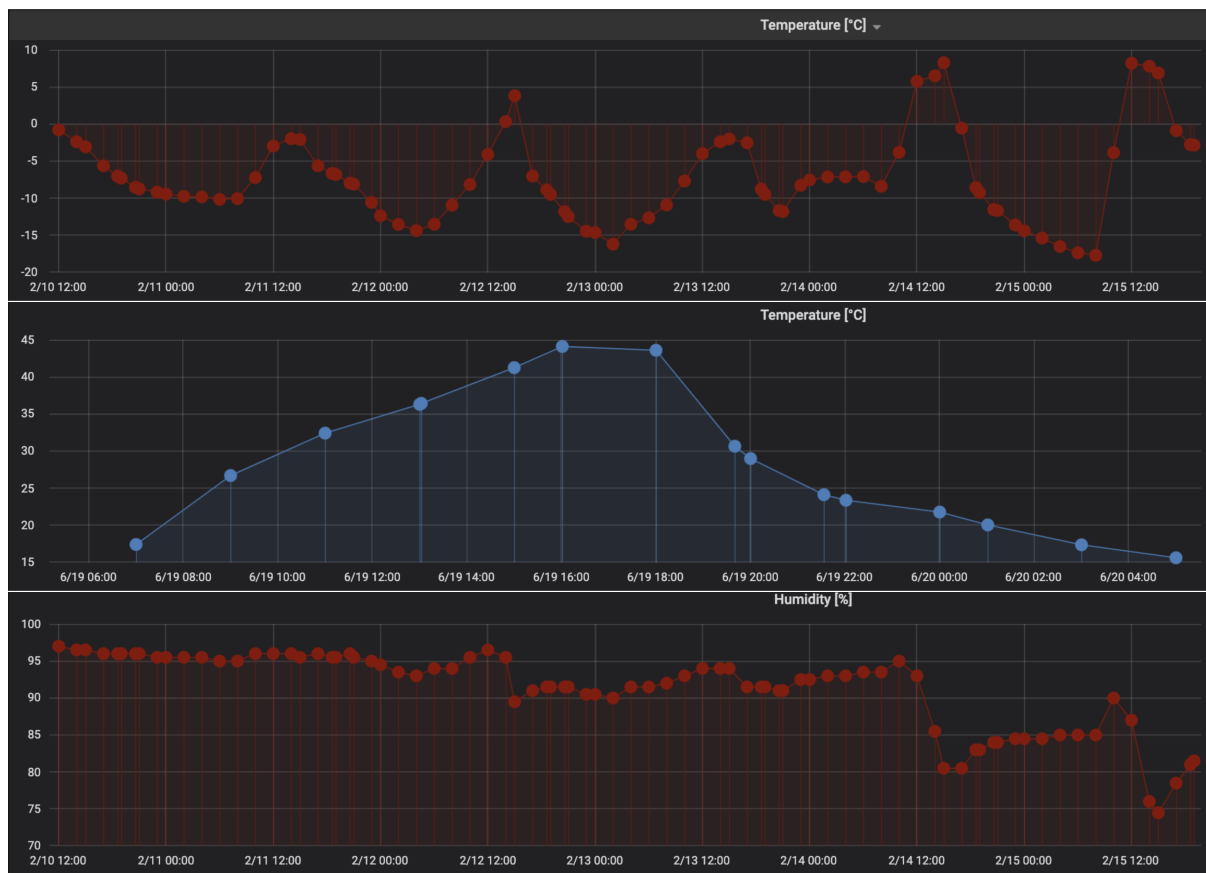
### 3. TEMPERATURE AND HUMIDITY MEASUREMENT IN DFB

As already mentioned, testing takes place in the vicinity of the city of Brno, where several DFB are monitored. This testing has been in operation for more than a year, so temperature and humidity extremes can be seen that can permanently damage DFB. The temperature range throughout the year was from -18 °C to 44 °C. The annual average of all sensors was around 10–20 °C. The humidity range was recorded



**Figure 2:** LoRaWAN network architecture

from 3 % to 100 %, where the annual average humidity was around 70 %, one sensor recorded an annual average of 85 %. All extreme values are shown in the Figure 3.



**Figure 3:** Extreme values of temperature and humidity in switch cabinets

Due to the location of some sensors, it was possible to note that some sent messages did not reach the backend system for visualization. The network was probably too busy at the time of the signal was not very strong to reach the Gateway. Correctly received messages out of the total number of sent messages

are mainly affected by the location as well as the distance between the sensor and the GW. Correctly received messages ranged from 80–94 %.

#### 4. CONCLUSION

Due to the requirements of the contracting authority, which was a battery-powered device that will be able to measure temperature and humidity at least every four hours, LoRaWAN communication technology was selected and tested. The devices were then placed in DFB around the city of Brno. After year-round measurements, it was found that more than half of the devices reach humidity of up to 100 %. It was also found that in winter the temperature can drop to almost -20 °C and during the warm summer months to almost 45 °C.

In the case of free-standing DFB, which are not protected by shade in the summer months, the primary casing may begin to degrade very soon. This degradation is probably caused by UV radiation hitting the DFB. The DFB cover can be damaged and, in extreme cases, can burst. This causes deformation of the cover and can rain or moisture easily gets into the DFB.

Measurement of LoRaWAN technologies inside DFB can demonstrably detect damaged cover due to degradation, or it can be warranty repair a poor quality piece from production or to change problematic piece for a new one.

#### ACKNOWLEDGMENT

I would like to thank Mr. Dominik Harman for his cooperation and for providing the measurement data for the elaboration of the article at the EEICT 2022 conference.

