

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

PROCEEDINGS I OF THE 29TH STUDENT EEICT 2023

ELECTRICAL ENGINEERING, INFORMATION SCIENCE,
AND COMMUNICATION TECHNOLOGIES

GENERAL PAPERS

April 25, 2023
Brno, Czech Republic

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION



Title	Proceedings I of the 29th Conference STUDENT EEICT 2023
Supervisor	Prof. Vladimír Aubrecht
Editor	Assoc. Prof. Vítězslav Novák
Publisher	Brno University of Technology, Faculty of Electrical Engineering and Communication
Year released	2023
	1st Edition

Content and language issues are the responsibility of the authors.

ISBN 978-80-214-6153-6

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION



STUDENT EEICT

SPONSORS

Diamond tier

Honeywell

Golden tier

ABB

onsemi

ThermoFisher SCIENTIFIC



STUDENT EEICT

SPONSORS

Silver tier

ASICentrum — A COMPANY OF THE SWATCH GROUP

Garrett — ADVANCING MOTION

FORVIA HELLA

NXP

TERADYNE

Valeo

Honorary tier

e·on

ŠKODA

vitesco TECHNOLOGIES

YUNEX TRAFFIC

TECHNICAL SPONSORS & MEDIA PARTNERS

IEEE

Elektrika.cz

VERONI kafestánek

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

Contents

Foreword

Bachelor Projects

B1 – Systems Modeling, Sensors, and Signal Processing I.

Tomáš Dvořák, Tomáš Sýkora
Digitization of teaching models 13
Tomáš Marcina
Comparison of parallel optimization methods 17
Mikhail Smolnikov, Tomas Horvath
Event Detection in Camera Recording 21
Štěpán Černý, Dominik Friml
Adaptive Angle Tracking Observer 24

B2 – Systems Modeling, Sensors, and Signal Processing II.

Adam Kečkéš
Analysis of BLDC motor Six-step and field oriented control 29
Tomáš Fryšták
Utilization of photovoltaic energy for heating the family house utility water 33
Libor Pavlík, Martin Paar
Design of the Energy Community Concept 37
Martin Ošťádal, Lukáš Radil
Daily Load Profile and Flexibility Aggregation 41

B3 – Biomedical Signal and Image Processing

Martin Kostial, Petr Marcoň, Peter Solár
The use of machine learning for non-invasive classification of brain pathologies 45
Michael Malík
Pathological tissue segmentation in brain MRI 48
Katerina Krejci, Jan Valosek, Jiri Chmelik
Segmentation of spinal cord structures from MRI images using deep learning 52

B4 – Communication and Information Systems

Jiří Jílek, Radek Možný, Petr Musil
The design of NB-IoT connected rain gauge 56
Alex Wittner, Marek Sikora, Sara Ricci
Dictionary Search Algorithm for Job Ads Profiling 60
David Carboch, Petr Dejdar
Real-Time Temperature Measurements in the Experimental Nuclear Reactor: A Data Analysis and Chart Creation 64
Miroslav Babeľa, Petr Musil, Radek Možný
OBD Diagnostic tool 68
Leida Ioani, Radovan Juran
Multi-Sensor Data Fusion Methods for Localization in Complex Environments 72
Pavol Repa, Milan Bubniak, Petr Musil
Principle Behavior of the Compensation Algorithm with Respect to the Input Data 75

B5 – Industrial Automation, Robotics and Artificial Intelligence

Martin Major, Tomáš Sýkora
Model of hot water exchanger station 79

Tomáš Pečinka, Petr Marcoň
Design of Asset Administration Shell testbed for Industry 4.0 83
Peter Bránecký, Petr Petyovský, Matej Pončák
Graphical user interface for embedded applications 87
Martin Sukdolák, Tomáš Benešl
Greenhouse manipulator 91
Tomáš Rotrekl, Petr Kadlec
Automatic Robot Path Planning Tool 95
Michal Marčonek, Tomáš Benešl
New home procedure on scara manipulator 99
David Mlejnek, Jakub Arm
Automated System for Creation of Gigapixel Images 103

B6 – Network Technologies, Machine Learning, and Wireless Comm.

Michal Stejskal, Petr Dzurenda
Deployment of Linear Programming for Creating Cybersecurity Profiles 108
Dominik Sabota
Network scanner for PowerShell 112
Radim Macho, Petr Musil
System for detecting connected devices to the power grid 115

Master's Projects

M1 – Microelectronics and Technology

Jan Klouda
Design of an integrated energy self-sufficient power supply for residential building consumption 119
Ngoc Nam Pham, Jan Leuchter, Jiří Háze
Acquisition Module for Non-Destructive Testing of Aircraft Composite Structures 123
Petr Zach
Design of a Hardware Accelerator of Demanding Computations with Multiple FPGAs 127
Petr Ohnút, Jakub Arm
Testbed for Simulation of MCU Application using RISC-V soft-core 131
Hana Hálová, Ladislav Chladil
Spray-dry Synthesis of Perspective Cathode Materials for Li-ion Batteries with Low Cobalt Content 135

M2 – Industrial Automation, Robotics and Artificial Intelligence

Grégoire Issassis, Petr Marcoň, Jiří Janoušek
Analysis of UAV sensors data 139
Denys Khmil, Jakub Arm
The Crane Manipulator For The 3D Printer 144
Petr Klement, Václav Kaczmarczyk
Manufacturing control algorithms for Industry 4.0 demonstrator 148
Martin Štípek, Tomáš Benešl
SCARA robot, program functions and use case 152
Van Phi Nguyen, Josef Bajer, Václav Křivánek
Precision landing of multirotor drones using computer vision 156
Martin Žiačik, Václav Kaczmarczyk
Multiagent systems for distributed production 160
Jan Wójcik
Emotion recognition system 164

M3 – Power Electronics, Electrical Machines, and Transmission

Michal Dvorak
Design of data logger for train safety management unit 168
Petr Navrátil, Radek Tománek
Instrument power supply unit 171
David Hýsek, Martin Paar
Evaluation of the Possibility of Adapting the Public Lighting Network in Brno for Electric Vehicles Charging 175

M4 – Network Security, Cryptography, Services, and Technologies

Matyáš Horký, Adrián Tomašov
Secure Transfer of Digital Messages 179
Vladyslav Shapoval, Petr Musil
Secure smart electrical socket for power metering 183
Kristián Klasovitý, Sara Ricci
Blind Issuance for Fast Keyed-Verification Anonymous Credentials 187
Norbert Lövinger
How can Privacy enhancing technologies solve car sharing privacy issues? 191
Veronika Fisarova, Tomas Stodulka
Cyber Arena Automation Component Modernization including Backup and Recovery Processes 195
Vladimir Spurny, Petr Munster
Enhancing the efficiency of fiber optic eavesdropping 199
Michal Baránek, Jan Kufa
Multifunction Speaker with Wireless Connectivity 202

M5 – Biomedical Engineering

Pavlína Sikorová, Lukáš Smital
PPG quality evaluation 206
Julie Nejezchlebová, Jana Schwarzerová, Darina Čejková
An insight into horizontal gene transfer in *Bacteroidetes spp.* using *in-silico* analysisof operon structures 210
Michal Švento, Ondřej Mokrý
Joint audio denoising and inpainting with plug-and-play proximal algorithm 214

M6 – Sensors, Measurement Systems, and Signal Processing

Martin Kubásek, Václav Kaczmarczyk
Visualization of system real-time data using software modules 218
Jakub Lukaszczyk
Influence of dataset on anomaly detector performance estimation 222
Milan Hornik, Jan Puskely, Tomas Mikulasek
Mutual Coupling Reduction Techniques in Microstrip Patch Phased Arrays 226
Jakub Nevláčil, Šimon Bilík, Karel Horák
Raspberry Pi Bee Health Monitoring Device 230
Daniel Haluška, Martin Rusz
Intelligent Wattmeter With Network Communication Option 233
Libor Matejek, Tomas Fryza
Video Summarization Using Deep Learning Approach 237
Adam Budac, Jan Kufa
Comparison of compression and decompression speeds for 360° images with resolutions up to 8K using conventional and deep learning algorithms 241

Ph.D. Projects

D1 – Biomedical Engineering

Inderjeet Bhogal, Vaishali Pankaj, Sudeep Roy
Exploration of potential inhibitors of CrtM enzyme of Staphylococcus aureus using squalene synthase inhibitors: An in-silico study 245
Michal Cicatka, Jan Lancos, Radim Burget
Approaches to clustering of microbial colonies on agar plates 250
Michal Cicatka, Jan Lancos, Radim Burget
Improvement of microbial colony segmentation using, synthetic data 255
Vaishali Pankaj, Inderjeet Bhogal, Sudeep Roy
Computational approaches to find potential phytochemical for ADHD treatment 260
Michal Nohel, Radim Kolář
Segmentation of optic disc and cup in retinal images using of deep learning approaches 265
Michal Gavenciak
Comparing Hough transform to other automated methods of feature extraction and matching for document correction 270
Mohammad Umair, Karel Sedlář
DNA Methylation Detection With Pacific Biosciences sequencing and Oxford Nanopore sequencing Technology 275
Lukáš Hrbotický
Accuracy of cycling power meters 279
Martin Hemzal
A simplified 2D model of reentry in cardiac tissue using the FHN model 284

D2 – Microelectronics and Technology

Martin Šedina, Tomáš Kazda
Short review of current EVs 289
Adam Dvorský, Martin Adámek
Experimental breeding station for edible insects 294
Pavel Tomíček, Jaroslav Boušek
Suppresion of FET gate parasitics using transient gate overvoltage 298

D3 – Power Electronics and Electrical Machines

Vitaliy Sizonenko, Petr Hutak
Design Adaptation of a 5-Phase Outer Rotor PMSM for Faulty Modes of Operation with Open Phases 302
Iveta Lolova
Topology Optimization of Line-Start Synchronous Reluctance Machine with Fully Filled Flux Barriers 307
David Rura
Performance Parameter Identification and Stability Analysis of Homopolar Electrodynamic Magnetic Bearings for Suspension System 312
Daniel Pribulla, Petr Klíma
Analysis of Permanent Magnet Segmentation Influence on Eddy-Current Losses in High-Speed PM Synchronous Machine 317
Martin Světlík
Comparing of Evolutionary and Non-parametric Algorithms and Data Preparing 322
Miroslav Mrajca
Eddy Currents Additional Losses Comparison in Oil Distribution Transformer Windings 327
Miroslav Mrajca
Advanced Turowski's Method for Eddy Loss Calculation in Transformer Foil Winding 332
Tomáš Macík
Control principles of a three-phase active rectifier and a three-phase DC/AC converter 337

D4 – Sensors, Measurement Systems, and Signal Processing

Marek Sedlařík, Petr Vyroubal
Modeling Li-ion battery pack for stationary energy storage systems 342

Serge Ayme Kouakouo Nomvussi, Jan Mikulka
Image Reconstuction in Electrical Impedance Tomography with Total Variation and Gauss-Newton Algorithm 347
Vojtěch Bednarský, Zbyněk Raida
Machine Learning for Antenna Design: Combining CST Studio Suite and Python 352
Jana Vyroubalová
Stereo Camera as a complementary sensor to LIDAR in mobile robot mapping 357

D5 – Network Technologies

Patrik Dobiáš, Lukáš Malina
Performance Evaluation of Post-Quantum Schemes on Mobile Platforms 364
Elham Younesian, Jiri Hosek
Towards 5G Networks Optimization through Machine Learning-aided Processing of Large Datasets 369
Balaji Kirubakaran, Jiri Hosek, Radim Burget
Efficient Multi-UAV Network Placement for Wireless Coverage in Challenging Environments 374
Petr Jedlička, Jan Hajný
NTT for Resources-Optimized Hardware Implementation of CRYSTALS-Kyber 379

D6 – Audio, Image and Video Signal Processing

Martin Kiac, Kamil Říha, Ondřej Krajsa
Application of YOLOv7 neural network model for control of laboratory processes 384
Stepan Jezek, Radim Burget
Visual defect detection in real-world industrial applications using convolutional neural networks 389
Šimon Skvaril
Loudspeaker Enclosure with Directional Radiation of Frequencies below 100 Hz 394
Matej Liska
FPGA Design with TDM Module for Real-Time Sound Acquisition from MEMS Microphone Array 399
Juraj Lagiň, Šimon Bilík
Structure Analysis of the FRP Rebar Using Computer Vision Techniques 404

D7 – Physical Sciences

Denis Misiurev, Vladimír Holcman
Direct Piezoelectric response miscroscopy of Bismuth ferrite thin films 410
Daniel Burda, Mohammad M. Allaham, Alexandr Knápek
Modifications of single-tip cold-field emitters by nanolayer thin dielectric coatings of refractory oxides 415
Tatiana Pisarenko, Stevan Gavranović
Design of all-inorganic lead-free cesium bismuth iodide perovskite 418

D8 – Industrial Automation, Robotics and Artificial Intelligence

Stanislav Svědiroh, Luděk Žalud
Real-Time Autonomous Vehicle Sensor Performance Assessment in Adverse Weather Conditions 423
Antonín Bohačík, David Kohout
Model of Distribution System for Energy Testbed 429
Eva Holasova, Jan Pospisil
Workload Tool for Identifying Communication Bottlenecks in Industrial Networks 434
Tomas Horelican
Towards a modular multi-platform ROS2-based navigation system for mobile robots 439
Ondrej Pospisil, Radek Mozny
Device behavior analysis based on OPC UA 445
Lukas Benesl, Radovan Juran
Multiprotocol Energy Load Management System: A Customer-Side Approach 450
Lukáš Kratochvíla
Point cloud obstacle detection with the map filtration 455

High School Projects

S1 – General Technical Focus

Michal Pomajba
Autonomously Controlled Car Track 460
Sebastian Matoušek
Bipedal Walking Robot 464
Mário Daniel Doskočil, Monika Pustková, Martin Holánik
Electric field web simulation 468
Ondřej Buček, Jan Mikulka
Design Methodology for 3D Printing of Spines 472



STUDENT EEICT

OPENING WORD OF THE DEAN

These Proceedings contain papers presented during the **29th annual STUDENT EEICT conference**, held at the Faculty of Electrical Engineering and Communication, Brno University of Technology, on April 25, 2023. 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, 167 peer-reviewed full papers distributed between 21 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 29th STUDENT EEICT (2023) 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

Digitization of teaching models

1st Tomáš Dvořák
Brno University of Technology
Faculty of Electrical Engineering and Communication
Department of Control and Instrumentation
616 00 Brno, Technická 3082/12, Czech Republic
E-mail: 220974@vut.cz

2nd Tomáš Sýkora
Brno University of Technology
Faculty of Electrical Engineering and Communication
Department of Control and Instrumentation
616 00 Brno, Technická 3082/12, Czech Republic
E-mail: 173753@vut.cz

***Abstract*—This article deals with the digital twin, one of the concepts of Industry 4.0. It is a technology that enables the digitisation of physical reality for monitoring, simulation and prediction. It is a trend that is already being used in many applications, not only in industry. This paper represents the conversion of school physical fixtures into digital form in order to simulate physical operation. It describes a guide to the creation and verification of the digital twin concept.**

***Index Terms*—Industry 4.0, Digital Twin, Siemens NX, TIA-Portal, OPC UA**

I. Introduction

Nowadays, the manufacturing company is faced with ever greater and more specific customer demands that established mass production cannot meet. This impetus behind the development of the Industry 4.0 idea is leading to the emergence of so-called smart factories, which ensure higher production efficiency. The digital twin, a virtual copy of the physical system, aims to use the current information from the physical system along with the necessary analyses to predict future behavior or problems. Based on this information, it can then take actionable interventions while keeping the operator constantly informed.

II. Industry 4.0

Industry 4.0, like previous industrial revolutions, has a profound impact on society itself. This idea was introduced by the Germans to make production more efficient. It is a concept of fully automated production without the presence of humans by interconnecting machines with each other using communication and information technology. Industry 4.0 will reduce downtime, save costs, ensure faster decision-making, increase demand, and overall move the manufacturing enterprise forward. At the heart of Industry 4.0 are so-called cyber-physical systems. [1]

A. Cyber-physical systems

Industry 4.0 requires a transition from physical components to cyber-physical systems. Cyber-physical systems consist of electronic, software, and physical components. Each physical component, i.e. the robot, the machine, and the product itself, must have a certain intelligence. The goal is for the product itself to know its manufacturing process and its specifications, to be able to report information about itself in real-time and to be able to control its production in communication with other smart devices in the manufacturing process. The principle of this decentralized control will then enable real-time decision-making, flexible response to change, feedback, and autonomous operation, thereby increasing production efficiency and productivity. [2]

B. Key pillars

The introduction of cyber-physical systems generates a huge amount of data in the production plant, which is then further processed. The analysis of this big data is used to detect threats or for predictive maintenance. At the same time, this data must be stored somewhere so that it can be accessed from anywhere and everywhere, which is made possible by online cloud storage.

Cybersecurity is related to this data to prevent its loss, which could be critical to the production process.

Another pillar is the Internet of Things, which is used to communicate between machines, products and robots. System integration, or horizontal and vertical integration, allows all elements of a manufacturing enterprise to be connected for immediate response.

Simulation translates physical equipment and processes into a virtual environment to simulate production so that there is no unnecessary downtime or loss of quality in real-world use.

Augmented reality displays a digital image in the real plant for repair or maintenance purposes. The development of autonomous robots has made great strides and they are currently filling the gap between conventional robots and humans, and one day there will be a time when they will be incorporated into manufacturing where they will safely interact with humans.

The final pillar is 3D printing, which produces a part from a 3D model using additive manufacturing, eliminating material waste and allowing more complex designs to be realised. [3]

C. Summary

Industry 4.0 brings with it many benefits such as increased productivity, faster response times, reduced waste, or predictive maintenance. However, all these benefits come at a cost. A company must make an initial investment to obtain them, as part or all of the business must be digitized according to the principles of Industry 4.0. In connection with this, it will then have to be more concerned with data and employee

security. In addition, some jobs will disappear and new jobs will be created, so it is necessary to deal with the retraining of employees. [4]

III. Digital twin

As already mentioned, a digital twin is a virtual representation of a physical system using data exchange for monitoring, prediction, optimization, control and decision making in the shortest possible time. In other words, it is a digital copy of the equipment that will facilitate the commissioning of the equipment by reducing the time to commission it and detecting any errors. In more complex cases, the environment in which the equipment is located must also be taken into account, as this can also affect its operation. Once the equipment is running, the digital twin is used to perform various simulations to help us predict future errors, including maintenance planning. All this is done on the basis of constantly updated data, i.e. there must be a constant exchange of data between the physical device and the digital twin. [5]

A. Distribution

The problem with Industry 4.0 is that companies are exploiting it for commercial purposes without using its concepts. It should be understood that the definition of a digital twin described above is a complex technology and it is always necessary to define whether it is a digital twin in the true sense of the word. In general, it depends on whether there is any data exchange between the physical fixture and the digital model.