#### REFERENCES

- [1] M. Curiale, "From smart grids to smart city," 2014 Saudi Arabia Smart Grid Conference (SASG), 2014, pp. 1-9, doi: 10.1109/SASG.2014.7274280.
- [2] F. Abate, M. Carratù, C. Liguori, M. Ferro and V. Paciello, "Smart meter for the IoT," 2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2018, pp. 1-6, doi: 10.1109/I2MTC.2018.8409838.
- [3] G. C. Lazaroiu and M. Roscia, "Model for smart appliances toward smart grid into smart city," 2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA), 2016, pp. 622-627, doi: 10.1109/ICRERA.2016.7884410.
- [4] Nová vyhláška o měření elektřiny č. 359/2020 Sb. MPO [online]. 2021 [2022-03-13]. Available from: <https://bit.ly/3MFDBoa>.
- [5] Implementace inteligentního měření v PREdistribuci, a.s. ENERGY-HUB [online]. Praha, 2021 [2022-03-13]. Available from: <https://bit.ly/37lcMFz>.
- [6] Mapy pokrytí. IoT Portál [online]. [2022-03-13]. Available from: <https://bit.ly/3JbPaRT>.
- [7] IOT TECHNOLOGIE: LORA A LORAWAN. Eman [online]. 2019 [2022-03-13]. Available from: <https://bit.ly/3I7wO36>.
- [8] E. Sisinni et al., "LoRaWAN Range Extender for Industrial IoT," in IEEE Transactions on Industrial Informatics, vol. 16, no. 8, pp. 5607-5616, Aug. 2020, doi: 10.1109/TII.2019.2958620.
- [9] Veřejné/privátní LoRaWAN sítě [online]. Ostrava [2022-03-13]. Available from: <https://bit.ly/3JbbXxk>.

# Introduction of Cyber Arena

T. Stodulka, Z. Martinasek

Department of Telecommunications, Faculty of Electrical Engineering and Communication,  
Brno University of Technology, Technicka 12, 616 00 Brno, Czech Republic

E-mail: [xstodu07@vut.cz](mailto:xstodu07@vut.cz), [martinasek@vut.cz](mailto:martinasek@vut.cz)

**Abstract**—This article describes the current state of a private cyber range named Cyber Arena that is developed by our university. First chapter briefly explains the current problematic and describes the sequence diagram of exemplary communication between the main Cyber Arena components that are OpenStack, ansible AWX, and web-based Training application. Second chapter focuses on Arena architecture and the following chapters briefly explain all mentioned main Arena components.

**Keywords**—AWX, Cyber Arena, cyber games, cyber range, OpenStack

## 1. INTRODUCTION

Nowadays, cyber systems and communication networks face many types of threats from the point of view of cyber security, such as ransomware, malware, computer worms, spams, distributed attacks, etc. Recently, they are also developing very rapidly and their growth will continue at the same or even higher pace. No one is immune to cyber attacks from government organizations to smaller family businesses. Therefore, it is necessary to educate both the public and professional experts to be able to map potential attacks, prepare for their consequences, and generally know how to limit such security risks. These purposes are greatly fulfilled by the cyber range [1, 2] – a platform responsible for the complete preparation of various simulated scenarios where users are exposed to diverse challenges.

As these platforms are usually expensive [3, 4], the Department of Telecommunications (UTKO) is working on its own platform called **Cyber Arena**. This platform will be used in the teaching of several university courses in cyber security education. At present, the software part of the Cyber Arena consists of 3 components that run on 2+ $x$  servers, where  $x$  represents all OpenStack servers: (1) OpenStack [5, 6, 7], running on physical servers; (2) AWX [8, 9], running in a virtual machine (VM) in the OpenStack; (3) training web application (a combination of user/admin interface for control and administration of Cyber Arena and playing cyber security games) with relational database [10], running again in VM in the OpenStack. Example of communication between Arena components is shown in the figure 1. Admin or student logs in the Training application, while the student is only allowed to play games and view

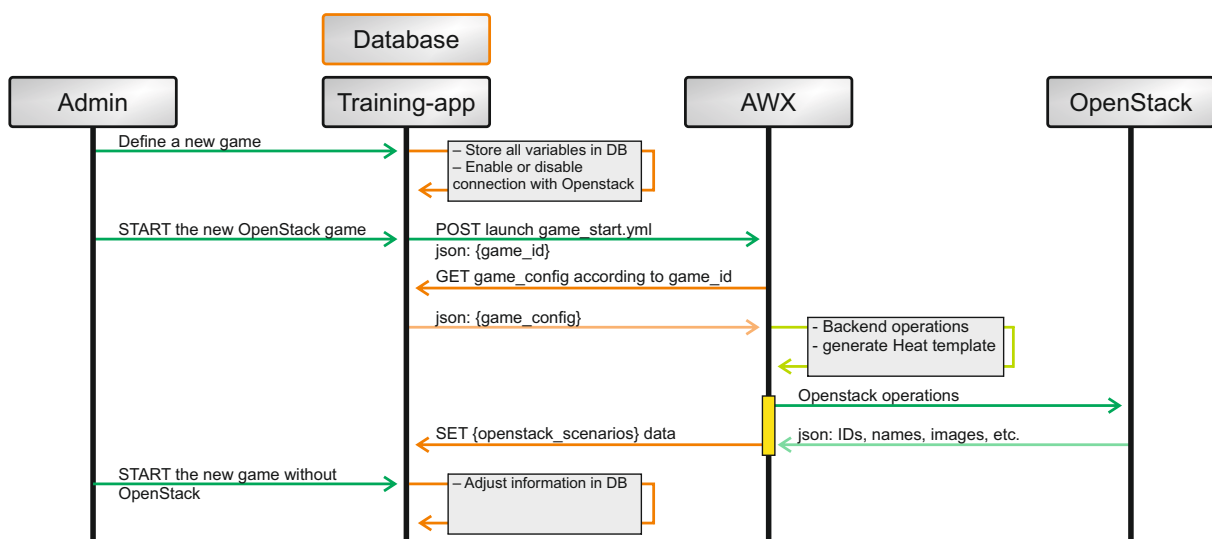
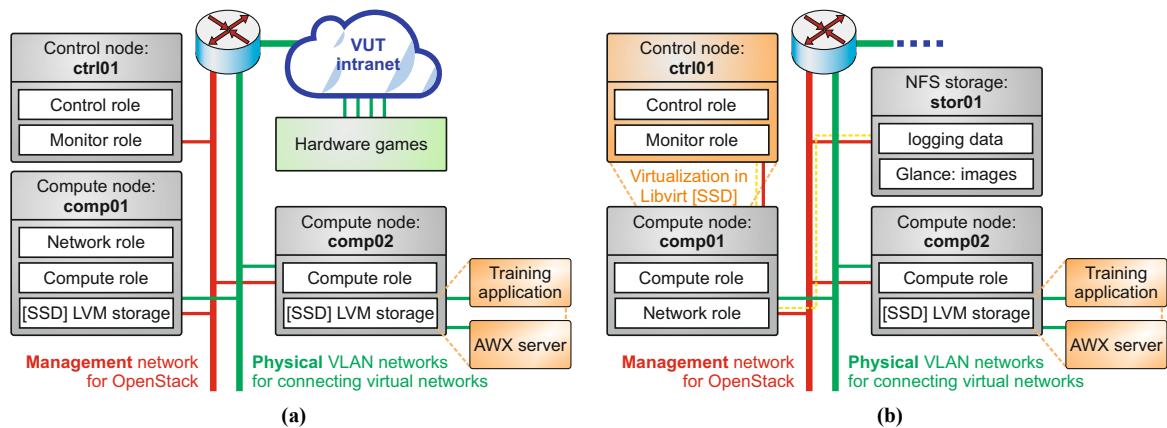


Figure 1: Sequence diagram of exemplary communication between the Cyber Arena components

their statistics. Admin then has access to more advanced options, eg turns on/off the newly defined game. Games can be independently playable (working with downloaded attachments, etc.) or dependent on OpenStack, where a separate virtualized environment is already being created for each student. AWX in the Cyber Arena is an intermediary between the Training application and OpenStack.

## 2. CYBER ARENA ARCHITECTURE

The architecture of Cyber Arena has undergone one major update, that is visualized in figure 2. In the legacy architecture, all OpenStack nodes were deployed on baremetal servers and critical Arena applications (AWX and Training application) were running on the SSD disks of compute nodes. The size of SSD disks in compute nodes (**comp0x**) was not fully utilized and OpenStack controller (**ctrl01**) had considerable performance issues – during the virtualization of 30+ scenarios, the controller was starting to be unstable due to weak processor and slow speed of HDDs. In the new architecture, **ctrl01** is virtualized on **comp01** and the original controller will be used as NFS (Network File System) storage (**stor01**) for **ctrl01**, as the controller still needs to store relatively large amounts of data. The remaining disk capacity of the **stor01** will be used to create periodic backups of the entire Cyber Arena. Moving OpenStack controller to VM brings many advantages, mainly: easy backup and recover, easy portability, and scalability of HW resources.



**Figure 2:** Block diagram of the old (a) and the current cloud of the Cyber Arena (b)

Each OpenStack node performs important roles, server **ctrl01** controls the whole OpenStack and additionally has implemented intelligent monitoring using Grafana and Prometheus exporters (node, openstack) and central logging of OpenStack services using Kibana and ElasticSearch. Compute nodes virtualize game scenarios and server **comp01** additionally manages virtual networking for all instances (VMs and containers). Physical and virtual networks are interconnected through VLAN networks that are mapped into OpenStack. Each VLAN network has its purpose: one VLAN is used exclusively for Internet connection (public network) and others VLANs connect instances to hardware games SCADA/ICS (Supervisory Control And Data Acquisition/Industrial Control System).

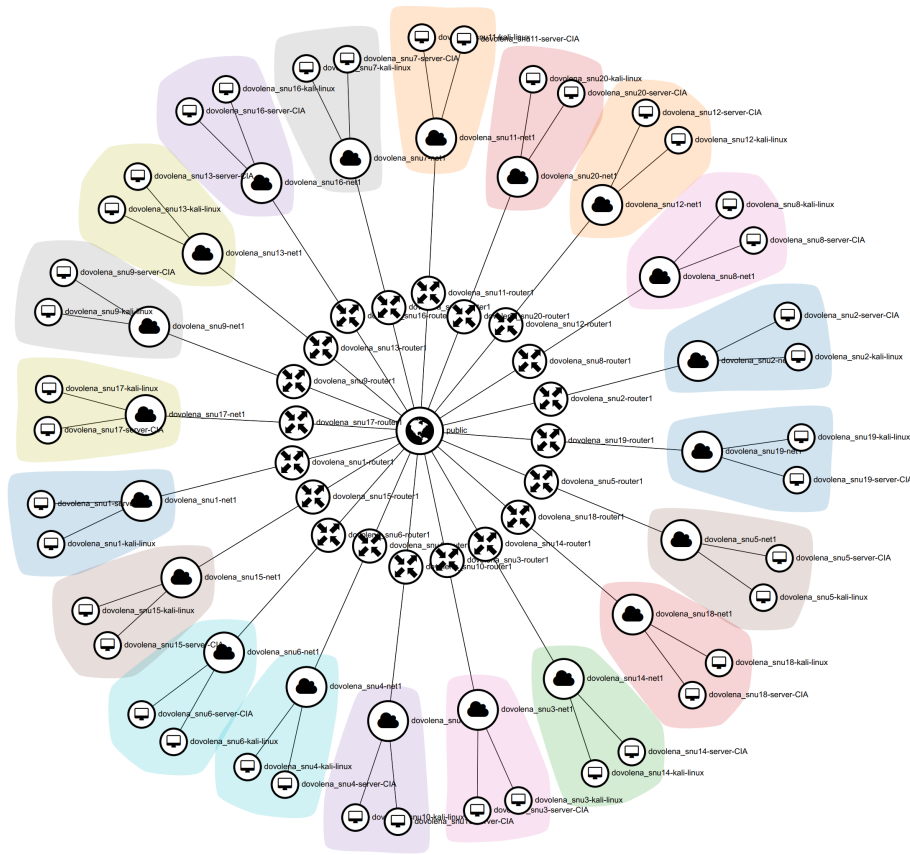
## 3. OPENSTACK HEAT AND HW CAPABILITIES