If there is no interaction between the physical fixture and the digital model, it is only a digital model to simulate the operation of the device without any intervention. In the case of a digital shadow, data flows from the physical fixture to the digital model when a change in the physical state is reflected in the digital model. In contrast, the digital twin allows data exchange in both directions. [6]

B. Utilization

The use of the digital twin is not only in industry but also in other areas of our lives. In healthcare, the digital twin is used to predict future health, and select the right therapy in the cancer treatment process. In meteorology, it is used to forecast the weather, and in the educational process, we are talking about finding the talents and gifts of students in particular fields. The most well-known use is the digital twin of an entire city, specifically Singapore, where the city is being redeveloped to improve the well-being of the population. [7]

IV. Physical products

A. Case sorter

The physical model of the Case Sorter is mounted on a rectangular plate. The model contains 3 conveyor belts which are tensioned on toothed pulleys and fixed with aluminum profiles. The belts are driven by permanent magnet DC motors. Also used are 2 solenoids which are mounted on a wooden base. To move the crates, a rectangular plexiglass is screwed to the end of the solenoid piston. In front of each solenoid, optical sensor holders are attached to aluminum profiles. These sensors have plates of umacart placed opposite each other for better light reflection. On the top of the fixture, there is a connector for the PLC connection. On the underside of the fixture are the control boards that control the motors. The short side belts are controlled in both directions, whereas the longest belt is only controlled in one direction.

On the main belt, crates are placed on the right side, which is sorted into pallets on the side belts based on size or scanning time. Once the pallets are full, they go to the other end of the belt for unloading and then go back for more crates.

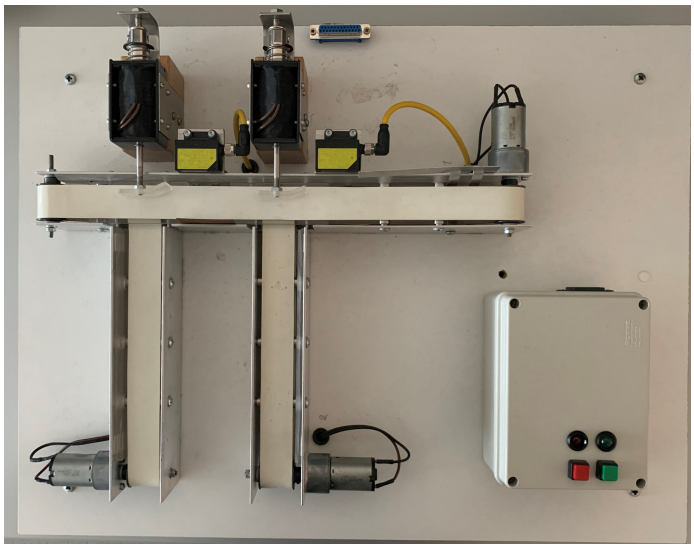


Fig. 1. Case sorter

B. Drilling machine

The physical model of the Drill includes 2 tall aluminum stands attached to a rectangular plate. On the right stand, there are 2 pulleys on which the belt is tensioned. Rails are attached to the front to move the guide rollers of the support plate. A motor holder with a chuck is bolted to the carrier plate. The belt is connected to the carrier plate by 2 small aluminum profiles, allowing the carrier plate to be moved up or down. It also has a permanent magnet on the inside. On the left stand, there are 5 reed contacts which are used to sense the position based on the permanent magnet's approach. Directly below the chuck, there is a defined space for positioning the workpiece. To identify the size of the workpiece, there are also 2 microswitches located exactly above each other under the chuck. The movement of the belt and the drilling itself is controlled by permanent magnet DC motors.

The presence of the workpiece is decided by the micro switches, which also give us information on whether the workpiece is large or small. In the case of a large workpiece, the drill drills twice, first from the top position to the middle position, then back to the top position, and drills to the bottom position. In the case of a small workpiece, it drills only

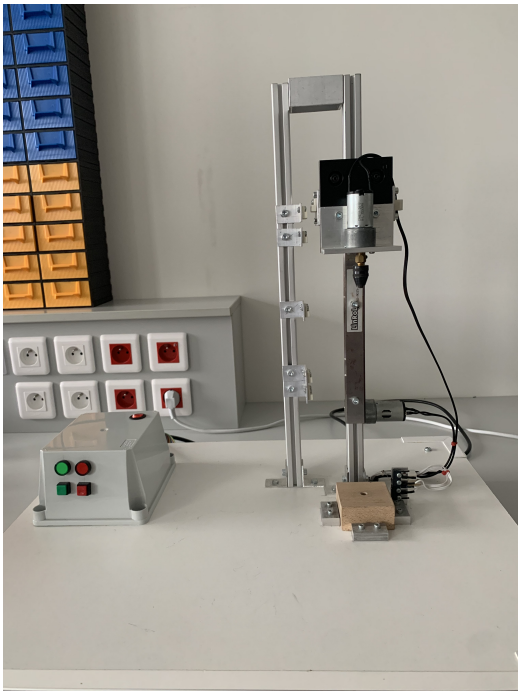


Fig. 2. Drilling machine

once, i.e. from the top position to the bottom position. These positions are identified by the reed contacts.

V. Creating a digital twin

NX Mechatronics Concept Designer (MCD) from Siemens was used to convert the physical fixtures into virtual form. The digital twin is supposed to be an exact copy of the physical jig, so it was first necessary to measure all the aluminium, plastic or metal components. For the electronic components, it was then enough to find their datasheet. MCD is a user-friendly 3D modelling environment where the entire assembly was created after the individual components were drawn. This assembly should mimic the physical jig in appearance. After this assembly, it was necessary to define the basic physics, i.e. the non-deformable and colliding solids. After defining the possible movement of the bodies and collision materials, the components were set in motion. The functionality of the whole model was tested using simple operations that can be created in MCD. The last step was to create signals for the PLC with the appropriate simulation parameters. For the case sorter model, figures 5 and 6 show a detail of the model creation. The figures show a list of non-deformable and colliding bodies along with constraints to ensure proper movement of the bodies.

The physical fixtures are controlled by a program from the PLC and so is the digital twin. For this, TIAPortal was used, where variables with the same names as the signals in the MCD were created. The physical PLCs already have built-in communication via an OPC UA server. In our case it was enough to enable this communication in TIAPortal and in the virtual PLC it was necessary to enter the same IP address as in TIAPortal to create it. Subsequently, the created

Basic Physics	
CB_levy_profil	Collision Body
CB_levy_profil_2	Collision Body
CB_podstavec	Collision Body
CB_pravy_profil	Collision Body
CB_pravy_profil_2	Collision Body
ObjectSource(1)	Object Source
RB_hlavni_pas_1	Rigid Body
CB_hlavni_pas_1	Collision Body
RB_pist_1	Rigid Body
CB_pist_1	Collision Body
RB_pist_2	Rigid Body
CB_pist_2	Collision Body
RB_plast	Rigid Body
CB_plast	Collision Body
RB_plast_2	Rigid Body
CB_plast_2	Collision Body
RB_puk_1	Rigid Body
CB_puk_1	Collision Body
RB_vedlejsi_pas	Rigid Body
CB_vedlejsi_pas	Collision Body
RB_vedlejsi_pas_2	Rigid Body
CB_vedlejsi_pas_2	Collision Body

Fig. 3. Basic physics

Joints and Constraints	
PreventCollision(1)	Prevent Collision
PreventCollision(3)	Prevent Collision
PreventCollision_2_levy_profil_hla...	Prevent Collision
PreventCollision_2_levy_profil_ve...	Prevent Collision
PreventCollision_2_pravy_profil_h...	Prevent Collision
PreventCollision_2_pravy_profil_v...	Prevent Collision
PreventCollision_levy_profil_hlavn...	Prevent Collision
PreventCollision_levy_profil_vedle...	Prevent Collision
PreventCollision_pasy	Prevent Collision
PreventCollision_pasy_2	Prevent Collision
PreventCollision_pravy_profil_hla...	Prevent Collision
PreventCollision_pravy_profil_ved...	Prevent Collision
RB_hlavni_pas_FJ(	Fixed Joint
RB_pist_1_SJ(1)	Sliding Joint
RB_pist_2_SJ(1)	Sliding Joint
RB_plast_2_SJ(1)	Sliding Joint
RB_plast_SJ(1)	Sliding Joint
RB_vedlejsi_pas_2_FJ	Fixed Joint
RB_vedlejsi_pas_FJ	Fixed Joint

Fig. 4. Joints and constraints

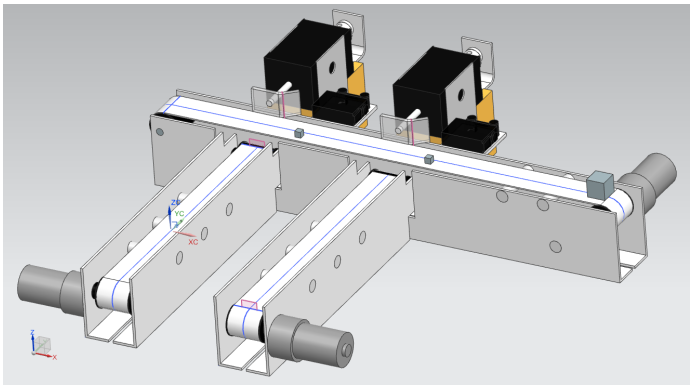


Fig. 5. Digital model of case sorter

program was uploaded to the virtual PLC and the verification of communication via OPC UA server was performed via the UA Expert software representing the test client. The last step was to connect the MCD to the created OPC UA server and to program the signals from the simulation to the external signals from the PLC. Then it was enough to run the simulation and the digital twin was controlled based on the program from the PLC via the OPC UA server.

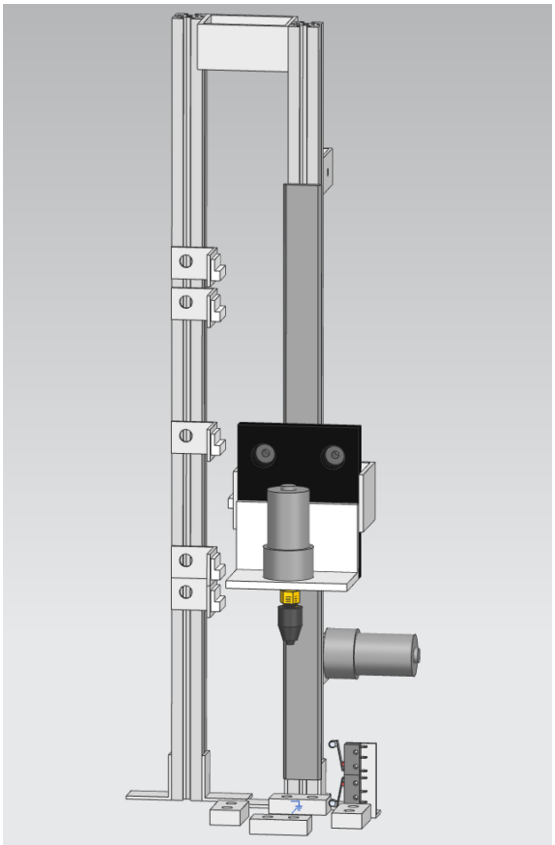


Fig. 6. Digital model of drilling machine

VI. Conclusion

The result of this work is digital twins of physical fixtures that allow simulation of physical operation. The work can be used for both in-school and out-of-school purposes, as it presents step-by-step instructions on how to create a digital twin and how to verify the functionality of the program before deploying it on a physical fixture to avoid destroying it. However, in terms of the definition of a digital twin, these are only digital models. There is no data exchange between the physical fixture and the virtual model, so this article is only about validating the idea of Industry 4.0 and the concept of a digital twin for its use in industry.

Acknowledgment

The completion of this paper was made possible by the grant No. FEKT-S-23-8451 - ”Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence, automation and measurement” financially supported by the Internal science fund of Brno University of Technology.

References

[1] Hermann Meissner, Rebecca Ilsen, and Jan C. Aurich. Analysis of control architectures in the context of industry 4.0. Procedia CIRP, 62:165–169, 2017. URL: https://linkinghub.elsevier.com/retrieve/pii/S2212827117300641, doi:10.1016/j.procir.2016.06.113.

[2] Kyberfyzické systémy. URL: https://www.iot-portal.cz/2016/08/22/kyberfyzikalni-systemy/.

[3] Gizem Erboz. How to define industry 4.0: The main pillars of industry 4.0. Managerial trends in the development of enterprises in globalization era, page s1–7, 2017.
URL: https://www.researchgate.net/publication/326557388_How_To_Define_Industry_40_Main_Pillars_Of_Industry_40.

[4] Mohamed Mamad. Challenges and benefits of industry 4.0: An overvie. IJSOM: 256 International Journal of Supply and Operations Management, 315(3):256–265, 2018. URL: http://www.ijsom.com/article_2767.html, doi: 10.22034/2018.3.7.

[5] Fei Tao, Qinglin Qi, Lihui Wang, and A.Y.C. Nee. Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0. Engineering, 5(4):653–661, 2019. URL: https://linkinghub.elsevier.com/retrieve/pii/S209580991830612X, doi:10.1016/j.eng.2019.01.014.

[6] AHELEROFF, Shohin, Xun XU, Ray Y. ZHONG a Yuqian LU. Digital Twin as a Service (DTaaS) in Industry 4.0: An Architecture Reference Model. Advanced Engineering Informatics [online]. 2021, 47 [cit. 2023-03-09]. ISSN 14740346. Dostupné z: doi:10.1016/j.aei.2020.101225

[7] Vše o průmyslu, 2016. URL: https://www.vseoprumyslu.cz/digitalizace/digitalni-prototypovani/jak-digitalni-dvojcata-a-technologie-iiot-prispivaji-provozu.html.

Comparison of parallel optimization methods

1st Tomáš Marcina
Bachelor Degree Programme (3)
FEEC BUT
Brno, Czech
230124@vut.cz

***Abstract*—The work is concerned on the comparison of different algorithms designed for the optimization of multidimensional functions. Specifically, it deals with population optimization methods, which are generally suitable for parallelizing the calculation of the optimal function value. The parallelization of the calculation leads to a significant acceleration speed of the calculations. This paper compares several different algorithms suitable for this.**

***Index Terms*—optimization, parallel optimizations, differential evolution, particle swarm optimization, firefly algorithm, artificial bee colony algorithm, grey wolf optimization.**

I. Introduction

Optimization is the process of finding the optimal value of a function, extremes (minimum or maximum values). Optimization is the basis for the use of mathematical models and statistics in decision-making processes. The effort to speed up this process leads to the parallelization of calculations, most suitable for this are population-based optimization methods.

The paper is organized as follows. The second chapter presents introduction to optimization. The third and fourth chapter describes implementation of algorithms. The last chapter compares all algorithms.

II. Optimization

A. Formulation of the basic optimization task

The goal of each optimization task is to find the optimal solution of given objective function. Finding minimum value of function can be written as

$$\min_x f(\mathbf{x}) \tag{1}$$

$$\text{subject to } x \in \mathcal{X},$$

where

- $\mathbf{x}$ is vector $[x_1, x_2, ...x_n]$, which defines a point in n-dimensional space
- $f(\mathbf{x})$ is cost function
- $\mathcal{X}$ is search domain of the objective function.

Any value $\mathbf{x}$ in search domain $\mathcal{X}$, that minimizes the objective function is called solution, is written in equation (2) [5].

$$\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) = \{\mathbf{x}^* \in \mathcal{X} : f(\mathbf{x}^*) \leq f(\mathbf{x}), for\, \forall\, \mathbf{x} \in \mathcal{X}\} \tag{2}$$

Figure 1 shows example of optimization in single dimensional space with solution x^*, which is local minimum.

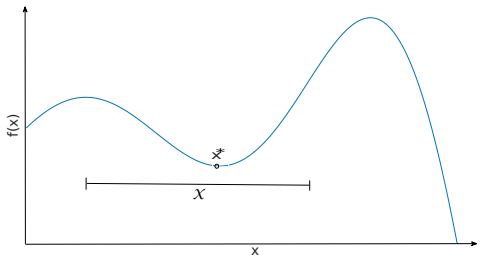


Fig. 1. Example of optimization in single dimensional space

For finding maximum value of objective function equation can be redefined as

$$\max_x f(\mathbf{x}), \tag{3}$$

and solution will be

$$\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) = \{\mathbf{x}^* \in \mathcal{X} : f(\mathbf{x}^*) \geq f(\mathbf{x}), for\, \forall\, \mathbf{x} \in \mathcal{X}\} \tag{4}$$

B. Constraints

For most optimization tasks is search domain $\mathcal{X}$ narrowed with constraint functions

$$g_i(\mathbf{x}) \geq 0, i = 1, 2, ...n, \tag{5}$$

which defines search domain for optimizations. Constraints are defined with operators $\geq, \leq, =, \neq$, possibly $>, <$. [1]

III. Population-based optimization methods (Algorithm 1)

The basic idea of population methods is to distribute a set of design points (agents) with the number of N throughout the searched domain. Agents subsequently influence each other according to a predefined principle, this principle is directly defined in a specific algorithm, and iteratively move towards the optimal value of the objective function. [1] A higher number of agents helps the algorithm to avoid getting stuck in a local minimum of the objective function. Most of the population methods have a stochastic nature of computation and in general it is easy to parallelize the computation [1].

Algorithm 1 Population-based algorithm pseudocode

```
i = 0
initialize population (positions of agents)
evaluate population
while i < i_max do
    agents influence each other
    evaluate temporary population
    save better solutions
    i += 1
end while
return best solution
```

IV. Implemented algorithms

Library containing author's implementation of methods described in following subsections is available. Algorithms were implemented in Matlab. Library includes all methods, test functions and scripts on which algorithms were tested.

A. Differential evolution (DE)

Differential evolution attempts to optimize each agent in a population by recombining other random agents.

For each agent $\mathbf{x}^{(m)}$ algorithm follows next principle:

1) Choose four random solutions $\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_3}, \mathbf{r_4}$, best solution so far $\mathbf{x_{best}}$ and combine them according to formula (6) [5],
$$\mathbf{u}^{(m)} = \mathbf{x_{best}} + w(\mathbf{r_1} + \mathbf{r_2} - \mathbf{r_3} - \mathbf{r_4}), \tag{6}$$
where w is input parameter in interval $< 0,4; 1 >$,
2) Choose random dimension $j \in [1, 2, ...n]$ in n-dimensional space,
3) Construct temporary vector $\mathbf{y}^{(m)}$, which every dimension follows
$$y_i^{(m)} = u_i^{(m)} \text{ if i=j, or with probability } p, \quad \text{otherwise } y_i^{(m)} = x_i^{(m)} \tag{7}$$
p is another input parameter in interval $< 0; 1 >$.

B. Particle swarm optimization (PSO)

In every iteration, agents move towards the best position it has ever been in and towards the best position found so far with velocity $\mathbf{v}$. Agent's new position is calculated according to (8) [4],
$$\mathbf{x}(k+1)^{(m)} = \mathbf{x}(k)^{(m)} + \mathbf{v}(k)^{(m)}, \tag{8}$$
values in vector $\mathbf{v}^{(m)}$ follows equation (9),
$$\mathbf{v}(k+1)^{(m)} = w\mathbf{v}(k)^{(m)} + c_1 r_1(\mathbf{x}(k)_{best}^{(m)} - \mathbf{x}(k)^{(m)}) + c_2 r_2(\mathbf{x}(k)_{best} - \mathbf{x}(k)^{(m)}), \tag{9}$$
where

- c_1, c_2 are input parameters with value $c_1, c_2 \in \mathbf{R}$,
- w is input parameter with value in interval $< 0; 1 >$,
- r_1, r_2 are random values in interval $< 0; 1 >$,
- $\mathbf{x}_{best}^{(m)}$ is agent's best solution so far,
- $\mathbf{x_{best}}$ is best solution in entire population.

C. Firefly algorithm (FF)

The algorithm is inspired by the behavior of these animals during the mating season, when they try to attract individuals of the opposite sex by flashing lights. Every agent watches everyone else and iteratively move to those with best light intensity according to equation (10) [1],
$$\mathbf{x}(k+1)^{(m)} = \mathbf{x}(k)^{(m)} + \beta I(||\mathbf{y}(k)^{(m)} - \mathbf{x}(k)^{(m)}||) \cdot (\mathbf{y}(k)^{(m)} - \mathbf{x}(k)^{(m)}) + \alpha\epsilon, \tag{10}$$
where

- α, β are input parameters in interval $< 0; 1 >$,
- ϵ is random move, calculated with normal distribution around zero with sigma 1.