Heat [11] is an orchestration tool for automating complex processes in OpenStack. Using this service, OpenStack creates scenarios for players in the Cyber Arena. Generally, it applies that 1 stack equals to 1 scenario for 1 student – then each student has its own virtual environment. Heat template is generated using jinja2 in YAML format. Template generation is defined based on the game configuration, which is inserted during the game definition in the Training application.

There are currently 3 games in the Cyber Arena ready for playing, where 2 of them are OpenStack dependent. An exemplary topology of a virtualized game “Dovolena snu” (Dream vacation) for 20 students is shown in the figure 3. It contains Kali linux servers, and vulnerable CIA websites, where each server is located in its own virtual network. Both networks are connected by a virtual router, which is the only one connected to the physical network public.

The total capacity of computing resources is currently: 512 CPU cores, 512 GB RAM and storage 11.45 TB HDD + 0.45 TB SSD (for critical applications). Speaking about the game Dream Vacation, its





**Figure 3:** Virtual topology of the game Dream vacation in OpenStack

servers need for 1 scenario: 8 CPU cores, 8 GB RAM and 120 GB HDD – that ideally corresponds to a total of 57/57/95 scenarios (taken into account: critical applications and virtualized OpenStack Controller).

#### 4. ANSIBLE AWX

AWX is a community version of Ansible Tower, that is developed and managed by Red Hat. Basically, it is extended Ansible, which has API interface implemented and so can be easily managed remotely. Ansible is an automation tool composed of **playbooks** – a component that combines complex and simple **tasks** performed on target servers. In addition, tasks can be grouped into **roles**, that are then called from playbooks. The roles can then be repeated in other playbooks. It is mainly a matter of keeping the code redundancy as small as possible.

AWX in the Cyber Arena is an intermediary between the Training application and OpenStack for games that require the creation of virtualized environments for students. Four playbooks are currently implemented in AWX:

- **Game start** – turning on the game in the Cyber Arena by the admin. Briefly, it creates a new OpenStack project with quotas, privileges, firewall, etc. Next, it generates authentication files and Heat templates according to the loaded configuration from Arena database, validates the template, creates the Heat stacks and collect all useful data about scenarios and stores it in database.
- **Game shutdown** – turning off the game in the Cyber Arena by the admin. The playbook loads all scenarios from Arena database, deletes them and release all reserved HW resources.
- **Student start** – turning on the game by the player. Reserves the free scenario by user ID, loads the scenario data and generates VNC (Virtual Network Computing) console URLs for player.
- **Student end** – turning off the game by the player/admin/system. Releases the scenario reservation and rebuild all VMs using fresh images.

## 5. TRAINING APPLICATION FROM USER VIEW

Training application is web-based user/admin interface to play, manage, and monitor cyber games. Currently, the login is authenticated through Microsoft Office 365, so the users don't have to create new accounts and can use the same credentials as for VUT. The user has to be assigned to some lecturing group (typically some university courses) and after that he/she should be able to see the provided cyber games on the home page of Training application, see figure 4.

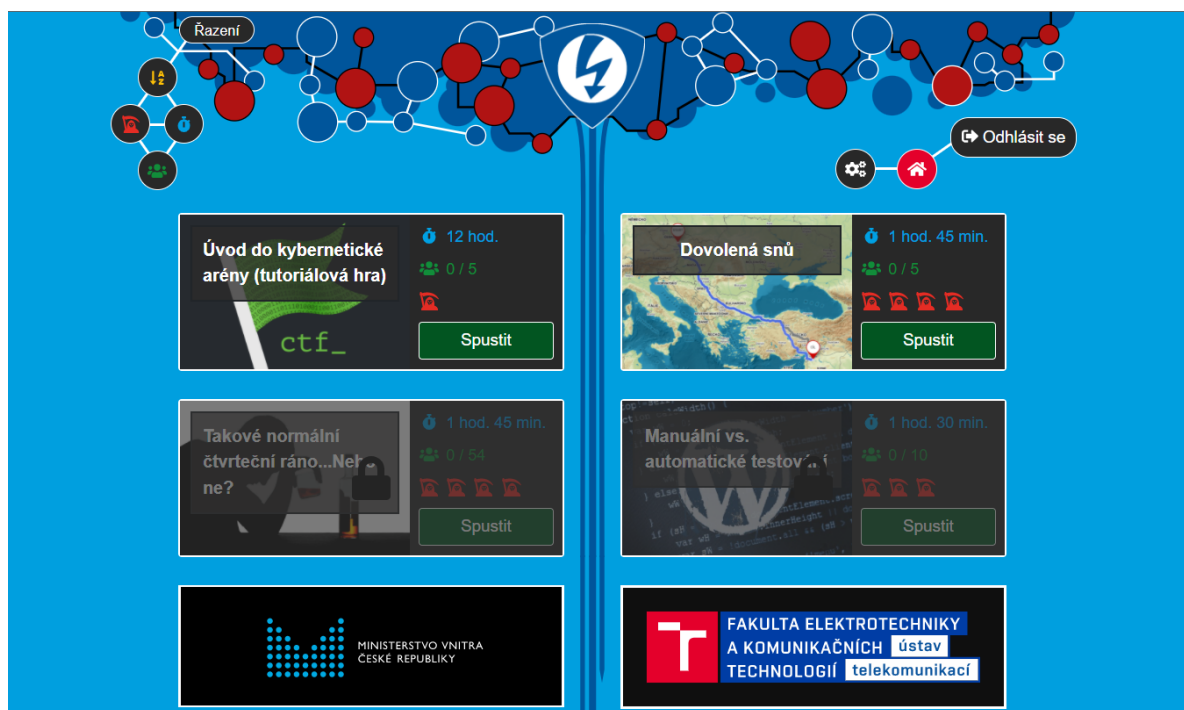


Figure 4: Home page of Cyber Arena with provided cyber games

Each cyber game composes of the following parts: **(1) tasks** – text guide, helps, window for taking notes, buttons for opening consoles; **(2) test** – typically 5 questions with one or multiple correct answers; **(3) summary** – detailed evaluation and the total number of points obtained; **(4) scoreboard** – display the score of all active players in graph and table format. An example from playing game Dream vacation is shown in the picture 5.

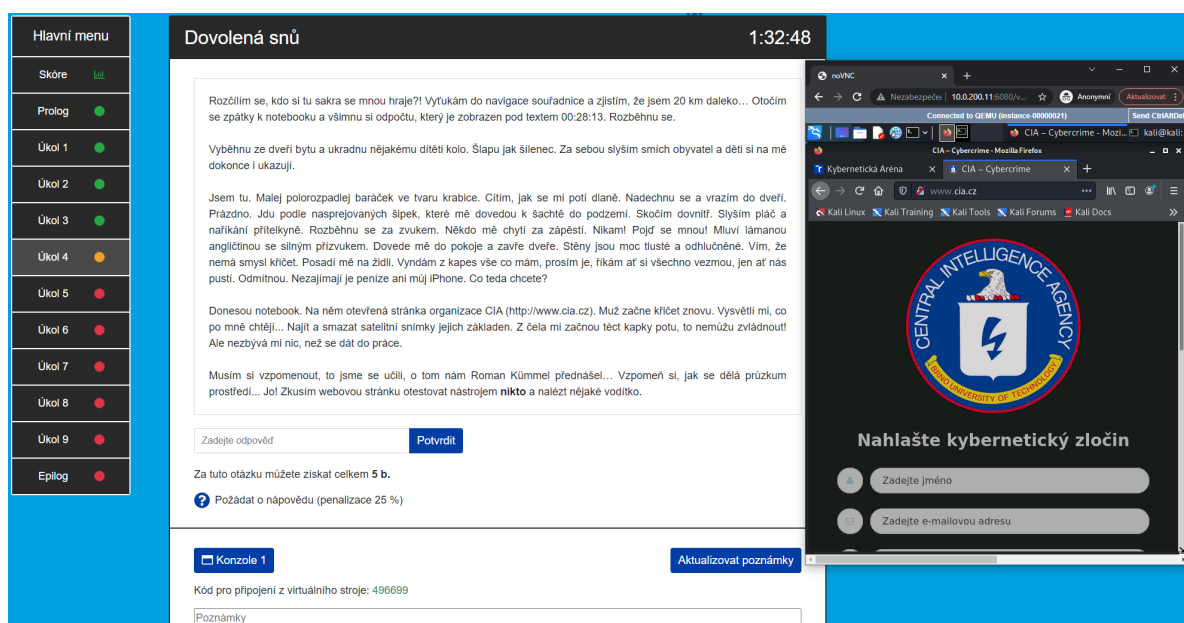


Figure 5: The example view from game Dream vacation

## 6. CONCLUSION

Cyber ranges represent an innovative solution designed for preparing security environments to effectively train users against today's cyber crimes. They can emulate thousands of scenarios for both beginners and professionals and in the academic sphere, the cyber range offers a great opportunity to try out theoretical information in practice. Within the article, the main components of private cyber range named Cyber Arena were described. It includes web-based Training application, AWX, and OpenStack. The whole structure has been developed for almost 3 years. Currently, it fully supports the playing of 3 cyber games and new 5 games are being prepared (OpenStack dependant) which will be used not only in one university course, or in the future it is planned to organize games for external students (for example from high school). For these reasons, it is essential that some games will need to run in parallel. Overall, it is necessary to optimize the usage of cloud resources (CPU, RAM, and disk capacity) so that it is possible to create the maximum amount of parallel running scenarios, given the unused full potential without affecting the running virtual machines. This optimization will be included in future works.

Compared with other cyber ranges, the biggest advantage is that the usage of Cyber Arena will be almost free, because most of cyber ranges are being paid for a lot of money or they are private. Of course, there are also many free cyber ranges, but they usually lack of support or some functionality is not supported (eg creating own games, virtualize Windows servers). In addition, Cyber Arena supports HW games based on SCADA/ICS and their implementation into OpenStack is currently being finalized. Besides cyber range functionality, Cyber Arena is also used for helping many students to work on their thesis by lending computing power or building experimental environments. In the future, there is a planned implementation of new features such as: personalized flags (students will not be able to cheat easily), move AWX out of OpenStack and extend its functionality for automated deployment, backup and restoration of the whole Cyber Arena.

## ACKNOWLEDGMENT

The presented research was financed by the Ministry of Interior under grant no. VI20192022132 and by the Brno University of Technology as an interfaculty project FEKT/FSI-J-22-7968.