Light intensity is function of distance of two agents, with input parameter γ, wich is positive number close to zero
$$I(r) = e^{-\gamma r^2}. \tag{11}$$

D. Artificial bee colony algorithm (ABC)

Algorithm consists of three groups of bees

- employed bees apply a random neighborhood structure to each agent and then calculate the solution fitness value,
- onlookers select an agent based on the solution fitness and apply the random neighborhood structure again,
- scouts, who replace a solution that can't be improved with a new random solution

Random neighborhood structure for agent $\mathbf{x}^{(m)}$ consists of three steps [2]:

1) Choose random agent $\mathbf{x}^{(n)}$, where $m \neq n$,
2) and for every dimension $i \in [1, 2, ...n]$ calculates
$$x(k+1)_i^{(m)} = x(k)_i^{(m)} + \Phi_i(x(k)_i^{(m)} - x(k)_i^{(n)}), \tag{12}$$
Φ is random number in interval $< -c, c >$, where $c \in \mathbf{R}$
3) better solution is saved, if solution wasn't updated for a while scouts will replace it.

Scouts will replace solution if it wasn't been for d-times. Abandonment decision d is input parameter, $d \in \mathbf{Z}$. Formula (13) defines principle of replacing solution $\mathbf{x}^{(m)}$.
$$x_i^{(m)} = a_i + r(b_i - a_i), \tag{13}$$
where

- a_i is lower bound of searched space $\mathcal{X}$,
- b_i is upper bound of searched space $\mathcal{X}$,
- $i \in [1, 2, ...n]$ in n-dimensional space,
- r is random number in interval $< 0, 1 >$

E. Improved gray wolf optimization (GWO)

Three best solutions are selected, named α, β and γ. For every agent $\mathbf{x}^{(m)}$ are calculated $\mathbf{X_P}$, which follows equation (14) [3]

$$\mathbf{X_P} = \mathbf{P} - A \times \mathbf{D_P}, \tag{14}$$

$$A = 2ar_1 - a, \tag{15}$$

where a goes down with every iteration ($2 \rightarrow 0$)

$$\mathbf{D_P} = \mid 2r_2\mathbf{P} - \mathbf{X} \mid, \tag{16}$$

where

- $\mathbf{P}$ are solutions α, β, γ
- $r_{1,2}$ are random numbers in interval $< 0; 1 >$
- $\mathbf{x}^{(m)}$ is position of actual agent.

New position is average of $\mathbf{X_P}$

$$\mathbf{x}^{(m)} = \frac{\mathbf{X}_\alpha + \mathbf{X}_\beta + \mathbf{X}_\gamma}{3}, \tag{17}$$

Algorithm can be modified to reduce the probability of getting stuck in a local minimum. The improved version consists in that new random agent $\mathbf{X}^{(n)}$ is calculated in radius R between the past and calculated position of agent. New value is calculated for each dimension according to the formula [3]

$$X(k+1)_i^{(m)} = X(k)_i^{(m)} + r(X(k)_i^{(n)} - X(k)_i^{(o)}), \tag{18}$$

where

- $\mathbf{X}^{(o)}$ is random agent in population,
- r is random number in population $< 0; 1 >$
- i is index of calculated dimension, $i \in [1, 2, ...n]$ in n-dimensional space.

V. Comparison of algorithms

All tests have been performed in $\mathbf{R}^2$ space with the same initial population for all algorithms. Initial population was generated randomly. In tables I and II minimum value of all test functions is 0.

A. Convergence comparison

The graphs in figures 2, 3 show average value of the found minima from 20 runs with 50 agents and 50 iterations. The firefly algorithm and particle swarm optimization tend to get stuck in a local minimum. In the case of Easom's function, algorithms remained stuck in global maxima, because there was no local minimum. Most of algorithms, appears that their convergence depends on the tested function, except for firefly algorithm, which has the fastest convergence.

B. Accuracy based on number of iterations

Table I shows the average value of the found minima from 10 runs of the algorithms with 50 agents. All algorithms already achieved a result quite close to the optimal solution with respect to the maximum functional value at the number of iterations 10, with the exception to Layeb10 function, which is difficult to optimize. With the number of iterations 50, the only exceptions were particle swarm optimization and the firefly algorithm. After 500 iterations only Particle Swarm Optimization was not suitable, which is probably more prone to getting stuck in a local minimum.

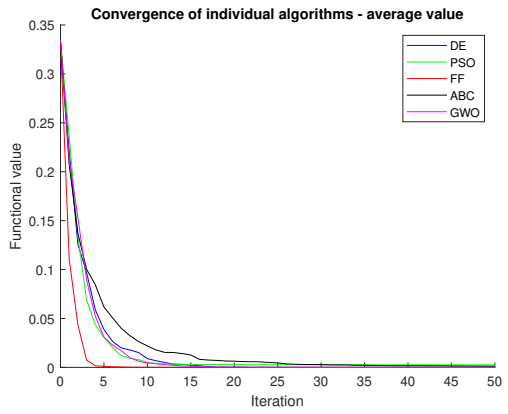


Fig. 2. Average value of algorithms on Schaffer function

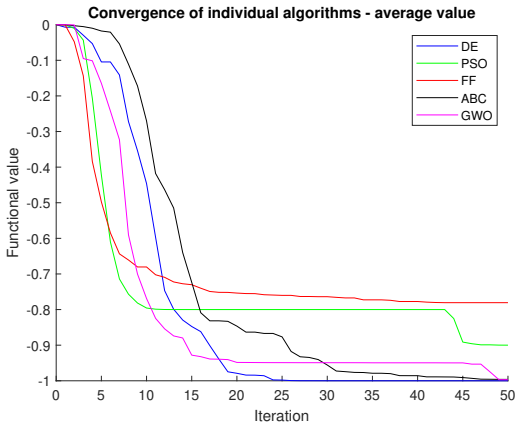


Fig. 3. Average value of algorithms on Easoms function

TABLE I
Accuracy based on number of iterations

Alg	Test function	Number of iteration			
		10	50	100	500
DE	Beale	2,35E-05	7,62E-02	2,29E-01	0,00E+00
	Bukin	7,32E-01	2,47E-02	2,96E-02	2,78E-02
	Layeb10	6,73E+00	3,78E-05	2,31E-08	0,00E+00
	Schaffer N.2	8,22E-03	0,00E+00	0,00E+00	0,00E+00
PSO	Beale	8,39E-02	2,37E-01	2,42E-01	7,62E-02
	Bukin	3,66E-01	3,01E-02	2,43E-02	2,69E-02
	Layeb10	1,73E+01	5,10E+00	7,25E+00	5,16E+00
	Schaffer N.2	4,56E-03	2,08E-10	0,00E+00	0,00E+00
FF	Beale	5,66E-05	2,08E-05	7,64E-02	2,33E-06
	Bukin	6,38E-01	2,98E-01	2,57E-01	8,26E-02
	Layeb10	9,06E+00	2,59E+00	3,06E-01	4,74E-03
	Schaffer N.2	1,38E-07	1,57E-08	5,19E-09	1,28E-09
ABC	Beale	6,18E-03	7,36E-10	4,71E-18	0,00E+00
	Bukin	7,32E-01	2,47E-02	2,96E-02	2,78E-02
	Layeb10	4,97E+00	1,25E-01	6,58E-03	0,00E+00
	Schaffer N.2	7,14E-03	1,37E-08	1,59E-12	0,00E+00
GWO	Beale	1,09E-05	1,71E-07	1,10E-08	3,98E-10
	Bukin	1,83E+00	5,69E-01	5,52E-01	2,14E-01
	Layeb10	3,27E+00	8,61E-01	5,74E-01	3,89E-01
	Schaffer N.2	4,56E-03	2,08E-10	0,00E+00	0,00E+00

C. Accuracy based on number of agents

The table II shows average value of the found minima from 10 runs at 50 iterations. With the number of agents 10, algorithms achieved worse results than with 10 iterations with the number of agents 50. At higher numbers of agents (100, 500) are results more robust.

TABLE II
ACCURACY BASED ON NUMBER OF AGENTS

Algorithm	Test function	Number of agents			
		10	50	100	500
DE	Beale	1,10E+00	7,62E-02	7,62E-02	5,62E-22
	Bukin	1,97E-02	1,92E-02	1,75E-02	2,30E-02
	Layeb10	7,03E+00	6,20E-05	9,45E-06	2,00E-06
	Schaffer N.2	1,33E+01	2,77E-05	5,71E-06	2,60E-06
PSO	Beale	1,44E+00	1,63E-01	1,63E-27	1,11E-32
	Bukin	2,23E-02	1,62E-02	1,64E-02	2,33E-02
	Layeb10	2,84E+01	4,20E+00	2,87E-01	2,87E-01
	Schaffer N.2	1,73E-01	1,04E-03	0,00E+00	0,00E+00
FF	Beale	2,34E-01	2,88E-05	3,81E-06	1,45E-06
	Bukin	1,72E+00	2,91E-01	1,68E-01	8,96E-02
	Layeb10	3,74E+01	3,07E-01	8,18E-03	3,25E-03
	Schaffer N.2	2,51E-01	1,49E-08	5,02E-09	6,92E-10
ABC	Beale	7,62E-02	5,00E-11	2,17E-11	3,37E-12
	Bukin	4,86E-02	5,34E-02	4,78E-02	4,34E-02
	Layeb10	4,95E+00	8,30E-01	7,51E-02	2,04E-02
	Schaffer N.2	2,11E-03	1,79E-08	1,58E-08	2,43E-09
GWO	Beale	7,64E-02	1,20E-07	2,79E-08	5,60E-09
	Bukin	1,60E+00	6,78E-01	4,94E-01	3,39E-01
	Layeb10	2,31E+00	9,75E-01	5,73E-01	2,87E-01
	Schaffer N.2	3,92E-04	1,72E-07	0,00E+00	0,00E+00

VI. CONCLUSION

Regarding the speed of convergence, the firefly algorithm is the best, but it simply gets stuck in a local minimum. The particle swarm optimization is slightly better in avoiding getting stuck in a local minimum, and converges as well as the remaining algorithms. According to robustness, differential evolution, artificial bee colony algorithm and gray wolf optimization has the best results.

ACKNOWLEDGMENT

I would like to thank the supervisor Mr. Kozubík for his professional guidance, consultations, patience and his unique incentives to motivate work.

REFERENCES

[1] Kochenderfer, M.;Wheeler T. *Algorithms for optimization*. The MIT Press Cambridge, 2019, ISBN 9780262039420.

[2] Ye, T.; Wang, W.; Wang, H.; Cui, Z.; Wang, Y.; Zhao, J.; Hu, M. Artificial bee colony algorithm with efficient search strategy based on random neighborhood structure. *Knowledge-Based Systems*, vol. 241, 2022, ISSN 0950-7051.

[3] Nadimi-Shahraki, M. H.; Taghian, S.; Mirjalili, S. An improved grey wolf optimizer for solving engineering problems. *Expert Systems with Applications*, vol. 166, 2021, ISSN 0957-4174.

[4] Gad, A.G. Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review. Arch Computat Methods Eng 29, 2531–2561 (2022). https://doi.org/10.1007/s11831-021-09694-4

[5] Tvrdlík, J. *Evoluční algoritmy*. Ostravská univerzita, Přírodovědecká fakulta, 2004: pp. 4-48.

Event Detection in Camera Recording

1st Mikhail Smolnikov
Bachelor Degree Programme (3.)
Brno University of Technology
Brno, Czech Republic
xsmoln01@vut.cz

2nd Tomas Horvath
Brno University of Technology,
FEEC, Dept. of Telecommunications,
Technicka 12, 616 00 Brno, Czech Republic
horvath@vut.cz

***Abstract*—This paper describes the internal architecture of a desktop application that has been designed to detect and to recognise motion in a video sequence. The system automates the monitoring of videos from security cameras, and processes videos up to 12 times faster than a human operator could do manually. The user is then presented with a text report, video clips, and images of any events that occurred in the uploaded video, allowing efficient and timely detection of any potential security issues.**

***Index Terms*—motion detection, object classification, image processing, neural networks, video analysis**

I. Introduction

Nowadays, there is an increasing need to process and analyse large amounts of media data. To address this need, neural networks are being trained and utilised for various tasks, such as recognising faces or objects in an image. Many of these tools are available as open source, making them accessible for use in various projects.

This article provides detailed information about a possible solution for detecting and recognising motion in a video sequence of a static scene. The main motivation behind this project was the need for automatic monitoring of the situation in the backyard where a security camera had been installed. The system had the ability to recognise any objects that appeared in the video feed, which could be useful in identifying potential intruders or unwanted visitors. Overall, this project demonstrated the potential of neural networks in solving real-world problems.

The rest of this paper is structured as follows: Section II provides detailed information about the technologies used to create the graphical application interface and manipulate frames in the project. In Section III, the individual steps involved in processing video files are described. These steps include preprocessing, motion and object detection, and representing the results. Section IV outlines the results of system performance measurements carried out in various operating scenarios. These measurements were used to evaluate the processing speed and efficiency of the application.

II. Technology stack

For the graphical interface of the application, PyQt5 (Python Qt version 5) was chosen as the library. PyQt is a popular Python binding for the cross-platform Qt framework [1] [2]. It was chosen thanks to its flexibility in building interfaces and controlling in-app events such as clicking, scrolling, and more. With the library, developers have access to a wide range of widgets and can easily customize the look and feeling of their application. Additionally, PyQt is a stable and reliable library that has been widely used in the Python community for many years [3].

To ensure accurate and efficient image processing control, one of the tools utilized is the opencv-python package, which conveniently wraps an OpenCV (Open Source Computer Vision) library originally written in C++ [4] [5]. By using this library, it is possible to implement various image processing techniques such as the Frames Differences algorithm. Moreover the library controls the pre-trained YOLOv5 (You Look Only Once version 5) convolutional neural network model used for object classification [6]. This model is particularly effective at accurately classifying objects and providing their position coordinates with a high degree of accuracy [7] [8].

The ffmpeg command-line tool is a program used to manipulate and process video and audio files [9]. In this context, it can be used to cut a video into smaller segments based on specific time intervals.

III. Video processing

A video is composed of frames, which are still images made up of pixels. Each pixel has a mathematical value that describes its color, which can be represented as an RGB (Red, Green, Blue) value. These values are used to manipulate the image for analysis or modification purposes using various image processing techniques.

The primary objective of the project was to achieve faster processing speeds compared to manual video analysis by human operators. To achieve this, the method was chosen where the images are processed only once and then the result is compiled from retrieved data.

To start processing the video frames, they need to be prepared for further analysis. This can be done by adjusting the frame resolution, converting colors, and applying some filters to eliminate any potential noise. Next to detect motion in a video, the frames in which motion occurred must be identified. Once this has been accomplished, it is time to identify the object that caused the movements. Finally, all frames can be grouped into clips of events, and a report for the entire video can be generated.

The process of analysing recorded video is summarised in Fig. 1.

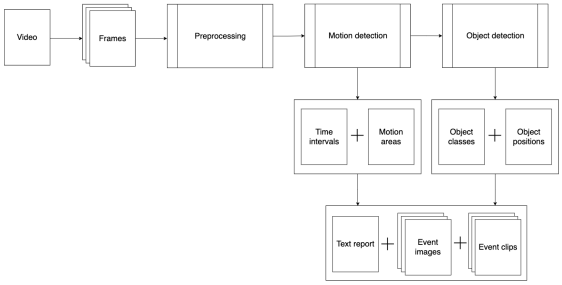


Fig. 1. The diagram of the video processing workflow in the application

A. Preprocessing

Each video frame is first reduced to 640×640 pixels to speed up and optimise processing. The reduction in resolution allows faster processing times and creates a more optimal environment for the next step of image processing. The frame colour map is then converted to grey using a technique called greyscale conversion, which simplifies the image and allows smoothier processing. After greyscale conversion, potential noises are removed by means of Gaussian blur, a technique which blurs the image in a way that removes any noise that may affect further processing. This ensures that the image is as clean as possible and ready for the next stage of processing [10].

B. Motion detection

After preprocessing and cleaning the video frames, the Frames Differences algorithm is applied to detect motion in the video. The algorithm first compares the first frame of the video with the current frame and identifies the areas where changes have occurred. The threshold value is an important factor to consider. If it is set to too low, even the smallest changes, such as the movement of leaves on a tree, will be detected as motion. Conversely, if it is set to too high, some true movements may be missed [11]. By detecting these changes, the algorithm determines where motion has occurred in the video [12].

Following the detection of motion, the overlapping and nearby regions are merged to create a more accurate representation of the location where the motion occurred. This helps to eliminate any false positives that may have been detected from the initial motion.

Finally, the data is segmented into time periods in which the event occurred. This allows the identification of specific intervals in which motion was detected, which can be useful for later analysis and processing of the video.

C. Object detection

In the object detection module, the middle frame capturing the object of interest is calculated for each time interval. The algorithm then analyses these frames and extracts key features to identify the object of interest. This process involves a series of complex calculations, including edge detection, feature extraction, and neural network classification. The YOLOv5 neural network is particularly effective at this task, as it can accurately classify objects and provide their position coordinates with a high degree of accuracy.



Fig. 2. Processing results of the Frame Differences algorithm. Above: Two consecutive frames in a video sequence. Bottom: Left, the difference in pixel values of the two frames. Right, the result after thresholding and gain-based filtering. The source of the video from which the frame is taken: [13].

However, not all objects that are initially detected are of interest. In fact, objects that are not in the list of the areas where motion has been detected are marked as stationary and filtered out. This helps to reduce the amount of unnecessary data and increases the overall efficiency of the object detection module. Besides, the filtering process can be fine-tuned to ensure that only the most relevant objects are being detected and processed, further enhancing the accuracy and effectiveness of the system.

D. Representation of results

The final step in the process is to create a text report and video files of events based on the time intervals and object data received from the previous modules. The report provides a detailed overview of the events that occurred in the selected video files, including information about moving objects and the times of occurrence. Timestamps for the start and the end of movements are used to create video clips, which are then saved to the results folder along with the text report. These video clips provide a visual representation of the events, which can be useful in understanding the context and sequence of actions that took place.

Apart from the report and video files, the system stores frames in which moving objects were detected. The objects in these images are outlined by bounding boxes and labeled with their corresponding class, such as ”person” or ”car”.

IV. Measures

The performance of the application was tested using two different video examples with varying properties, as shown in Table I. The goal of these measurements was to evaluate the speed and efficiency of the application in processing motion data from the recorded videos. The results of the measurements are presented in Table II.



Fig. 3. Example of an image result that shows the detected object. The source of the video from which the frame is taken: [14].

Motion detection report

Processed video: *video_720p.mp4*

Start processing: *19.02.2023 18:13:33*

End processing: *19.02.2023 18:15:05*

Processing time: *0:01:32*

Number of detected motions: *3*

Detected motions

Start time	End time	Objects number	Objects	Clip of motion	Image of motion
0:02:28	0:02:59	2	person, person	video_720p_0_02_28-0_02_59.mp4	video_720p_0_02_28-0_02_59.png
0:10:35	0:11:06	1	auto	video_720p_0_10_35-0_11_06.mp4	video_720p_0_10_35-0_11_06.png
0:12:36	0:13:08	1	person	video_720p_0_12_36-0_13_08.mp4	video_720p_0_12_36-0_13_08.png

Fig. 4. Text report on the results of the motion detection process

The tests demonstrated that the application is capable of processing videos at a much faster rate than a person could do manually. Depending on the resolution of the video, the application can check videos between 3 to 12 times faster than a human operator. This makes the application ideal for the situations where there is a large amount of video data to analyse or where time is of the essence.