## REFERENCES

- [1] YAMIN, M., M.; KATT, B.; GKIOULOS, V. Cyber ranges and security testbeds: Scenarios, functions, tools and architecture. *Computers and Security*. 2020 (88), ISSN 0167-4048. DOI: <https://doi.org/10.1016/j.cose.2019.101636>.
- [2] PRIYADARSHINI, I. *Features and Architecture of The Modern Cyber Range: A Qualitative Analysis and Survey*. Master thesis. Faculty of the University of Delaware. 2018.
- [3] KUREK, T. 2020. Private cloud price-performance analysis. *Virtual Open Infrastructure Summit*, 19<sup>th</sup>–23<sup>th</sup> October, virtual event.
- [4] *Cyber Range for Higher Education - Build or Buy?: White Paper*. CYBERBIT [online]. c2022 [cit. 7.3.2022]. Available at: <https://www.cyberbit.com/resource/cyber-range-for-higher-education-build-or-buy>.
- [5] *Welcome to OpenStack Documentation*. OpenStack docs [online]. [cit. 7.3.2022]. Available at: <https://docs.openstack.org>.
- [6] *Overview, Openstack Components, Client tools, Deployment Tools*. OpenStack [online]. [cit. 7.3.2022]. Available at: <https://www.openstack.org/software>.
- [7] BUMGARDNER, V. K. Cody. *OpenStack in action*. Shelter Island, New York: Manning Publications., 2016. ISBN 978-161-7292-163.
- [8] *Red Hat Ansible Tower*. Ansible docs [online]. c2020 [cit. 7.3.2022]. Available at: <https://docs.ansible.com/ansible-tower>.
- [9] *Ansible Documentation*. Ansible [online]. c2021 [cit. 21.12.2021]. Available at: <https://docs.ansible.com/ansible/latest>.
- [10] SILBERSCHATZ, A.; KORTH, H. F.; SUDARSHAN, S. *Database system concepts. Seventh Edition*. New York, NY: McGraw-Hill, 2020, s. 12–15. ISBN 9780078022159
- [11] *Welcome to the Heat Documentation*. OpenStack docs [online]. [cit. 14.5.2021]. Available at: <https://docs.openstack.org/heat/latest>.

# Performance Metrics Evaluations for Selected Proactive Routing Protocols in Smart Grid Neighbourhood Area Wireless Mesh Network

Enock Anderson

Brno University of Technology, Czech Republic

E-mail: [enock.anderson@vut.cz](mailto:enock.anderson@vut.cz)

**Abstract**—The smart grid communication network has improved energy optimization delivery from generation, transmission, and consumption. The introduction of wireless mesh networks in the area of Neighbourhood ad hoc Networks has contributed to the quick and cost-effective way of deployment in the areas where back-haul technology like fiber optic is expensive and challenging and may not be timely. The IEEE 802.11s standard-based wireless mesh network associated with the hybrid wireless mesh protocol is highly recommended due to its extended functionality. However, the hybrid wireless mesh protocol based mesh networks can not guarantee the quality of service in such an environment. The OLSR (Optimised Link State Routing) protocol outperforms the Reactive protocol in ad hoc Networks when transmission quality is required. This paper proposes the adaptation of Extended Transmission Control with OLSR as the optimized protocol for such an environment in neighbourhood ad hoc networks. The simulation study was carried out using the NS3 simulator to evaluate the performance of OLSR and OLSR-ETX (OLSR with Expected Transmission Count). The results confirm the viability of this protocol for neighbourhood ad hoc networks and also outperform similar simulation using NS2. The implementation of OLSR-ETX in the neighbourhood ad hoc networks rural environment will increase reliability and improve Demand Response DR in Smart Grid.

**Keywords**—Demand Response, Extended Transmission Count, NANET, neighbourhood Adhoc Network, Network simulation, NS3, OLSR-ETX, Wireless Mesh Networks

## 1. INTRODUCTION

Smart Grid (SG) technology has contributed a lot to the 21st Century power generation. It has successfully enhanced the capabilities of the traditional energy sector to ensure advanced and automated energy delivery for the economic, and industrial growth of nations [1]. The infrastructure key elements, such as smart meters, circuit breakers, transformers, feeders, substations, control centers, grid stations are all equipped with SGCN (Smart Grid Communication Network) technology architectures and software applications to provide feedback and for monitoring of the grids in real-time for utility management and DP (Demand Response). The SGCN architecture uses a multi-layer approach of dividing the communication layer into HAN (Home Area Network), NAN (Neighborhood Area Network), and WAN (Wide Area Network) [2, 3]. The NAN is a critical layer deployed to cover the distribution and the transmission domain using communication technology like wireless, wired, optical, or a combination of two or more of these technologies are deployed to ensure reliability and high throughput in the network. Its primary function is to move a large bi-directional volume of diverse data and control information between the service providers in WAN and the HAN smart devices. NAN architecture also provides a platform for the DMR (Demand Response Management) systems of the SG. However, deploying HAN in a rural and challenging environment cost-effectively and quickly is always a challenge for utility companies. Concerning the above, WMN (Wireless Mesh Network) deployment has been researched extensively in [3, 4, 5], and developed to guarantee connectivity by building a multi-hop wireless backbone to support the fast, liable, and cost-effective means to deploy NAN, especially in the remote and rural areas where access to Back-haul technology like fiber optic cable lines will be expensive and challenging to deploy within a short period. The WMN architecture uses MAC, PHY, and routing protocols to move packets among devices.