TABLE I
MEASURED VIDEO PROPERTIES

Video num.	Duration	Frames	Resolution	Format
1	20 min	37355	1280×720 px	MP4
2	20 min	40130	2560×1920 px	MP4

TABLE II
THE MEASUREMENT RESULTS

Measurement type	Video num.	Speed	Time
Motion detection only	1	393.21 fps	95 s
	2	118.03 fps	340 s
Motion and object detection	1	377.32 fps	99 s
	2	113.68 fps	353 s

V. CONCLUSION

This article introduces a desktop application for analysing recorded video clips and extracting motion data. The application is written in Python and utilises the YOLOv5 neural network model along with the Frame Differences motion detection algorithm. By using this solution, videos can be checked between 3 to 12 times faster, depending on the resolution, compared to manual inspection. Future plans for the application include adding the ability to select specific objects for tracking, in order to provide users with desired results and enhance their overall experience.

REFERENCES

[1] "Qt for Python," [online] https://doc.qt.io/qtforpython
[2] "Qt — Development Framework for Cross-platform Applications," [online] https://www.qt.io/product/framework
[3] W.Joshua, "Overview of PyQt5," In "Modern PyQt: Create GUI Applications for Project Management, Computer Vision, and Data Analysis," Apress, Berkeley, CA, 2021, pp. 1-42, doi:10.1007/978-1-4842-6603-8_1",
[4] "OpenCV: OpenCV-Python Tutorials," [online] https://docs.opencv.org/4.x/d6/d00/tutorial_py_root.html
[5] "About - OpenCV," [online] https://opencv.org/about
[6] "YOLOv5 - Ultralytics — Revolutionizing the World of Vision AI," [online] https://ultralytics.com/yolov5
[7] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.
[8] M. Horvat, L. Jelečević, G. Gledec, "A comparative study of YOLOv5 models performance for image localization and classification," 33rd Central European Conference on Information and Intelligent Systems (CECIIS), Dubrovnik, Croatia, 2022
[9] "ffmpeg Documentation," [online] https://ffmpeg.org/ffmpeg.html
[10] P. Benický and J. Ladislav, "Real Time Motion Data Preprocessing," Journal of Electrical Engineering, vol.61, no.4, 2011, pp.247-251, doi: 10.2478/v10187-010-0035-2.
[11] A M Husein, Calvin, D. Halim, R. Leo and William, "Motion detect application with frame difference method on a surveillance camera," Journal of Physics: Conference Series, vol.1230, no.1, 2918, pp. 12-17, doi: 10.1088/1742-6596/1230/1/012017
[12] G. Shi, J. Suo, C. Liu, K. Wan and X. Lv, "Moving target detection algorithm in image sequences based on edge detection and frame difference," 2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 2017, pp. 740-744, doi: 10.1109/ITOEC.2017.8122449.
[13] Vibrant Films, "Cruise Self Driving Car," 2015, Vimeo https://vimeo.com/145559428/
[14] SlashGear, "Nest Cam Outdoor security capture video sample," 2016, YouTube https://www.youtube.com/watch?v=kjRmgDJC3ZM

Adaptive Angle Tracking Observer

1st Štěpán Černý
Brno University of Technology
Faculty of Electrical Engineering and Communication
Department of Control and Instrumentation
616 00 Brno, Technická 3082/12, Czech Republic
E-mail: xcerny73@vutbr.cz

2nd Ing. Dominik Friml
Brno University of Technology
Faculty of Electrical Engineering and Communication
Department of Control and Instrumentation
616 00 Brno, Technická 3082/12, Czech Republic
E-mail: dominik.friml@vutbr.cz

***Abstract*—This paper deals with rotor velocity estimation using an algorithm called Angle tracking observer, which is designed and tested in MATLAB simulation environment. It then focuses on tuning the Angle tracking observer parameters using an optimization criterion, analysis of parameters and their use for adaptive Angle tracking observer. Then the Angle tracking observer is converted into a discrete form.**

***Index Terms*—Angle Tracking Observer, Velocity estimation, ITAE criterion**

I. Introduction

To estimate velocity and position sensors are used , but their measurements are not good enough, so they need to be improved. One way to improve the measurement is by using Angle tracking observer (ATO).

The aim of the work is to design and simulate an ATO, that will be able to estimate the engine velocity, to tune suitable ATO parameters using the ITAE criterion and test their performance. The subsequent intention is to replace the conventional ATO with an improved adaptive ATO and followingly to modify the ATO for practical use.

II. Basic Angle Tracking Observer

The most commonly used type of ATO is shown in Fig. 1. The ATO contains two integrators, which means it is a second order system. The input to the ATO is the signal from the velocity sensor, the goal of the ATO is to convert this signal to velocity. So the output of the controller is the angular velocity and by integrating it we get the roll angle. This circuit can also be approached as a PI controller in series with an integrator. [1] [2] [3]

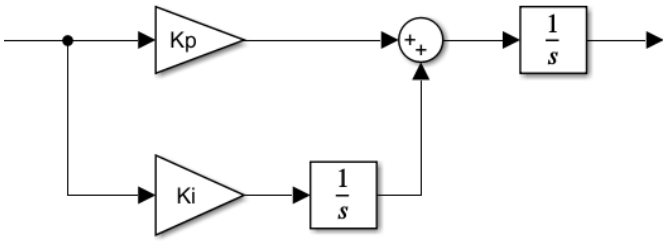


Fig. 1. Basic ATO scheme

A. Simulation of ATO

The proposed simulation circuit is shown in Fig. 2. The transfer function determining the motor dynamics was taken from [1]. It is approximated by $\frac{w^3}{(s+w)^3}$, $w = 100$. The simulation assumes the use of an incremental velocity sensor with 1024 bits per revolution. The output of this encoder is a discrete sawtooth waveform ranging from π to $-\pi$. The discontinuity from π to $-\pi$ present a problem for the PI controller to cope with, so the signal from the sensor is decomposed into $\sin$ and $\cos$. [1] [2] [3]

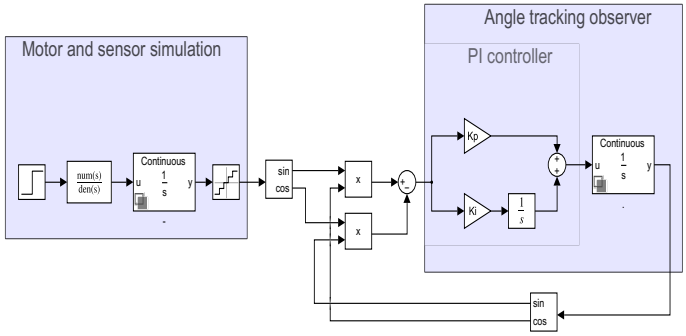


Fig. 2. ATO scheme for simulation

B. Parameter tuning

It is crucial for the accuracy of the ATO that the PI controller parameters are chosen appropriately. The first step is to find parameters for a particular desired velocity. These were experimentally chosen to be 50 for the proportional component, and 20 for the integration component of the controller at a velocity of 24 rad/s. In Fig. 3 we can see how the ATO with these parameters (blue waveform) can track the reference velocity (red waveform).

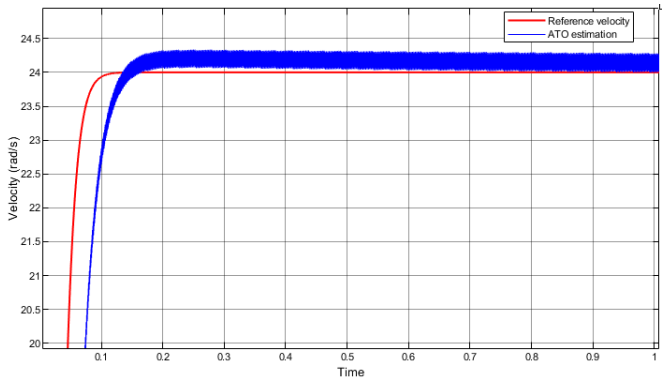


Fig. 3. Control with functional parameters

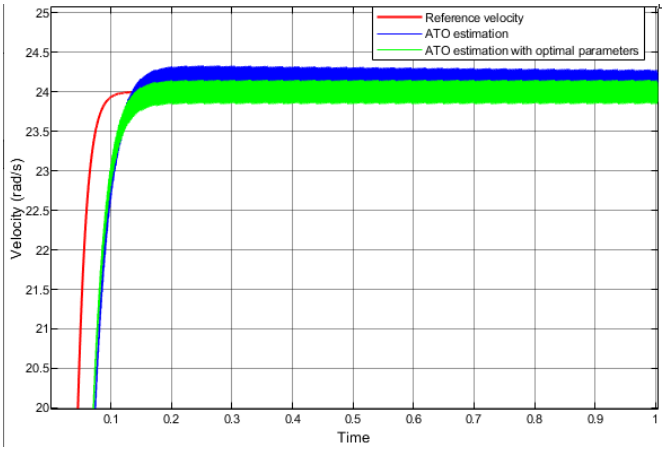


Fig. 4. Control with optimal parameters

The goal now is to find the optimal parameters and thus improve the ATO properties.

There are several methods for finding the optimal fitting parameters. In this paper, the ITAE (Integral Time Absolute Error) criterion was chosen for optimization.

$$\text{ITAE} = \int_0^\infty |e(t)| t \, dt \tag{1}$$

The equation (1) shows that the integral of the absolute value of the time-weighted deviation e(t) is calculated. This means that the smaller the deviation, the smaller the ITAE value and the better the estimation.

In general, this criterion does not have an analytical solution due to the presence of a non-linear function of absolute value. Solution can therefore only be obtained numerically. In practice the integration is not computed in $t \to \infty$, it ends when the response reaches steady state, or when the simulation ends. Since the simulation runs in discrete time, the integration is replaced by summation and the time by n steps. The modified equation is then (2).

$$\text{ITAE} = \sum_{m=1}^{n} |e(m)| \tag{2}$$

The ITAE value alone however does not say anything, it is necessary to compare several values among each other that were obtained under the same conditions such as the system, step and simulation time used. As such, the ITAE does not serve directly to obtain the optimal parameters, but determines the quality of the control.

The simulation was carried out in MATLAB environment, where the implicit function "fminsearch" was used to find the optimal parameters for the PI controller based on ITAE. The fminsearch function requires initial parameters where the original experimentally selected parameters were just used. The simulation time was 5s and the simulation step was 10^{-4}.

The result of the optimization is shown in Fig. 4. Comparing the waveforms, it can be seen that the dynamics of both ATOs is similar, but in the case of using the ATO with optimal parameters (green waveform) the non-zero deviation is removed. [4] [5] [6]

Different optimal parameters correspond to different velocities. As it is not possible to obtain optimal parameters for all various velocities, a vector of velocities, from two to fifty with step two, was chosen for which the parameters were found.

So at this point we have 25 different velocities and the ideal controller parameters for each of them. However, we are limited by the fact that we have to define the reference velocity we want to track and choose the parameters based on that. The assumption is that the motor velocity will change dynamically, so an analysis was made to see how an ATO with the proposed parameters for a given velocity can estimate the other velocities.

As mentioned before, ITAE determines the quality of the estimation, so we have to define the correct ITAE value ourselves according to the quality of estimation we require. Let's say for our purposes this value has to be lower than 1.5×10^4.

If we again use the optimal parameters for a speed of 24 rad/s, we see (Fig. 5) that the ITAE increases linearly with increasing speed, but the maximum value is 13 000, so these parameters meet our requirements.

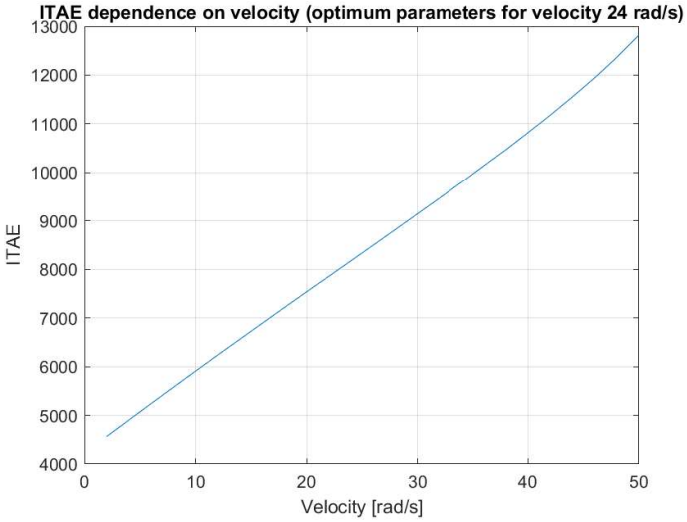


Fig. 5. ITAE dependence on velocity with optimum parameters for velocity 24 rad/s

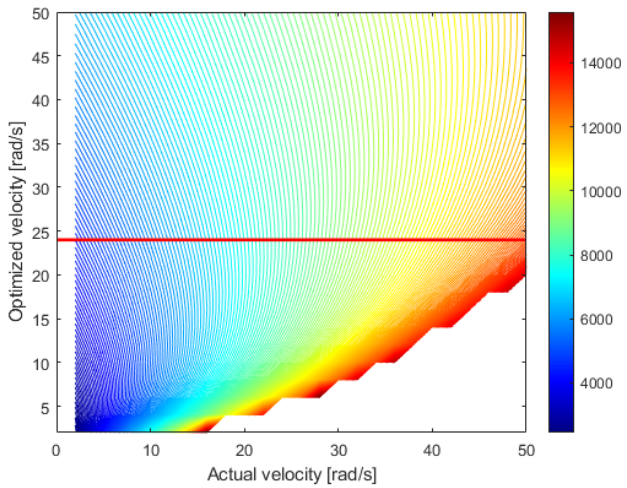


Fig. 6. Relationship between optimal parameters and different velocities for compliant ITAE

Subsequently, a more extensive analysis was performed for the combination of all speeds with all parameters. The results of the analysis are shown in Fig. 6. The X-axis represents the vector of selected velocities and the Y-axis represents the optimized velocities, respectively optimal parameters for the given velocities.

It can be seen that using these parameters, and other parameters for lower velocities, ITAE exceeds the maximum allowed value at higher velocities. This means that the ATO is unable to estimate under these conditions. Therefore, it is necessary to choose parameters of higher velocities, which although have larger ITAE at lower velocities, allow sufficient angular velocity estimate for all velocities.

The specific choice of parameters depends on the application. Assuming that the motor velocity will change frequently, it is most appropriate to select parameters for medium velocity. If we know that the motor velocity will be, for example at high velocities most of the time, it is better to choose parameters for higher velocities.

C. Adaptive ATO

Another option for tuning parameters is adaptive ATO, which retrieves optimal parameters from a look-up table based on the corresponding velocity. The lookup must be delayed by one step due to the algebraic loop, meaning that the parameters from the look-up table will never exactly match the current velocity. However, the assumption is that adaptive tuning is preferable to using constant parameters.

A linear interpolation of the values in the look-up table can be used, allowing the parameters to be calculated even for velocities outside the selected velocity vector.

In Fig. 7 comparison between the adaptive ATO (red waveform) and the ATO with parameters for a velocity of 24 rad/s (yellow waveform) at different velocities is shown. At the beginning, at 7 rad/s, the waveforms of the two ATOs are almost the same, but as the velocity increases, the dynamic properties of the adaptive ATO improve significantly. This is also true when going from high to low velocitiess. In the ramp case, the waveforms are again almost identical.

III. Discrete substitutes for continuous integration

Another goal of this work is to replace the continuous integrators of ATO with discrete integrators in order to allow practical implementation where discrete integration is necessary. Trapezoidal replacement (3), backward Eulerian integration replacement (4) and forward Eulerian replacement (5) have been proposed. For our purposes, the replacement is implemented only for the PI controller integrator, but the replacement of the second integrator will be similar.

$$y(n) = y(n-1) + K(t(n) - t(n-1))\frac{(u(n) + u(n-1))}{2} \quad (3)$$

$$y(n) = y(n-1) + K(t(n) - t(n-1))u(n) \quad (4)$$

$$y(n) = y(n-1) + K(t(n) - t(n-1))u(n-1) \quad (5)$$

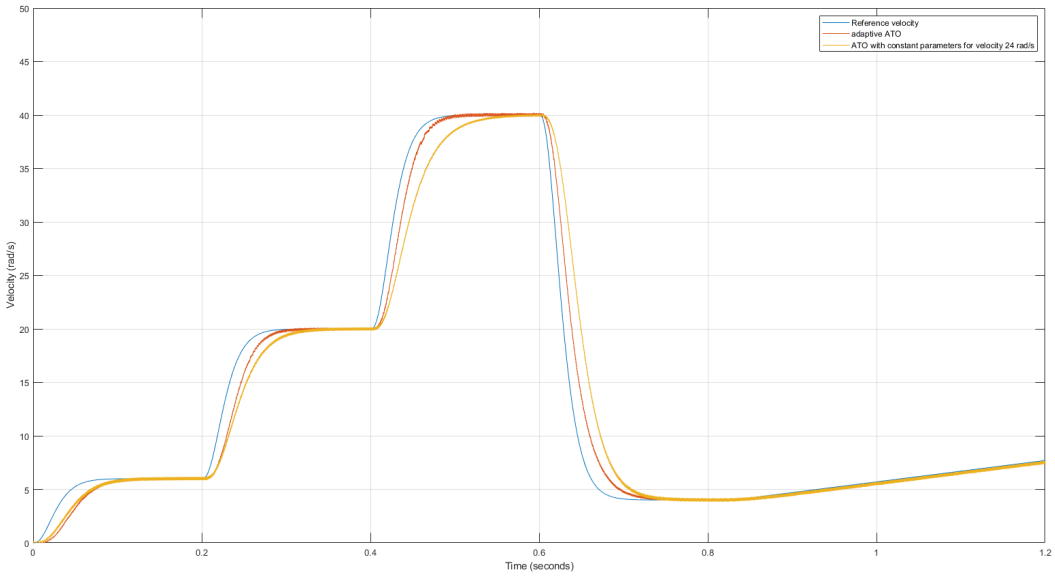


Fig. 7. Adaptive ATO vs ATO with constant parameters

The discrete substitutes were designed in MATLAB Simulink environment (Fig. 8) according to the formulas where $y(n)$ is the current output, K is the gain of the integration component of the PI controller, the time difference is the step size which is 0.0001 and $u(n)$ is the input. The correctness of the designs was verified with Simulink discrete integrators. [7] [8]

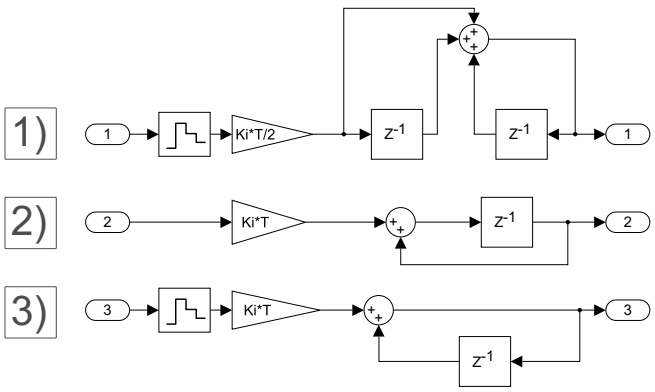


Fig. 8. 1) Trapezoidal replacement 2) Forward Eulerian replacement 3) Backward Eulerian replacement

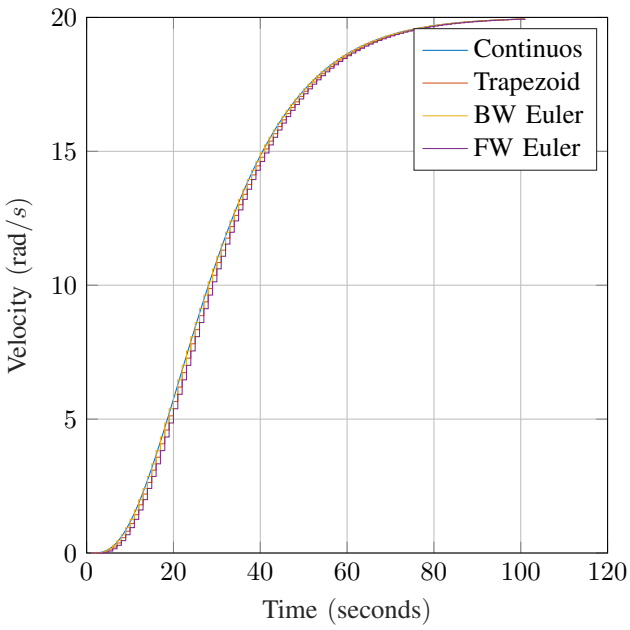


Fig. 9. Discrete integrators vs continuous integrator

A. Comparison of ATO results with discrete and continuous integration

Simulations of all ATOs with discrete replacement integration were performed under the same conditions as continuous ATOs. The results obtained for all three substitutes are similar and do not differ in principle. Comparison of discrete substitutes with a continuous integrator is shown in the Fig. 9.

Since the results of the discretizations are similar, only the Euler backward method is published. From the results (Fig. 10 and Fig. 11), it can be argued that the discretization of the integrator did not significantly affect the properties of the ATO and in practical application, the proposed algorithm will be functional.

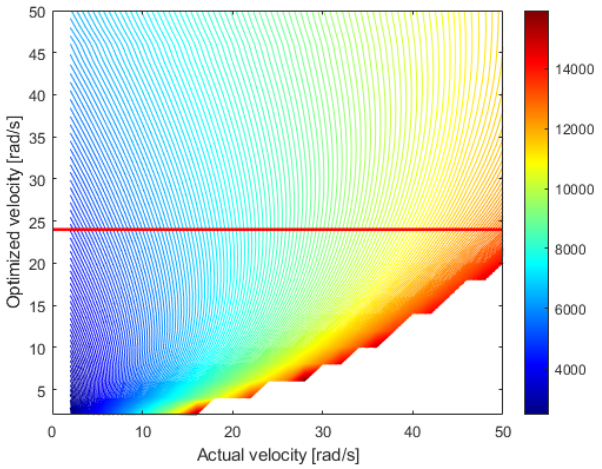


Fig. 10. Relationship between optimal parameters and different velocities for compliant ITAE - Backward Euler

IV. Conclusion

The ATO was designed and the optimal parameters were designed using ITAE criteria. Based on the simulation, it was found that the parameters for low velocities, obtained specifically by this criterion, are inoperative for high velocities. Hence, either larger parameters have to be chosen or adaptive parameter tuning must be used.

It was found that the proposed adaptive ATO can better estimate the velocity during the transient, especially at high velocities, than the ATO with constant parameters, which is a significant improvement. The only drawback is its computational cost.

All three discrete replacements were designed correctly. Their results are similar and consistent with the properties of the continuous ATO.

Acknowledgment

The completion of this paper was made possible by the grant No. FEKT-S-23-8451 - ”Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence, automation and measurement” financially supported by the Internal science fund of Brno University of Technology.

References

[1] BARTIK, Ondrej a Ludek BUCHTA. Angle tracking observer for the velocity estimation with the filtered control variable. In: ELSEVIER. 2019, s. 6.

[2] JARZEBOWICZ, Leszek, Maciej CISEK a Artur OPALINSKI. Angle Tracking Observer for Filtering Rotor Position Estimates in Sensorless Electric Drives [online]. 4 [cit. 2023-03-12]. doi:https://doi.org/10.5755/j01.eie.22.5.16341

[3] P. Modesto, R. Cordero and A. D’Estefani, ”Design of an Angle Tracking Observer for Resolver Sensors using a Type-II System and Root Locus Technique,” 2020 IEEE XXVII International Conference on Electronics, Electrical Engineering and Computing (INTERCON), Lima, Peru, 2020, pp. 1-4, doi: 10.1109/INTERCON50315.2020.9220224.

[4] BLAHA, Petr a Petr VAVŘÍN. Řízení a regulace I [online]. Brno, Česká republika: Vysoké učení technické v Brně, 2004 [cit. 2023-03-13].

[5] STRNAD, Martin. Srovnání metod regulace stejnosměrného pohonu. Plzeň, 2020. Bakalářská práce. ZÁPADOČESKÁ UNIVERZITA V PLZNI.

[6] MARTINS, G. Fernando. Tuning PID Controllers using the ITAE Criterion [online]. 7 [cit. 2023-03-13].

[7] Euler integration method for solving differential equations. X-engineer [online]. [cit. 2023-03-13].

[8] MathWorks [online]. [cit. 2023-03-13].

Analysis of BLDC motor Six-step and field oriented control

Adam Kečkéš
dept. of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
xkecke01@vutbr.cz

***Abstract*—This paper presents two approaches to motor control. Firstly, a Six-step algorithm is presented, followed by a field-oriented control approach. Both algorithms are demonstrated with the BLDC machine as the simulation done in Simulink. A simple description of the BLDC is given as well. The paper concludes with the simulation results.**

***Index Terms*—six-step, field oriented control, foc, bldc, dc motor, motor control**

I. Introduction

When it comes to modelling and control of brushless direct current (BLDC) machines it is necessary to ensure information about the immediate position of the rotor and the motor speed. For this reason motors are equipped with resolvers, incremental sensors or even hall sensors. This sensor is based on the principle of Hall effect and it measures the size and the orientation of magnetic fields, thus it responds to the rotating magnets, which are mounted on the motor. In most applications it is quite necessary so that the used motors achieve ideal or best possible performance with appropriately selected and efficient choice of the control algorithm. Due to this fact, there are several types of control algorithms for these machines.

As mentioned, this article is focused on the control of BLDC machine. Such a machines are derived from simple DC machines which are quite easy to control and which are stable in low operating speeds. Classical structure of DC motors consists of a stator with electromagnets excited by current, which has the task of creating magnetic field. The windings of the rotating armature are then connected to the mechanical commutator by graphite metal brushes which then slide over it. Subsequently, they create electrical connection between moving armature windings and external circuits. Such a mechanical commutator fulfils the role of a mechanical rectifier that converts alternating AC current by the frequency corresponding to the rotor speed of the direct current outside the armature. [1] [5]

An improvement of the simple DC motors is the creation of BLDC motors, which are also referred to as brushless. Compared to classic DC motors, they do not have the mentioned mechanical commutator over which the brushes forming the electrical node used to slip. In its place a static electronic commutator is used, whose role is the commutation of the current in armature windings in the order of the rotor angular position θ_r. Hence the principle of DC machines is preserved, i.e. the machine is still self-commutating. [1]

The second difference consists in replacing electromagnets exciting the main magnetic field with adequately selected permanent magnets. This solution is only possible due to similarity of magnetic parameters and also thanks to the lower prices of this choice design. Such permanent magnets contain the rare earth elements, such as neodymium (Nd), samarium (Sr) along others. Use of the permanent magnets has a significant impact on the efficiency due to the fact that there are no power losses in earlier used electromagnets, as well the weight of these machines is significantly lower. [1]

II. CONTROL ALGORITHMS

A. Six-step commutation

This type of algorithm is often referred to as trapezoidal commutation. It is applied especially in the control of mentioned BLDC machines and also permanent magnet synchronous machines. When it comes to the principle of this control, it consists of controlling the stator currents to ensure required motor speed and direction of rotation. When applying such an algorithm in simulation, there is the possibility of using several types of commutation modes, such as the 120 °, which is later used in the paper. This commutation mode conducts current always in only two of the stator phases. [3] [4]

Figure 1 represents the principle of Six-step commutation for BLDC motor. Especially, the mode in which there are three phases of BLDC motor excited in 120 ° sequences. The current therefore always flows through 2 of the three windings of the motor, with one winding held at the upper potential. The second winding is held at the upper potential and the last, third, winding has floating potential. Using this method, 6 possible commutation states per revolution are given. In the timing of these upper and lower arrangements, it is essential to know the information about the position of the rotor relative to the stator winding. It can be also seen, that each sector corresponds to a specific rotor angle. The simplest way of how to determine the position of the rotor is to use Hall sensors. [3] [4]

Table I contains HA,HB and HC signals which represent signals from the Hall sensors, T1-T6 represent the switching signal and the sector column represents the combination of hall sensor states which determines the sector, where the

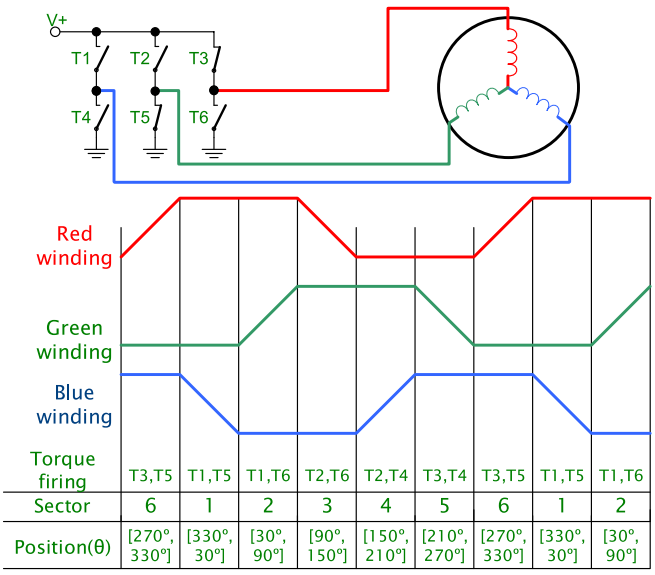


Fig. 1. Six-step commutation

Sector	H_A	H_B	H_C	T1	T2	T3	T4	T5	T6
6	0	0	1	1	0	0	0	1	0
5	0	1	1	0	0	1	0	1	0
4	0	1	0	0	0	1	1	0	0
3	1	1	0	0	1	0	1	0	0
2	1	0	0	0	1	0	0	0	1
1	1	0	1	1	0	0	0	0	1

TABLE I
TABLE OF SECTORS AND HALL SENSOR OUTPUTS

rotor is present. This sequence of sectors must be in specific arrangement, which is derived from the detection of each of the Hall sensors. By using the sequence mentioned above, the motor rotates clockwise.

B. Field oriented control (FOC)

This type of control is often referred as field oriented control (FOC), it is a control method in which the stator currents of a three-phase AC or brushless DC electric motor are identified as two orthogonal components. [2]

Currently the FOC algorithm is one of the most widely used methods of controlling alternating current (AC) and brushless DC electric motors. This method guarantees optimal operation of the motor, both in steady and transient states, it is optimal only with current controllers implemented. The goal of such control is to separate from the stator current the torque generating part and the magnetizing (flux linkage) part. Thanks to this, the same behaviour of the alternating current is achieved, as with a DC drive, where the torque can be individually controlled and magnetization of the motor. [2]

When using this method, information of the rotor position measurement is needed. Using these measured values, it is possible to decide in which position the rotor is located and subsequently it is necessary to calculate a new set of three-phase voltages which are then fed to the power inverter. As a result a new magnetic field is created, which then shifts rotor by a certain position until further measurements are taken. This process is repeated gradually and thus the smooth rotation of the motor is achieved. It is also necessary to express the spatial vectors that rely on the individual voltages, currents and magnetic fluxes in the coordinate systems which are coupled with an orthogonal system. [2]

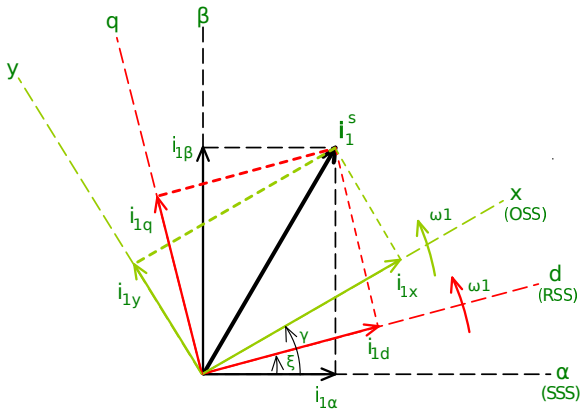


Fig. 2. Distribution of coordinate systems and their quantities

Figure 2 shows the distribution of coordinate systems and their quantities. The planes of rotor coordinate system(RSS) and oriented coordinate system (OSS) are coupled to the stator coordinate system (SSS), they also clamp the angle between them which is ϵ for SSS and RSS and the angle γ (for SSS and OSS). Due to the knowledge of the oriented angles, absolute values of the vectors, and the used motor quantities we can create a conversion between the individual coordinate systems. [2]

III. SIMULATION SCHEMES AND RESULTS

In this section, a simulation solutions for individual types of motor control algorithms will be presented. From the knowledge gained in the previous sections regarding engine management, and thus specifically the before mentioned Six-step and FOC algorithms it is possible to verifying their applicability and some real differences. The BLDC motor selected for this simulation was created by modifying the initial settings of the model, which is part of the Simscape library in the MATLAB-Simulink. The default Simscape model was modified according to the Table II. Then, the inertia is

Maximum permanent magnet flux linkage	0.03	Wb
Rotor angle over which bemf is constant	pi/12	rad
Number of pole pairs	3	-
Stator d-axis inductance, L_d	0.00022	H
Stator q-axis inductance, L_q	0.00022	H
Stator resistance per phase, R_s	0.013	Ω
Rotor inertia	0.001	$kg \cdot m^2$

TABLE II
TABLE OF BLDC PARAMETERS

represented by a separate block in the modeling scheme, such

a step is necessary for the subsequent measurement of the motor torque. The configuration for this simulation: sampling period was 1 ms, the used solver method was Euler and the duration of the simulation was 0.5 s.

A. Six-step simulation

For the correct function of the Six-step algorithm, it was necessary to create a function that converts the signal from Hall sensors distributed in the motor to the switching sequence of individual voltages according to sectors. The created model was designed with these considerations in mind, it is shown in Figure 3.

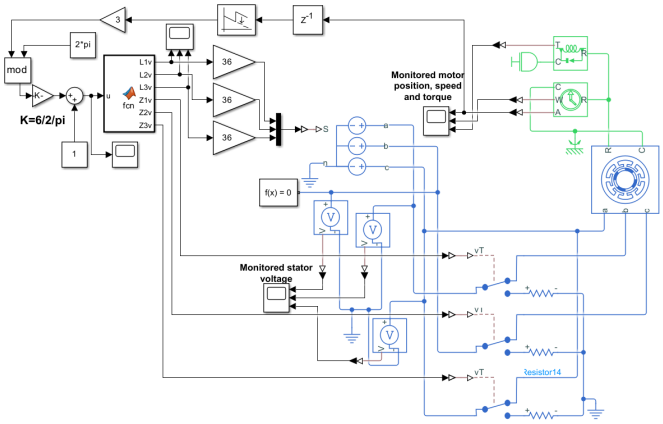


Fig. 3. Six-step control model of BLDC machine

The Matlab function block fulfils the task of switching individual sectors and then its output is sent to controlled voltage source, which creates voltage for individual motor windings. This algorithm connects always one phase to the positive pole, second to negative and the third one is disconnected. Core of this function is implemented via state machine. Due to the fact that Simulink does not work natively with a 'floating' phase, it was necessary to create it. In the case of this state, a value (0.5) is sent to the output of the phase, which creates a distinguishes from the positive and negative values of the phases. It is not necessary to set this value to 0.5. The reason why it is that way is so the signals are differentiated.

Subsequently, it is necessary to use an single pole double throw (SPDT) switch that ensures this 'disconnection' of phase by pulling the voltage across the resistor (10 MΩ) to the ground. Due to the conversion between mechanical to angular position the Unwrap block is used and multiplied by number of pole pairs. Further, the position feedback is limited to $< 0; 2\pi >$, and it is biased by one to simulate number of quadrants (one to six).

B. Field oriented control

For the realization of vector control for modified BLDC motor in MATLAB-Simulink environment, the simscape library was also used. The principle of this control is described in the previous subsections II-B. The simulation scheme that was used for this control is displayed on the Figure 4.

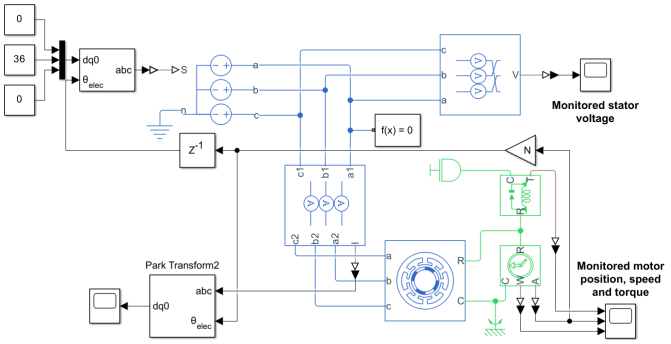


Fig. 4. FOC model of BLDC machine

The feedback loop consist of current position of rotor multiplied by number of pole pairs. It is also necessary for every Simscape component to have its own solver configuration as seen in both Figure 3 and 4. This model scheme is much simpler than Six-step model scheme because,the most of the used blocks are included in MATLAB libraries. Same as in the case of Six-step control, the BLDC motor is excited by 36 V.

C. Results

The main difference between both algorithms is visible in the measured magnitudes of revolutions and torque. These outputs from both of the control algorithms are shown in Figures 5, 6. Both of these simulations are used only to show the differences, they also depend on the choice of BLDC motor and they do not consider additional application.

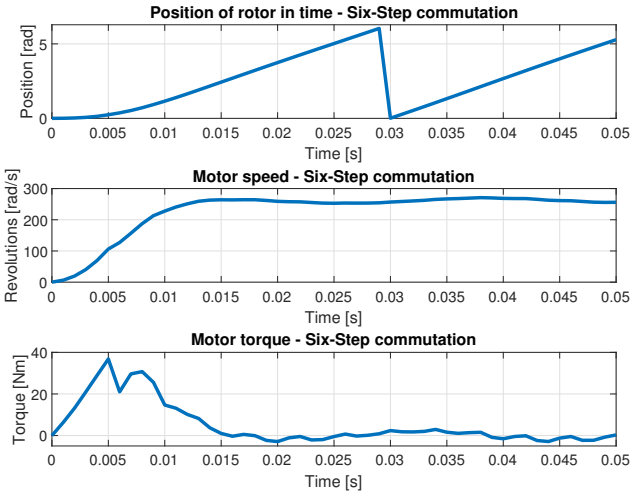


Fig. 5. Measured characteristics of BLDC motor for six-step control

The primary difference can be seen in the dependence of revolutions on time. When using Six-step commutation maximum output speeds are much higher then in Vector control. From the start of the simulation, the output start to oscillate in the range of ± 7 rad/s, however this oscillation settles down to a smaller value. It is advisable to use a controller, which could

limit this oscillation and thus improve the transient behaviour. The main problem with the Six-step method can be seen in the engine torque graph. In this case large torque pulsations occur, which are among the main disadvantages of this solution. These pulsations are mainly caused by the principle of this control, where motor is controlled by switching of the phases, the full principle was explained earlier, in Section II-A.