## 2. STATE OF THE ART

### 2.1. Related Works

Related research in WMN routing protocols and routing metrics deployed in NAN, MANET (Mobile Ad hoc NETWORK), WSN (Wireless Sensor Networks), and VANET (Vehicular ad hoc network) are used to update and maintain the network topology in the devices [3]. The Wireless Meshed Network uses numerous sets of protocols which are proactive, reactive, and hybrid and their goal is to optimize the network performances. They all give different results regarding throughput, latency, packet loss probability, end to end delay, network gain just to mention a few. Many research works have been done on routing protocols for WMN, MANET, VANET in [3]. In [3, 4], the goal of the authors was to analyze the performance metrics of MANET and WSN using reactive protocols such as AODV (Ad hoc On-Demand Distance Vector) and DSR (Dynamic Source Routing). The authors of [5] performed a testbed experiment of protocols HWMN (Hybrid Wireless Mesh Network). In the area of proactive protocols, authors in [4] performed an analysis of proactive OLSR and OLSR-ETX on a simulation environment but the simulation used was an old version. Other ad-hoc experiments were limited to identifying MAC (Medium Access Control) problems, by providing insights on the one-hop MAC dynamics. In [6], the authors present an experimental comparison of OLSR using the standard Hop-Count and the ETX (Extended Transmission Count) metric in an 8 by 8 grid of closely spaced Wi-Fi (Wireless Fidelity) nodes with high mobility in MANET to obtain some results. This work is more close to this research which gives us a benchmark for our results. The choice of OLSR which is a link-state proactive routing protocol that performs reliably is researched in [7] for rural and agricultural environments. This paper aims to analyze the performance metrics of proactive OLSR and OLSR-ETX in a simulation environment to confirm their suitability for deployment in a rural NAN architecture. Also to compare if it is consistent with previous research using a similar simulation tool; NS2. The simulation environment is tweaked using a few network characteristics that match dispersed, heavy interference rural environments where the NANET is deployed.

### 2.2. Simulation Description

Simulation is a good software platform and methodology to experiment with and evaluate different network topologies scenarios without or before real-world implementation. It can be used to have a fair idea of a real-life experience of how a system will perform over a longer period as well as reduce its associated cost of live implementation [4]. Most of the available simulators, such as Omnet++, Glomosim, and NS2, are realized as dual-language simulators. One language form of simulators usually C++ or Python, used for modeling provide a better learning curve and quick implementation of the experiment. The paper [8] named NS3 as a single language simulator, written in C++, which makes it a good platform choice for this experiment. Similarly, the simulator also has some preprogrammed proactive and reactive protocol classes, other MAC/PHY parameters, and wireless standards are also preconfigured to facilitate the timely results of the experiment. The addition of ETX metric algorithms was reconfigured and also implemented on methods from [3, 9, 10]. The experimental parameters are shown in Table I.

## 3. EXPERIMENTAL RESULTS

### 3.1. Performance Measurement and Analysis

The simulation was done to match the hash agricultural environment of 500–1500 square meters. The OLSR routing and the OLSR-ETX simulations were done and the results were compared in Table II. Sample codes of the ETX routing metric for AODV from GitHub were adopted and reconfigured for the OLSR protocol routing table in NS3 and it was also compared to previous works using NS2 in [11]. Each simulation was run for 120s for both protocols. Data traffic was generated from node S to node D, and simulation results (throughput, delay, and packet loss ratio) were obtained using the FlowMonitor. Also, another simulation was performed to select a route based on the ETX metric. The initial ON seconds were meant for sending ETX probe packets to the destination node, the second Off timer handled the

**Table I:** Experimental parameters

<b>Protocol</b>	OLSR/ETX
<b>Number of Nodes</b>	10
<b>Logical Link</b>	mesh
<b>MAC</b>	IEEE 802.11g
<b>Flow Type</b>	CBR
<b>Bandwidth</b>	2 Mbps
<b>Packet Size</b>	1024 bytes
<b>Number of Trials</b>	20
<b>Duration</b>	120 sec
<b>Propagation Method</b>	Friis Loss Model

probe. On receiving an ETX probe packet the forward delivery ratio and reverse delivery ratio of a link was calculated to find the value of ETX. The routes chosen for routing the packets consider the ETX value while updating the routing table with the obtained link quality. The process guaranteed the ETX values of its constituent links. In all the simulations performed, the preconfigured ETX for OLSR was able to identify the links with less hops and with better reliability. In the case of the OLSR protocol, it consistently chose the shorter path regardless of the poor link speed. Because of this, the OLSR-ETX outperformed the previous protocol. The performance of the simulation was calculated and measured based on the metrics listed below:

1. Packet Delivery ratio (PDR)

$$Deliveryratio = \frac{TotalPacketreceived}{TotalPacketsent} \cdot 100 [\%] \quad (1)$$

2. Average Throughput

$$Throughput = \frac{D}{T} [kbps] \quad (2)$$

3. Packet Delay (source node (sent at Pds) to the destination node (received at Pdr))

$$Delay = Pdr - Pds [s] \quad (3)$$

4. Expected Transmission Count (ETX), proposed by De Couto et al, is a prediction routing metric of a link.

$$ETX = \frac{1}{pf \cdot pr} \quad (4)$$

Table I the parameters for the simulation that depicts a rural NANET deployment by other APs disseminated within the campus. In general, the interference from other APs is non-controllable.