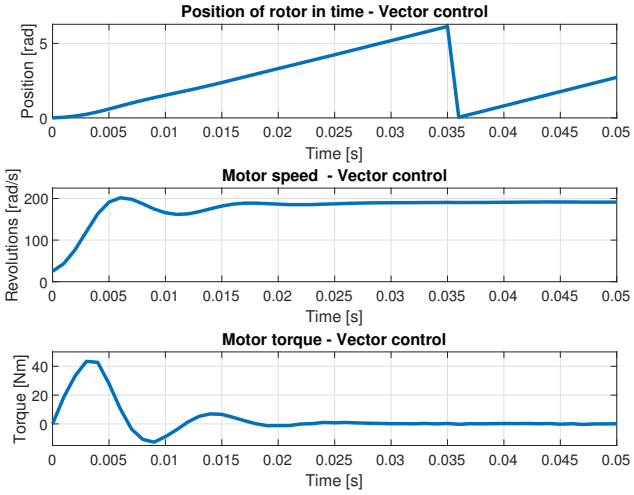


Fig. 6. Measured characteristics of BLDC motor for FOC

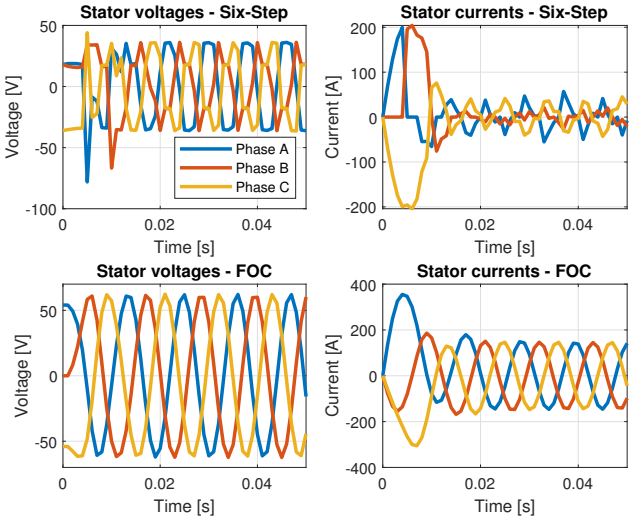


Fig. 7. Measured stator voltages and currents of BLDC motor

When it comes to the use of FOC, from Figure 6, it can be seen that there are no longer those oscillations in such an extent, as with the previously mentioned Six-step commutation. As mentioned in Section II-B, it is also advised to use some sort of a controller to improve the transient behaviour.

To calculate the efficiency of both control algorithms, it was necessary to measure mechanical output power and electrical input power. This dependency of electrical input power was acquired by multiplying stator voltages and currents. The mechanical output power was then calculated by multiplying mechanical revolutions and torque. Then the average value of these dependencies was taken. Final efficiency of motor was obtained by dividing the mechanical output power by the electrical input power. It can be said that the efficiency of motor controlled by Six-Step commutation is lower than that of FOC. It can also be said that FOC consumes less energy to perform the same task as the Six-step commutation. The rotor speed of the machine with FOC is kept at a lower $\mathrm{rad/s}$, as seen in Figure 6. In terms of maximum power output, it manages to handle higher peak torque, as seen in Figure 6.

From Figure 5 it can be seen that the Six-step commutation has a lower peak torque and a higher rotor speed. This is because the energy is getting lost in the visible torque swings. Even the transient behavior of this commutation is slower, this is connected to the peak output power.

IV. Conclusion

Thanks to these simulation outputs, it can be seen, why Field oriented control is more frequently used that Six-Step commutation.

Experimental data has shown the possibility of Vector control to manage higher peak output powers, when control algorithm is concerned, despite the same supply voltage connected to BLDC windings.

The results open new possibilities for experiments, e.g., testing presented models on a real physical systems, other BLDC motors, or even PMSM machines. There is also a possibility for further applications and more complex simulations.

Acknowledgment

The completition of this paper was made possible by the support of Ing. Ondřej Bartík. His useful advice and suggestions were really helpful to me during the project´s completion.

References

[1] Brushless DC Motor Drives (BLDC). In: WACH, Piotr. Dynamics and Control of Electrical Drives. Berlin, Heidelberg: Springer, 2011, s. 281-381. ISBN 978- 3-642-20222-3.

[2] CHAMRÁD, Petr. *Vektorové řízení synchronního motoru s permanentními magnety* [online]. Ostrava, 2015 [cit. 2023-03-08]. Available from: https://theses.cz/id/5kpz33/. Master's thesis. VŠB - Technical University of Ostrava, Fakulta elektrotechniky a informatiky. Thesis supervisor Martin Kuchař.

[3] Neborák, I. 2002. *Modelování a simulace elektrických regulovaných pohonů* Ostrava: Vysoká škola báňská - Technická univerzita, 2002. ISBN 80-248-0083-7.

[4] *Six Step Commutation* [online]. The MathWorks, 2020, [cit. 2023-03-08]. Available from: https://www.mathworks.com/help/mcb/ref/six-stepcommutation.html

[5] SUL, Seung-Ki. *Control of electric machine drive systems.* Hoboken: John Wiley, c2011. IEEE press series on power engineering. ISBN 978-0-470-59079-9.

Utilization of photovoltaic energy for heating the family house utility water

Tomáš Fryšták
FEKT
Brno University of Technology
Czech Republic
E-mail: xfryst02@vutbr.cz

Abstract— This article focuses on the issue of utility water heating in the family house and the design of a system that would be feasible in the layout of the given house and constitute a fully-fledged replacement, preferably an improvement of the existing system of utility water heating. The issue is viewed from two perspectives. From the point of view of using a photovoltaic system with a converter and an electrical spiral, and from the point of view of using a photothermal system where the sunlight is directly converted to heat and its subsequent storage.

***Keywords*— Heating of utility water in the house, photovoltaics, photothermal**

I. INTRODUCTION

The situation in the energy market has changed more than dramatically in the last year. Not exactly in a good direction for the average consumer. Where previously one would not have seen a problem in financing the running of a household, or a particular complication, one was suddenly looking for a way to reduce the cost of running it. It should be noted that the situation seems more stable with the last few months and energy prices are also close to their original levels. However, the project that this article deals with began to emerge just when the energy price trend was exactly the opposite. It was the rising gas prices at the beginning of last year that triggered the rebuilding of the technology for heating and heating utility water in the family house. Heating of utility water and heating was realized by a gas boiler with water heating technology using an 18 kW flow heater with a gas burner. One of the requirements influencing the subsequent heating technology was full or at least partial independence from the supply of energy from the utility networks. A second important criterion that had to be met by the new system was the fitting of new technology into the existing space of the family home, with minimal need for potential building works. When the first condition was fulfilled, it was considered from the start using solar energy, so it was chosen between photothermal and photovoltaic water heating system. Each of these water heating systems operates on a completely different principle. After considering the individual pros and cons of the individual system, during the final implementation, the option of heating hot utility water was chosen using photovoltaic panels, which, with the help of a changer, power the electric heater in the boiler. However, this system will only provide hot water if solar power is available. And that's not always the case. Overnight and possibly day-to-day absence of sunlight can cover the accumulation tank in which hot water accumulates. But typically, in the case of winter, when there is not much solar energy for heating and, for that matter, general sunshine, and for that the house must be heated somehow, the solar-powered system needs to be supplemented with a back-up source of thermal energy. In this case, it is a stove burning biomass with hot water exchanger. This ensures the heating of the house together with the heating of the utility water.

II. REASON FOR THE CHANGE IN THE INSTALLED WATER HEATING AND HEATING SYSTEM

The heating of the house was done using a fireplace and the gas boiler already mentioned. The tricky part, though, was that only three rooms in the house could be heated with a fireplace. The heat distribution was dealt with by hot air distribution systems that were not well done and heating in this way was inefficient. If the other rooms in the house needed to be heated, the gas boiler had to be started. The heat to these rooms was distributed by means of hot-water distribution and subsequently connected radiators. The heating of the utility water was not very efficient either, as it took some time to heat the water to the desired temperature, this was due to some temperature inertia of the flow heater. Before from the tap started running water at the desired temperature, a few liters had to be forgiven. This has led to inefficient use of flowing water and some user discomfort. It is clear from the description of the situation that the heating of the house was partially and the heating of the water entirely dependent on the use of gas as fuel. This was the main reason for changing the way heating and hot water heating. The average gas consumption in the winter period was 85 m^3 per month and in the summer period 25 m^3. Then the average annual gas consumption was 642 m^3. This was changed a lot after the new system was installed. The gas is now used only for cooking and its monthly consumption is equal to 2 to 3 m^3.

III. SELECTION OF APPROPRIATE TECHNOLOGY OFHEATING UTILITY WATER

As mentioned earlier, several aspects were called for in the selection of the appropriate utility water heating technology. The first and main requirement was partial or complete independence of the system from the supply of primary energy from utility networks and from the space demands of the new system. The new system must be fit into the existing premises of the family house and make use of the already constructed components of the old heat and hot water distribution system. Finally, the cost and operating costs of the new system play an important role. The family house mentioned here is located in a small village in the south of Morava, therefore the use of solar energy was considered right from the outset. The sun in this area reaches the highest average annual solar impact values in the Czech Republic per square meter. Specifically, the amount of energy hitting the horizontal level in these parts of the republic is around 1 000 to 1 200 kW/m^2 per year. Similar values can be achieved for a south facing installation with optimal slope. The south side is also oriented to the part of the roof that is designed for the installation of equipment using this solar energy, either solar panels or hot-water collectors. The decision on the final solution will be decided in the following sub-chapters.

A. Photovoltaic system

In the case of the use of a photovoltaic system for the heating of utility water, the following configuration is considered. As an element ensuring accumulation and water heating itself, a boiler with an electric heater was chosen as a source of heat. The electrical energy needed to power the heater will be generated by photovoltaic panels. To avoid the need to adjust the electric connection of the boiler and to avoid dc switching, a converter will be inserted into the assembly between the photovoltaic panels and the boiler.

1) Boiler

Several input parameters need to be determined to select the appropriate boiler. The first requirement is a compliant boiler volume. This is essentially based on the number of people using hot water and the nature of the use of hot water. Other uses of water will be in offices, where water is used to wash hands and coffee mugs, and vice versa in the home. Here, hot water is used for many more things, from washing hands, to washing dishes, to washing like a shower, or a bath. This clearly implies that water consumption in the home is slightly higher. Another important aspect is where the boiler itself will be placed, whether it will be able to anchor itself to the wall or need some sort of support structure. Whether it will be installed horizontally or vertically. And if it even gets its dimensions into the spaces it's supposed to be installed in.

According to the previous criteria, a 148 l boiler with a 2.2 kW installed electric heater and a 9 kW hot-water exchanger, at a water temperature of 80 °C and a flow rate of 310 l/h was selected with the commercial designation OKC 160. This boiler is designed for vertical installation with a wall anchorage [1].

When selecting the boiler, these facts were considered. The house is occupied by 4 people with an average hot water consumption of 40 liters per person. Warm water should always be enough for one day. Heating water from 10 °C to 75 °C will require 11.1 kWh of energy, calculated by „(1)" and this heating will take 5.1 hours to consider the performance of the installed heating coil.

The following equation was used to calculate:

$$Q = m \cdot c \cdot (t_h - t_c) \quad (1)$$

Where Q is the amount of energy that needs to be delivered to a certain amount of water (J), m is the mass of water that is heated (kg), c is the specific thermal capacity of water (J/K), t_h is the temperature of hot water (°C) and t_c is the temperature of cold water (°C).

2) Converter

The performance of the heating coil is already known, so the design of the changer may be initiated. An uncountable number of changers can be found on the market. The first thing to determine here is what is required of the convertor. First and foremost, the changer should be able to use the maximum available performance of the panels, to have the MPPT function, or its equivalent. Next, it is important to note what this changer will be used for. In this case, it will power a resistance heater of 2.2 kW. The value of the output voltage shall be the same as the value of the foaming on which the heating spiral of the boiler is built, in this case the effective value of this voltage is 230 V. However, such a changer, which produces a sine waveform on the output clamps, costs something like 15,000 CZK. However, only a signal that periodically passes through zero at a certain interval is sufficient to feed the resistance load. It is this fact that will allow the heating coil to be connected without the necessary modifications to the electric connection of the boiler and the original switching mechanism can be used without risk of excessive wear and subsequent destruction. These changers are more affordable and fully adequate to power the heating coil, however, their disadvantage is that these changers cannot be powered by anything other than resistance load, which in this case doesn't matter, as it will be its main and only function. As for the price of these convertors, this starts somewhere around 1/3 of the price of converters of similar power with sinusoidal output voltage, around 5,000 CZK.

The MR4316AC NG changer has been selected for use in this installation. This transducer has two outputs, either DC, directly from the photovoltaic panels, or an alternating rectangular shape with a frequency of 121 Hz and a maximum voltage of 250 V. Panels with a maximum voltage of 60 to 430 V DC and a current of 16 A can be connected to the input of the changer. Photo of the converter and switchboards is on the figure 2 [2].

3) Photovoltaic panel

To use the maximum power of the panels even under changing lighting conditions, the already mentioned convertor is used in the assembly. It is tasked with monitoring or actively searching for Maximum Power Point. The MPP is the point on the Volt

Ampere characteristic of the solar panel in which the panel delivers the highest performance. Without the use of a convertor with a function that allows this point to be searched, the use of available energy would be inefficient. Since the selected transducer can only reduce the voltage, the input DC voltage from the panels must be higher than the voltage required on the AC output. The implication is that the panels will need to be connected to the series, as the voltage of a single panel usually moves lower than the voltage required to power the chosen boiler. Next, the input voltage of the panels shall not exceed 430 V, even in an empty state, and the input current shall not exceed 16 A.

Determination of the electrical parameters of solar panels takes place under certain measurement conditions. And this in laboratory, this method is called Standard Test Conditions and Nominal Operational Cell Temperature. With the STC method, the intensity of the incoming solar radiation is 1,000 W/m2, the temperature of the panel is +25 °C, the spectrum AM = 1.5. Naturally, these conditions can rarely be achieved, which is why this measurement is carried out in laboratories. Conversely, NOCT measurement is intended to represent normal conditions in the operation of FV panels [3].
Photovoltaic panels Amerisolar AS-6M144-HC 450 Wp measured at STC were used for the assembly. In the table 1 are mentioned parameters needed for the calculations will be given to determine how many panels will be needed to get the desired performance [4].

TABLE I. Amerisolar brand photovoltaic panel parameters

P_{STC} [W]	450
P_{NOCT} [W]	335
U_{OC} [V]	50,2
U_{pmax} [V]	41,8
I_{sc} [A]	11,28
I_{pmax} [A]	10,77

The final number of installed panels was chosen for 7 pieces. The final installation of the photovoltaic panels is shown in figure 1. The voltage to empty in the case of seven pieces and a temperature of -25 °C is 401 V. This suits the maximum input voltage of the converter. Since the panels are connected to a series of currents supplied by them, the current will be the same as that of a single panel and this is also smaller than the maximum allowable current of the changer. The total installed power of the panels is equal to 3150 Wp. The panels will be placed on the roof with a slope of 17°, a steel structure has been used to fix them, thanks to which the panels are now installed at an angle of 38°.

B. Photothermal system

The photothermal system uses completely different principles than the photovoltaic system. While photovoltaic panels convert solar energy into electrical energy, which must be converted to thermal energy by another device, photothermal panels directly transform solar energy into thermal energy. The heat energy generated is then transferred by means of a heat transfer agent, in the case of family homes, mostly to a heat storage tank. Collectors range from flat, to tubular, to concentration. However, the principle is very similar, with sunlight hitting the absorber, either directly or by bouncing off a spotlight, and here it transmits its energy to the heat transfer substance. Applications for single-family house make the most of a mixture of water and antifreeze. However, applications can only be found with water, oils, gas media and, finally, salts [5].

The accumulated water in the storage tanks of the photothermal system generally reaches lower temperatures than can be achieved when the water is heated by an electric heater. The implication is that the storage tank must be larger. In this case, this would demand installing a 400 liter storage tank. It was again based on the presumption that the house was occupied by 4 people. Considering an annual solar radiation of 1090 kWh/m2, collectors of 5.36 m2 would need to be installed. Which means 2 to 3 commonly available photothermal panels.

C. System selection

Both systems were designed to cover the hot water consumption of a 4 member family. In the case of the photovoltaic system, the volume of the storage tank was 148 litres, the output power of photovoltaic panels is 3.15 kWp, which occupy a roof area of 14 m2. A 400-litre accumulation tank had to be chosen for the photothermal system, and the area taken up by the flat vacuum photothermal panels would be approximately 5.4 m2. A more accurate area would depend on specific type of panel. While there was no major problem with the size and load capacity of the roof, panels of both variants can be placed on it, the interior spaces of the building are already restrictive. While the accumulation tank for the photovoltaic system, the volume and dimensions of which are smaller due to the higher temperature of the accumulated water, can be placed in the bathroom without significant restrictions. A larger accumulation tank for the photothermal system would have to be located in the basement, which is the only room in the house where it could be installed and in the process do not disrupt the character of the house aesthetically and functionally. In addition, it would be necessary to bring all connections to the water mains to the storage tank, which would mean necessary building improvements that would affect half the house. There was no need for that in the case of the photovoltaic system. In addition, the accumulation tank could be connected to the original water supply in the house without any major complications, without the need for major interventions. Because of these aspects, the photovoltaic system was chosen for implementation. Another advantage of this system is its potential extensivity, which would allow the use of electricity that would not be used to heat the water. One could use the solution of a single-phase island power plant with an accumulation of electricity in the batteries, powering certain appliances in the house.

IV. EXPERIENCE IN OPERATION

The original planned launch date for the new photovoltaic water heating system was planned for June 2022, unfortunately due to high demand for the convertor and the components needed to operate, whose delivery was delayed, the system was start up

to 22. October 2022. The season when conditions were good for power generation using photovoltaic panels passed and there was a cold and damp autumn. The sun did not appear in the sky for weeks at a time. But even in such weather, the system could produce 0.5 to 3 kWh per day. Which is definitely not enough to fully heat up the volume of the boiler, but at least it was some part of heat. The system behaved similarly during the winter, although there were a few days, mostly two to four a month, when the boiler could only be fully heated from solar energy. In the rest, the water needed to be heated by using biomass energy. The situation started to change with the end of February and the beginning of March. The sun shines longer, and the sun days are more. Indeed, several times the water in the boiler has been heated to the maximum temperature during the day and the extraction circuit has been triggered, consisting of a convection heater and a circuit that switching between the convection heater and the boiler. On such days, the amount of energy produced easily approached 19 kWh. It is the unused energy that could be used to power some appliances in the house after the system has been modified. Work is underway on the design of this extension and its economic valuation. It needs to be established if the amount of unused energy will be large enough to make this extension worthwhile. Overall, from the launch date to 12. March 2023, system produced 445 kWh.

V. CONCLUSION

The aim of this project was to displace the heating of the house and the heating of utility water from the dependence on gas, increasing the self-sufficiency of household operations and reducing operating costs. The first point was met thanks to a suitable choice of installed technology, the photovoltaic system complete with biomass-burning furnace is fully independent of energy from utility networks except off to the circulatory pump. The reconstruction itself was superbly successful, as the new system was able to connect to the original system without any major complications. The necessary construction work was kept to a minimum, which was welcoming for the actual running of the house. How much the cost of heating the house and the heating of the utility water has been reduced cannot be said now, as the heating season is not yet over and the summer season with the greatest number of sunny hours is yet to happen.