**Table II:** Simulation results

<b>Protocol</b>	<b>Data Load [Kbps]</b>	<b>Hop Count /EXT</b>	<b>Average Throughput [Kbps]</b>	<b>Packet Delivery Ratio [%]</b>	<b>End to End Delay [ms]</b>
OLSR-EXT	10	8/11	8.72	87.21	34.23
OLSR	10	8/-	7.84	78.42	36.77



OLSR-EXT	20	8/11	19.65	98.26	23.12
OLSR	20	8/--	16.94	84.70	35.54
OLSR-EXT	30	8/11	29.66	98.87	21.53
OLSR	30	8/--	26.54	88.45	33.87
OLSR-EXT	40	8/11	39.69	99.23	21.20
OLSR	40	8/--	37.65	94.13	32.67
OLSR-EXT	50	8/11	49.93	99.87	19.68
OLSR	50	8/--	47.38	94.77	29.88

### 3.2. Discussion

To assess the viability of proactive standard OLSR and OSLR ETX protocols for use in rural HANET environments, was the focus of this simulation. With that said, both routing protocol's simulation outputs were compared using tools like NetFlow, Wireshark, and trace metrics. The multi-route and hop were captured and analyzed and the ETX metrics performance was captured and analyzed according to the performance of various metrics calculations stated above. The results in Table II show that the initial bandwidth and resources were being utilized by the protocol to establish acquaintance of the network nodes and also it was using both multicast and broadcast messages which caused higher delays and low throughput. However, after the links and structure of the network had been established in various routing tables, the output was very much improved. Again it was observed that the ETX outperformed the standard OLSR in terms of End to End delay, Packet Delivery Ratio, and throughput.

## 4. CONCLUSION

This paper used the NS3 simulation tool to design a network topology of ten nodes with IEEE 802.11g devices with multipath and channel capabilities. The simulation has reconfigured the standard OLSR protocol to adapt the ETX metric in multipath/hop decisions that match the deployment of HANET. We compared the performance of the proactive OLSR routing protocol with Hop-Count and OLSR protocol implementation with the widely used ETX metric in the NS3 simulator. As expected, OLSR-ETX performed better than Hop-Count OLSR in terms of the Average Throughput and Packet Delivery Ratio and performed similarly in End to End Delays. Since the simulation has shown results consistent with related testbeds and simulations, it can be concluded that the implementation of the ETX metric in the NS3 simulator was successful and can guarantee reliable results when used in HANET for SGCN. However, in the simulation parameters, it was assumed that the energy source for the wireless nodes was reliable, but in the simulations, it was observed that the OLSR protocol draws much power from the nodes and also does not consider the energy state of the route when forwarding packets. This may lead to loss of links or reliability challenges in areas where nodes experience power losses. In this regard, there is a need to research further the energy management of proactive protocols in areas where there will be energy challenges for the wireless node devices. Future research should consider energy in the assessment of neighboring links when updating the route before forwarding packets.

## REFERENCES

- [1] W. Meng, R. Ma, and H.-H. Chen, "Smart grid neighborhood area networks," *IEEE Network*, vol. 28, no. 1, pp. 24–32, 2014. [Online]. Available: <http://ieeexplore.ieee.org/document/6724103/>
- [2] N. Raza, M. Q. Akbar, A. A. Soofi, and S. Akbar, "Study of smart grid communication network architectures and technologies," *Journal of Computer and Communications*, vol. 07, no. 03, pp. 19–29, 2019. [Online]. Available: <http://www.scirp.org/journal/doi.aspx?DOI=10.4236/jcc.2019.73003>
- [3] N. J. Jevtic and M. Z. Malnar, "The ns-3 simulator implementation of etx metric within aodv protocol," *2017 25th Telecommunication Forum (TELFOR)*, pp. 1–4, 2017. [Online]. Available: <http://ieeexplore.ieee.org/document/8249315/>

- [4] T. Wollenberg, "Performance measurement study in a wireless olsr-etx mesh network," *2012 IFIP Wireless Days*, pp. 1–7, 2012. [Online]. Available: <http://ieeexplore.ieee.org/document/6402882/>
- [5] J. Dusek and J. Mikulka, "Measurement-based domain parameter optimization in electrical impedance tomography imaging," *Sensors*, vol. 21, no. 7, 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/7/2507>
- [6] A. Abu Salem and H. Awwad, "Mobile ad-hoc network simulators, a survey and comparisons," 06 2014.
- [7] S. Lalar and A. Yadav, "Comparative study of routing protocols in manet," *Oriental journal of computer science and technology*, vol. 10, no. 1, pp. 174–179, 2017-03-28. [Online]. Available: <http://www.computerscijournal.org/vol10no1/comparative-study-of-routing-protocols-in-manet/>
- [8] S. Alam, M. F. Sohail, S. A. Ghauri, I. Qureshi, and N. Aqdas, "Cognitive radio based smart grid communication network," *Renewable and Sustainable Energy Reviews*, vol. 72, pp. 535–548, 2017. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1364032117300977>
- [9] A. Zakrzewska, L. Koszalka, I. Pozniak-Koszalka, and A. Kasprzak, "Analysis of routing protocol performance in wireless mesh networks," *2010 International Conference on Computational Science and Its Applications*, pp. 307–310, 2010. [Online]. Available: <http://ieeexplore.ieee.org/document/5476663/>
- [10] M. Malnar and N. Jevtić, "A framework for performance evaluation of vanets using ns-3 simulator," vol. 32, no. 2, pp. 255–268, 2020-03-15. [Online]. Available: <https://traffic.fpz.hr/index.php/PROMTT/article/view/3227>
- [11] M. Hiyama, E. Kulla, M. Ikeda, L. Barolli, and M. Takizawa, "Performance comparison of olsr with etx\_float and etx\_ff in a manet testbed," *2013 Seventh International Conference on Complex, Intelligent, and Software Intensive Systems*, pp. 8–14, 2013. [Online]. Available: <http://ieeexplore.ieee.org/document/6603859/>



# STUDENT EEICT



BRNO UNIVERSITY OF TECHNOLOGY  
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION  
TECHNICKÁ 3058/10, 616 00 BRNO, CZECH REPUBLIC



BRNO FACULTY OF ELECTRICAL  
UNIVERSITY ENGINEERING  
OF TECHNOLOGY AND COMMUNICATION

[www.eeict.cz](http://www.eeict.cz)



[www.fekt.vut.cz](http://www.fekt.vut.cz)