What can be based on is the fact that the average annual gas consumption in the house is 642 m3. This at the current gas price of 3304 CZK per 1 MWh and the permanent monthly payment of 292.57 CZK makes 25 890 CZK per year. This means that this amount can be saved annually. Of course, the resulting total will vary, as gas prices will not be constant [6].

VI. PHOTOS OF IMPLEMENTED SYSTEM



Fig. 1 Installation of photovoltaic panels on the roof of the house



Fig. 2 Converter and switchboards of the photovoltaic system

REFERENSES

[1] "OKC combined heater", *Dražice*, 2023. [Online]. Available: https://www.dzd.cz/ohrivace-a-zasobniky-teple-vody/kombinovane/zavesne/okc. [Accessed: 2023-03-31].

[2] "Converter MR4316AC NG", *BEL shop*. [Online]. Available: https://www.bel-shop.eu/regulator-napajeni-boileru-ze-solarnich-panelu-mr4316ac-ng/. [Accessed: 2023-03-31].

[3] V. Quaschning, *Renewable energy sources*, First Edition. Prague: Grada Publishing a.s., 2010.

[4] "Amerisolar Catalog", *Amerisolar AP*, 2022. [Online]. Available: https://www.weamerisolar.eu/wp-content/uploads/2022/10/Catalogo-AMS-rev4-2022.pdf. [Accessed: 2023-04-05].

[5] *Solar equipment in examples*, First Edition. Prague: Grada Publishing, a.s., 2013.

[6] "Price lists and terms and conditions for households", *innogy*, 2023. [Online]. Available: https://www.innogy.cz/ceniky/#plyn. [Accessed: 2023-03-31].

Design of the Energy Community Concept

Libor Pavlík
Brno University of Technology
Faculty of Electrical Engineering and Communication
Brno, Czechia
230411@vut.cz

Martin Paar
Brno University of Technology
Faculty of Electrical Engineering and Communication
Brno, Czechia
paar@vut.cz

Abstract—The article describes an analysis of energy community. The first part focuses on the introduction of energy community using the principles of generation and energy sharing at the end-use points. It describes emission-free and low-emission sources commonly used in energy community considerations and their possible applications. It further presents the municipality of Bánov, which will serve as a demonstration example for the community model and discusses the model description with a description of the consumption places for analysis. It also focuses on the description of the seasons and their influence on energy production and consumption at each place of consumption. The scenarios for which the analysis will be performed are also described. The article concludes with the basis for building a future community.

Keywords—Energy community, Municipality, Solar energy

I. INTRODUCTION

The main purpose of energy community is to decarbonise and decentralise the energy sector and to helps to increase the use of renewable energy. Its application also brings with it benefits for the members themselves in the form of more efficient use of their energy production facilities.

II. ENERGY COMMUNITY

The concept of energy community started in the 1990s in Denmark. Citizens purchased wind turbines and arranged for wind farms to be built as part of energy cooperatives and citizens' associations. The constructions were mainly financed by the members of the cooperatives and associations. The electricity produced was shared among themselves. Besides wind parks, there was also the construction of thermo-solar panels and biomass [1].

Nowadays, energy community is already represented by 9 000 communities within the European Union [2] and implementation is just beginning in the Czech Republic.

A. Definition of energy community

An energy community is an association of actors who produce energy together, from low-carbon sources, whether thermal or electrical, and also share and consume it among themselves. In the event of extreme circumstances, the energy community may resort to selling energy, but this may no longer be their primary objective [3].

B. The main goal of energy community

Communities focus primarily on the production of affordable energy. Energy is mostly sourced from zero-emission or low-emission sources. Typically, photovoltaic power plants, wind farms or biogas plants are used. Communities tend to produce most of the energy they need themselves, to be as little dependent as possible on supplies from outside the community. The creation of energy communities leads to the decentralization of generation resources and the development of an internal electricity market [3].

Motivations for the creation of an energy community can be considered as a sense of energy self-sufficiency, environmental conservation and last the possibility to reduce energy bills.

C. Possible source power

The options offered to communities for energy sourcing can be divided according to the commodity needed.

For electricity generation, photovoltaic power plants are offered as a viable option because they can be installed on already constructed rooftops. Once a PV power plant is installed, it has relatively low operating costs compared to other forms of electricity generation. PV power plants can be designed with a modular approach, allowing for easy scaling and expansion. Other sources of electricity can be wind turbines. Wind turbines are complex machines that require a lot of engineering expertise to design and manufacture. They also require careful consideration of factors such as blade length, blade shape, and turbine height, which can impact the efficiency and output of the turbine. The motivation for building wind farms is to achieve more power per built-up area. Wind power plants are complementary to photovoltaic sources mainly in covering the power deficit during the period of reduced sunlight in winter. The disadvantage of using wind power plants can be the noise level during operation [4].

The utilization of biogas plants for heat generation is highly suitable and advantageous for community use. The main reason is that the municipalities can produce the material for biogas production themselves. The methods of obtaining biogas are diverse, ranging from the use of gas from wastewater treatment, biomass processing and others. It could also be advantageous to combine a biogas plant with a photovoltaic power plant to ensure that the lower electricity production from the

photovoltaic power plant can be covered in winter. The advantage of biogas is the possibility to use cogeneration units to produce both heat and electricity [4].

III. The Analysed Municipality

The village of Bánov was selected for the application of energy community. It is located in the east of the Czech Republic, approximately 10 km from Uherské Hradiště near the Czech-Slovak border. The village has approximately 2100 inhabitants. The average age of the inhabitants is 42.4 years.

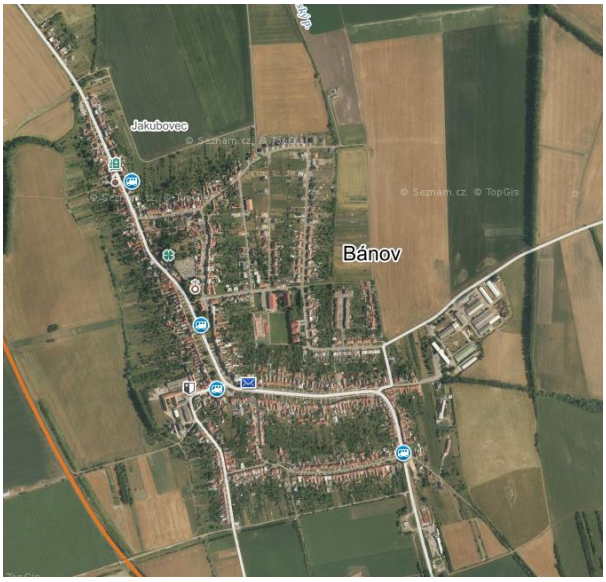


Fig. 1. Orthophoto of the municipality of Bánov ©Seznam.cz © TopGis

The village has a municipal building, kindergarten, primary school, library, sports hall, health centre, post office.

Since 2013, wastewater is treated in a wastewater treatment plant on the outskirts of the village, which is owned by the municipality. Drinking water is supplied by the group water supply system Uherské Hradiště - Uherský Brod - Bojkovice. The whole village is connected to a high-pressure gas pipeline. The village is connected to the 22 kV electricity network, which is further distributed at 0.4 kV level. In the village there is a transformer station and a regulation station for the gas pipeline. The Internet is provided in the village in the form of wireless connection and in 2020-2026 the construction of high-speed optical internet should be carried out [5].

Public lighting in the village was renovated in 2016 to ensure lower energy consumption. Light sources were replaced, and switchboards were replaced with the installation of lighting controllers. Network extension is planned for the next years [5].

IV. Description of the Model

The core of energy community can be represented by a model that will be composed of buildings with different consumption diagrams and equipped with renewable energy generation plants. The model for this analysis will use only with the resource in the form of photovoltaic power plants [6]. The generating plant's capacity will be shared among the objects. After aggregating the load diagrams of all participants, the load imbalance should be partially compensated. The resulting balance of grid overflows will be obtained by subtracting the total energy produced from the total energy consumed.

In the analysis it is important to record the different types of consumption and their behaviour under normal conditions, but also under non-standard operation, which may be caused for example by holidays and public holidays.

Buildings for municipal facilities are an essential part of the municipality. The consumption places of municipal buildings are usually scattered throughout the municipality and have different load diagrams. The consumption can vary both in size and in the distribution of load over time, or consumption pattern [6].

Types of consumption place:
- Municipal building
- Primary school
- Health centre
- Apartment house
- Sports facilities
- Cultural facilities
- Fire station

Divisions within the municipal infrastructure:
- Sewage treatment plant
- Water supply facilities
- Public lighting
- Waste collection yards
- Electric vehicle fleet and municipal charging stations
- Heating of municipal buildings by heat pumps

Division of consumption points is taken from [6].

In addition, the places of consumption that will be used in the analysis will be discussed below.

A. *The municipal building*

For the municipal building, a consumption diagram is similar to diagram of any office buildings is assumed. The building will consume energy every day except weekends and holidays when it is closed, according to opening hours. The consumption diagram of the post office will match that of the municipal office.

B. *The primary school and kindergarten*

The power consumption of primary schools will be mainly in the morning and in the afternoon until 16:00. The maximum consumption is expected to be between 10:00 and 13:00 during the operation of the school canteen. For the afternoon hours, it is anticipated that there will be more generation than consumption and thus more energy delivered to the community.

The non-standard type of operation is the period of holidays, public holidays and weekends when the consumption at the places of consumption will be almost zero. On a given day it will act mainly as a source of energy for the community.

It is assumed that the most of the energy will be consumed for lighting, water heating or heat pump heating [6].

C. The apartment houses

The natural of consumption of apartment houses corresponds to the nature of detached houses consumption. The peak value will be in the morning and evening hours [6].

Apartment buildings play an important role in energy community. In the case of the Czech Republic according to a study prepared by the EGÚ Brno [4], apartment houses suitable for the installation of a photovoltaic power plant have a potential generation capacity of 1 230 MW of electrical power. The advantages of apartment buildings include the easiest financing for the purchase of a PV plant due to the fact that more people will contribute to the purchase. Disadvantages can be more complex energy budgeting than for single-family homes.

A non-standard operation for apartment buildings is the holiday season when most residents are expected to be at home and thus the resulting house consumption will be higher [6].

D. The sports center

These are buildings designed for sports activities. They can be sports halls, gyms, etc. Their use is usually occasional. In case the heating is designed in a way other than by a heat pump, the production of the photovoltaic power plant does not usually overlap with the consumption of the building [6].

The use of battery storage is offered as a solution. Another option for the use of the energy produced is space conditioning.

E. The detached houses

Another large and integral part of the energy community sector is prosumers. These are mainly place of consumption that have their own energy source installed to cover their consumption.

Their consumption diagram will be similar to an apartment building but will vary depending on the sociological situation of the house owner. The diagram will be different for families where one parent is on maternity leave, families without children or homes occupied by retired people. Electricity can be used for water heating, heating/cooling, lighting, etc.

Similar to apartment buildings, holidays are a non-standard operation for single-family homes. The difference may arise when the occupants of the house also own, for example, in time of vacation. In this situation, the consumption of the house is only minimal and large overflows to the network will occur [6].

V. Analysis days

The use of electricity is variable. Its value depends on the power drawn from the grid and the time for which the power is used. Then, after collecting enough data, we can construct a consumption diagram. With the help of consumption diagrams we can describe the production or consumption of a given object. Their shape depends not only on the type of facility and its user, as mentioned earlier, but also on the day, season or weather.

For the purpose of analysis, weekly intervals will be selected. However, first it is necessary to establish a base week, which will be used for comparison with other weeks in which there will be non-standard operations.

A. The base week

A base week will be used to compare weekly extremes. To determine this, a week in October will be selected in which energy consumption is already increasing after the summer months, but the PV panels are still producing higher volumes of energy. Consumption of all participants is absolutely normal without any distinction.

The result should be a week that is not affected by any extreme.

B. The week during the summer holiday

The production of electricity from photovoltaic power plants reaches its highest values in the months from May to August. The peak production months are June and July. A primary school will be nested in the theoretical community, which has a large roof area for the installation of PV. We can therefore expect a large amount of energy produced from this building in the summer months. Primary schools usually have summer holidays in the months of July and August. The summer periods are also usually when energy consumption is lowest.

The assumption of this scenario will work with a large amount of energy produced from the school building and the consumption of the building reduced to an operational minimum. This is assumed to result in an energy overflow to the community, which would consume the majority of the energy.

C. The week on January

Winter months are characterised by shorter days and less sunlight. Depending on the colder weather, energy consumption tends to be higher than in other months.

This scenario will focus on a situation where all community members will be in normal operation and a higher energy supply will be required for heating and lighting. Electricity generation will be very reduced compared to the summer months.

D. The week during the Christmas holidays

The Christmas holidays are usually days when consumption at the municipal building, school and sports facilities will be limited to space tempering, while electricity consumption on the household side will be high. This is due to the winter weather and the high use of domestic appliances.

The assumption for this week is the lowest electricity production from photovoltaic power plants and thus lower volumes of energy flowing from unused buildings.

VI. Analysis Data

In this study, we utilized data from the distribution network operator to build a model that describes individual consumption points' energy usage with a 15-minute resolution. The data obtained also included the electricity generated from photovoltaic power plants at specific intervals per 1 kWp.

By utilizing this data, we generated load diagrams on a weekly basis, allowing us to compare the surplus or shortfall of energy when community power is used versus when it is not used.

VII. The Analysis community

The selection of off-take points in the energy community is a crucial aspect of developing a reliable and efficient energy infrastructure. In this respect, the available facilities in the community should be taken into account. In addition, the interconnection between municipal buildings and single-family homes should also be considered as a definitive selection criterion. Municipal buildings, including the residences in the municipality, are part of the municipality's energy system and must be integrated with other facilities to achieve optimal results. The community will be composed of the four members described in Chapter IV, A-E. The members of the community will be the municipal office, an elementary school, an ice arena and several houses.

Once the community has been formed, an analysis will be performed for each consumption point separately and the balance of energy flows to the grid will be determined for the weekly intervals and scenarios mentioned in Chapter V points A-D. Subsequently, the analysis will be performed for all consumption points, but it will be assumed that the energy is first consumed by the community members and the surplus is sold to the grid.

The ultimate goal of this study is to develop a model that describes the behaviour of the community under non-standard conditions. The proposed model will provide insight into the energy consumption patterns of the community during periods of high demand or low supply. By understanding how the community responds to non-standard conditions, energy planners can design better systems that are more resilient and can deal with unexpected events.

VIII. Conclusion

The main objective of this analysis is to simulate the behavior of the energy community and determine the amount of energy that flows into the distribution network. This data will help us determine whether the energy community concept is financially beneficial to the user. In particular, we will focus on critical times of the year, such as peak and off-peak periods.

However, certain issues may arise in the analysis that could affect the results. One of these is estimating valid results when considering the use of battery storage and electric vehicle batteries as potential energy storage. The reason for this is that these technologies may vary over different time periods, which may affect the amount of energy that will flow into the grid.

Dynamic tariffs are another factor that can affect the results. These allow users to determine when they want to take energy based on the current price. These tariffs are very flexible and can vary over time, which can have an impact on how much energy will flow into the grid.

Overall, this analysis will be very important in assessing the financial viability of the energy community. However, it will require careful planning and consideration of various factors that may affect the results. If the results are positive, energy communities may represent a new way to use energy more efficiently and reduce costs.

References

[1] *Denmark community power*, n.d., [online]. Available: https://www.communitypower.eu/en/denmark.html

[2] *European Commision, In focus: Energy communities to transform the EU's energy system*, 2022 [online]. Available: https://energy.ec.europa.eu/news/focus-energy-communities-transform-eus-energy-system-2022-12-13_en

[3] *Europea parlament, Directive (EU) 2019/944, Of the European Parliament and of the council on common rules for the internal market for electricity and amending Directive 2012/27/EU*, EUR-Lex, 2019 [online]. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L0944&from=CS

[4] M. Hrubý, M. Kocůrek, P. Liedermann, *EGU Brno study - community energy* [Studie EGU Brno - komunitní energetika], EGÚ Brno [EGU Brno], Frank Bold a. s. [in Czech], 2019 [online] Available: https://frankbold.org/sites/default/files/publikace/studie_egu_brno_-_komunitni_energetika.pdf

[5] *Development Programme of the Municipality of Bánov* [Program rozovje obce Bánov], Obec Bánov [Municipal council], [in Czech], 2019 [online] Available: https://www.banov.cz/download-file/PRO_Banov_final.pdf

[6] P. Novotný, M. Ander, M. Sedlák, *Case studies on the application of innovative solutions for energy-active consumers* [Případové studie aplikace inovativních řešení energeticky aktivních spotřebitelů], Svaz moderní energetika [Unition of Modern Energy], [in Czech], 2022 [online] Available: https://www.modernienergetika.cz/wp-content/uploads/2022/06/V2-Pripadove-studie-aplikace-inovativnich-reseni-energeticky-aktivnich-spotrebitelu.pdf


PROČ ČEKAT
NA BUDOUCNOST,
KDYŽ JI MŮŽETE
VYTVÁŘET?
Honeywell
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

careers.honeywell.com/brno
www.honeywell.com
e-mail: CZuniversityrelations@Honeywell.com

Daily Load Profile and Flexibility Aggregation

Martin Ošťádal
Dept. of Electrical Power Engineering
Brno University of Technology
Brno, Czech Republic
226969@vut.cz

Lukáš Radil
Dept. of Electrical Power Engineering
Brno University of Technology
Brno, Czech Republic
radil@vut.cz

***Abstract*—This paper deals with the concept of flexibility aggregation and daily load profile. After defining the concepts associated with flexibility and aggregation, there is a presentation of real-world data of electricity consumption from different households in New York City measured on the same day 30th of September 2019 and their impact on overall baseline graph.**

Keywords—flexibility, aggregation, baseline, peak shaving, peak shifting

I. Introduction

With current global rise in digitalization and automation affecting different sectors of industry, power engineering is not an exception. Rising number of renewable resources and the associated expected growth of decentralised production, accumulation and consumption of electricity is only contributing to this development [1][2]. Together, these factors add to increased liberalisation of energy sector and competition in the electricity market throughout the emergence of prosumers, aggregators and virtual powerplants. Rising energy prices and political efforts to reduce the volume of fossil fuels used not only for production of electricity require consumers to reduce their electricity consumption or use generated electricity as efficiently as possible [2]. Using peak shaving, load shifting, or even replacing electric home appliances with more modern, higher-efficiency devices present one of the possible solutions to achieve this goal.

II. Flexibility

A. Basic Understanding

Reference [3] defines flexibility as an ability of an electrical device to change the amount of electricity drawn from, delivered to, or accumulated within itself at a given point in time in response to price signals or commands. Biggest use of flexibility can be found in deviation coverage or regulation. Current standard management tool for electric power system is the use of flexibility of conventional large-scale power plants [4].

B. Terms Related to Flexibility

Given that there are many terms related to flexibility, only few of the most important ones will be defined, as shown below.

Baseline – the estimated value of consumption or production that would have occurred without the activation of flexibility and without the rebound effect [5]. In this paper the baseline will be used as an energy consumption reference for energy savings.

Rebound effect - a shift of consumption or production over time due to the activation of flexibility, which reduced the price of electricity, which in turn increased consumption thus partially cancelling out the initial savings. It can be said that the rebound effect is an unwanted peak shifting [1][6].

Peak shaving – a way of balancing frequency deviations on the grid to maintain grid stability, but also to reduce the price of electricity for large consumers, which in many countries is set according to their peak load. The customer reduces his consumption for short period of time by switching off certain equipment, covering his consumption with another source of electricity or by using energy stored in batteries as shown in Fig. 1. This way consumer can avoid peak load on the grid and lower its costs [7].

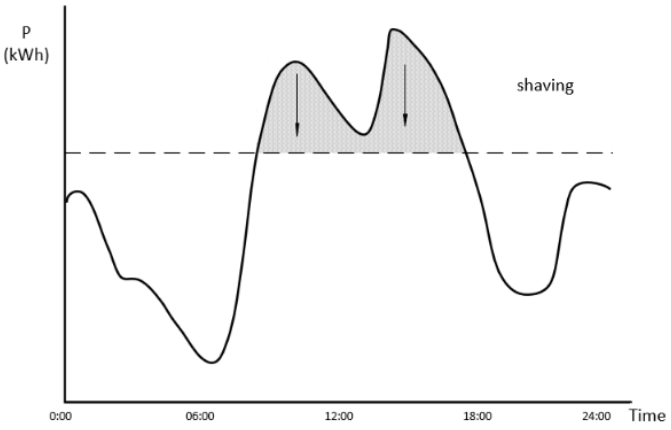


Fig. 1. Peak shaving [7]

Load shifting – a short-term reduction or increase in electricity consumption in response to price signals. The price of electricity varies throughout the day depending on the demand for electricity. During peak periods, the price of electricity is high, so by shifting consumption to off-peak periods, consumers can significantly reduce their costs. This method does not reduce electricity consumption as peak shaving does, but the consumption is just being shifted to a different time of the day, see Fig. 2. Freezers, refrigerators, heating, or dryers

can be used to evenly distribute electricity consumption throughout the day [8][9].

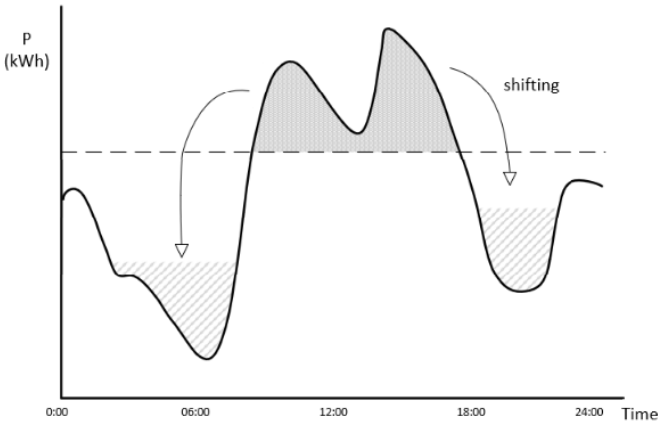


Fig. 2. Peak shifting [7]

Frequency Containment Reserve (FCR) – reserve used to restore the balance between supply and demand. It reacts to short-term frequency imbalances in the grid. The complete activation of this reserve must be available within 30 seconds and cover 15 minutes of delivery period per incident [10][11].

Automatic Frequency Restoration Reserve (aFRR) – used to stabilize the frequency towards the nominal value and to restore the balance between supply and demand when the FCR has failed to do so. It gradually replaces the FCR after 30 seconds. Complete activation of this reserve must be available within 5 minutes [11].

Manual Frequency Restoration Reserve (mFRR) – used when there are longer lasting deviations of nominal frequency in the power grid that cannot be resolved by FCR or aFRR alone. Complete activation of this reserve must be available within 12.5 minutes [12].

C. Types of Flexibility

Explicit demand flexibility – activated based on commands from individual market participants or through an aggregator. The need to use flexibility may be driven by the needs to compensate for deviations in the power grid [13].

Implicit demand flexibility – the flexibility provider directly controls this flexibility based on variable electricity prices or customer tariffs that change with changing transmission and distribution electricity costs [13].

III. AGGREGATION

The word aggregation is a synonym for grouping or merging. Hence the name of the new electricity market participant, aggregator. Aggregator may be a natural or legal person that aggregates the flexibility of individual flexibility providers, whose flexibility would be too small for them to offer in the electricity markets alone. Its purpose is to sell products on the electricity or ancillary services markets. Flexibility aggregator can also be described as an intermediary between the provider and the recipient of the flexibility [3].

Based on types of sources that the aggregator aggregates, there is aggregation of sources that consume or produce electricity.

The ability of renewable energy sources to generate electricity is strongly influenced by the weather and the seasons. This means that the stability of the electric power system itself will be also greatly affected by the growth of renewable energy sources. By aggregating the flexibility of smaller flexibility providers with, e.g. solar, wind or hydroelectric power plants, aggregator gains more power under itself and creates a 'virtual power plant (VPP)'. With different sources of electricity aggregated under the aggregator, these problems can be removed, because VPP would simply increase the generation of, e.g. hydroelectric power plants to cover the deviation in the case of bad weather [14].

IV. ANALYZED DATASET

This section provides a closer look at power consumption in different households in the City of New York on 30th of September 2019. Data used in this part were obtained from Pecan Street Inc. that provides access to consumer energy and water consumption behaviour [15]. These data contain power consumption measured in one second intervals for standard and also some unusual appliances and equipment that can be found in people's households, e.g. washing machine, air conditioning, dishwasher, water heater, refrigerator, water pump, jacuzzi, all light and plug circuits and many more. Power consumption of some of these appliances is shown in Table I.

TABLE I. APPLIANCES AND THEIR POWER CONSUMPTION

Machine type	Power (kW)
Washing Machine	2.01
Refrigerator	0.21
Water Heater	1.04
Dish Washer	1.34
Freezer	0.19

When all the electricity consumed in a household is aggregated, a baseline graph is created. Fig. 3 shows a baseline graph for a household, where 24 appliances were measured.

An example of a household generating electricity is shown in Fig 4. This graph shows how a household equipped with solar panels generates electricity during the day and it also shows an impact of solar production on a baseline graph. During period from 6:00 to 16:00 solar panels overproduce electricity that causes overflows into the grid. This is the time where flexibility would come to use. These overflows can be reduced by appropriate power sizing of the solar panels [16].

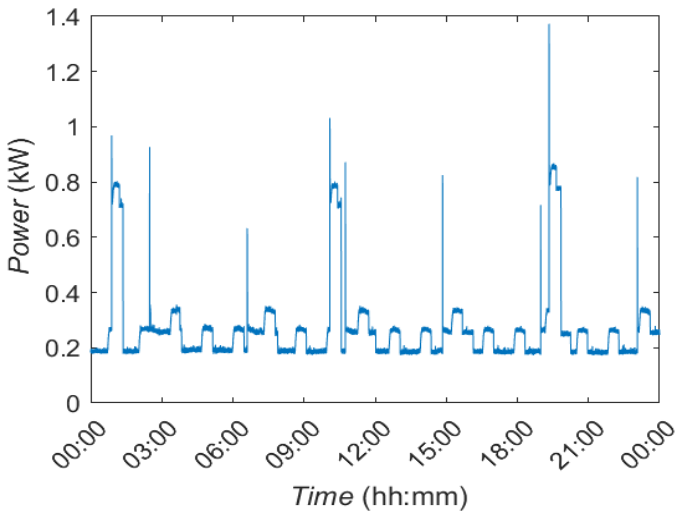


Fig. 3. Baseline graph for household 3

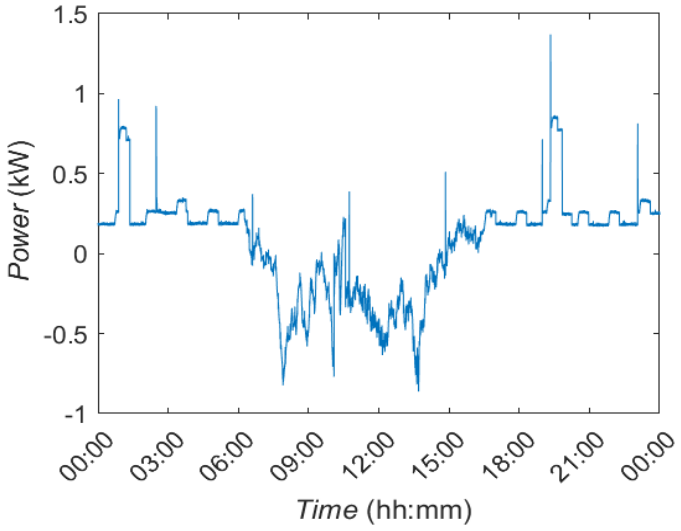


Fig. 4. Impact of solar production on a baseline graph

The shape of the baseline graph is determined not only by the appliances used, but also by the age or type of work of the people in each household. The number of appliances used is also affected by the time of the year and by the type of the day (weekday/weekend).

Fig. 5. shows how the energy consumption changes during the day with prices that would the household pay for each hour. These prices are based on prices from day-ahead market available from Czech electricity and gas market operator (OTE) [17]. Presented household would consume 17,59 kWh and pay €2.64 for this consumption based on 30th of September 2019 prices.

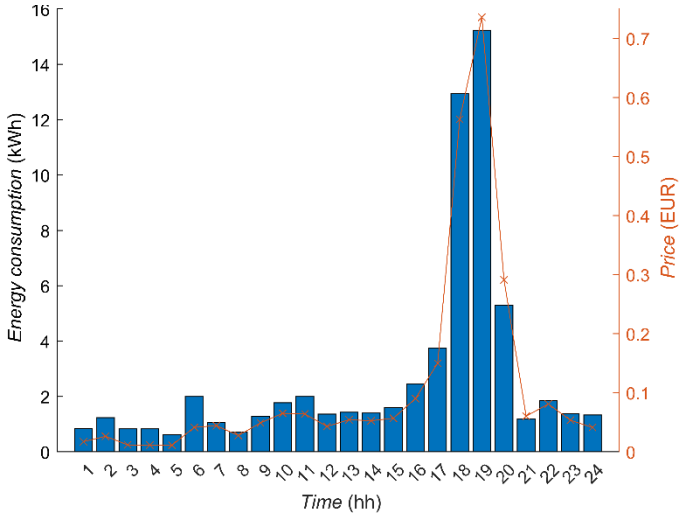


Fig. 5. Energy consumption with prices in each hour

It can be said that the power consumption in Fig. 6 is similar to the load diagram in the Czech Republic in Fig. 8. This is significant because the rhythm of life in the US and the Czech Republic could also be considered similar. Thus, flexibility is therefore possible to examine on different data.

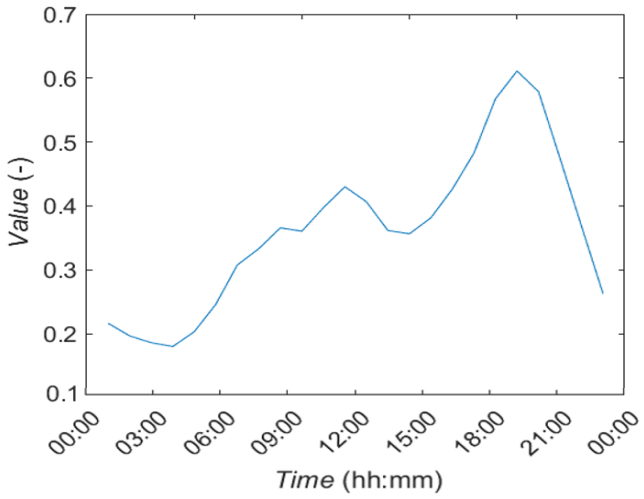


Fig. 6. Load diagram in the Czech Republic [18]

Flexibility should allow the daily load profiles to be balanced so that there are no large power imbalances, which could result in electricity supply constraints or failure to deliver electricity to all customers. With the use of flexibility requirements on the power lines can be lowered. This is why flexibility is important.

Based on the above, if appliances with higher energy consumption e.g. vehicle chargers or washing machines are moved to times when electricity demand is lower, costs can also be reduced. Fig. 7. shows the same household as in Fig. 5., but with use of implicit flexibility i.e. the energy consumption of some appliances is shifted to a later time by the flexibility provider. It is important to note, that with implicit flexibility, the number of appliances that the consumer (flexibility provider) will be willing to make available for flexibility is limited. The

consumer will only be willing to provide appliances, whose time shift in switching on has no or minimal impact on his convenience. In this case, the household would pay €2.33 for the same energy consumption.

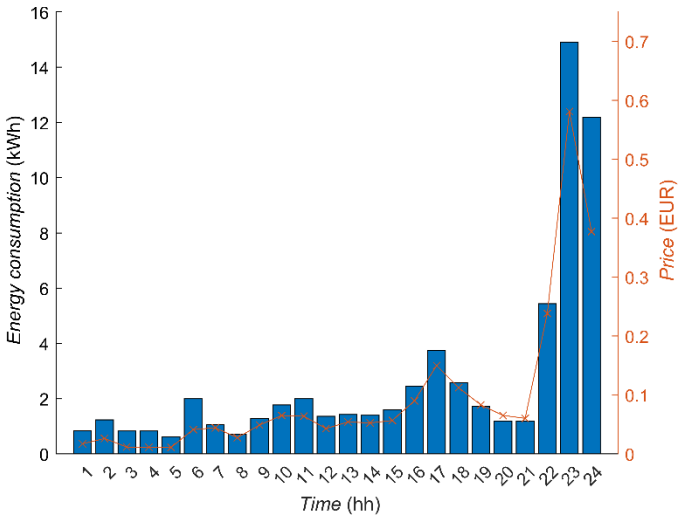


Fig. 7. Load shifting based on the price signal

V. Conclusion

The increasing number of renewable energy sources (RES), decentralised production and the liberalisation of the energy market are among the terms that are increasingly heard in the energy sector. A way to increase the deployment of RES is to provide financial or other benefits to potential flexibility providers. With this flexibility could be a clear tool for the efficient operation of sources and appliances on the grid.

It is also important to recognise that new models of consumer (residents, companies) behaviour will need to be created based on actual real-life behaviour and this data is an excellent tool for this.

The potential of flexibility will be determined not only by the technical design, but also by changes to legislation on national and international level. Legislation will be as important as the technical design, as it will have to define new rights and obligations for aggregators, distribution system operators and transmission system operators in a way that is non-discriminatory to one side or other while not hindering development and expansion.

Another way forward in this work could be to focus more on the flexibility of household appliances and their impact on the daily profile.

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 of the Czech Republic under BUT specific research programme (project No. FEKT-S-23-8403).

References

[1] Ministry of Industry and Trade. “Model of decentralised generation, storage and consumption including electromobility into the management process of the electricity system of the Czech Republic” [online]. Prague, 2018. Available: https://www.mpo.cz/assets/cz/energetika/strategicke-a-koncepcni-dokumenty/narodni-akcni-plan-pro-chytre-site/2020/2/Flexibilita.pdf

[2] F. Panetta. Greener and cheaper: could the transition away from fossil fuels generate a divine coincidence? [online]. Rome, 2022. Available: https://www.ecb.europa.eu/press/key/date/2022/html/ecb.sp221116~c1d5160785.en.html

[3] Deloitte. “The role of the aggregator in the Czech energy sector” [online]. Prague, 2018. Available: https://www.mpo.cz/assets/cz/energetika/strategicke-a-koncepcni-dokumenty/narodni-akcni-plan-pro-chytre-site/2020/1/Role-agregatora-v-ceske-energetice.pdf

[4] Ministry of Industry and Trade. “National action plan for smart grids 2019 – 2030” [online]. Prague, 2019. Available: https://www.mpo.cz/assets/cz/energetika/strategicke-a-koncepcni-dokumenty/narodni-akcni-plan-pro-chytre-site/2019/9/Aktualizace_NAP_SG_2019-2030.pdf

[5] A. Černý. “Think Tank - National action plan for smart grids” [online]. 2021. Available: https://www.mpo.cz/cz/energetika/strategicke-a-koncepcni-dokumenty/narodni-akcni-plan-pro-chytre-site/jednani-think-tanku-nap-sg-dne-11--5--2021--261369/

[6] German Environment Agency. Rebound effects [online]. Dessau-Roßlau, 2019. Available: https://www.umweltbundesamt.de/en/topics/waste-resources/economic-legal-dimensions-of-resource-conservation/rebound-effects

[7] Next Kraftwerke. “What does peak shaving mean?” [online]. Cologne. Available: https://www.next-kraftwerke.com/knowledge/what-is-peak-shaving

[8] Next Kraftwerke. “What is load management?” [online]. Cologne. Available: https://www.next-kraftwerke.com/knowledge/load-management

[9] Next Kraftwerke. “What is frequency containment reserve (FCR)?” [online]. Cologne. Available: https://www.next-kraftwerke.com/knowledge/frequency-containment-reserve-fcr

[10] C. Diaz-Londono, C. A. Correa-Florez, J. Vuelvas, A. Mazza, F. Ruiz, and G. Chicco. “Coordination of specialised energy aggregators for balancing service provision” [online]. Elsevier, 2022 [cit. 2022-03-10]. 23524677. Available: https://doi.org/10.1016/j.segan.2022.100817

[11] Next Kraftwerke. “What is aFRR (automatic frequency restoration reserve) and how does it work?” [online]. Cologne. Available: https://www.next-kraftwerke.com/knowledge/afrr

[12] Next Kraftwerke. “What is mFRR (manual frequency restoration reserve / R3)?” [online]. Cologne. Available: https://www.next-kraftwerke.com/knowledge/mfrr

[13] Ministry of Industry and Trade. “Model of integration of decentralised generation, storage and consumption, including electromobility, into the Czech electricity system management process - interim report for 2018” [online]. Prague, 2018. Available: https://www.mpo.cz/assets/cz/energetika/strategicke-a-koncepcni-dokumenty/narodni-akcni-plan-pro-chytre-site/2020/2/Flexibilita.pdf

[14] Next Kraftwerke. “What is a virtual power plant?” [online]. Cologne. Available: https://www.next-kraftwerke.com/knowledge/what-is-a-virtual-power-plant

[15] Pecan Street Dataport. https://dataport.pecanstreet.org/. March 2023

[16] M. Tomeš, “Prevention of electricity overflows from PV plants to the grid”. Energie21 [online]. 2016. Available: https://energie21.cz/zabrana-pretoku-elektriny-z-fve-do-site/

[17] OTE. Day-Ahead Market CZ Results [online]. Prague. Available: https://www.ote-cr.cz/en/short-term-markets/electricity/day-ahead-market?date=2019-09-30&set_language=en

[18] OTE. “Normalized load diagrams” [online]. Prague, 2020. Availabe: https://www.ote-cr.cz/cs/statistika/typove-diagramy-dodavek-elektriny/normalizovane-tdd?date=2022-09-2

The use of machine learning for non-invasive classification of brain pathologies

Martin Kostial
Faculty of electrical engineering and communication
Brno University of Technology
Brno, Czech Republic
237478@vut.cz

Petr Marcoň
Faculty of electrical engineering and communication
Brno University of Technology
Brno, Czech Republic
marcon@vut.cz

Peter Solár
Department of Neurosurgery
St. Anne's University Hospital
Brno, Czech Republic
peter.solar@fnusa.cz

Abstract–Accurate classification and spatial delineation of brain pathologies is crucial for correct diagnosis and effective treatment. Currently, CT (computed tomography) and MRI (magnetic resonance imaging) are used for this purpose, and when suspected, a biopsy is performed. The aim of this study is to demonstrate the potential of using machine learning to identify areas of a given pathology based on the diffusivity of individual tissues. KNN (k-nearest neighbours) and SVM (support vector machine) models were learned on a dataset containing data from patients with Glioblastoma Multiforme and Abscessus Cerebri, and then their performance was investigated. The obtained results indicate the high accuracy of the models, which only supports their possibilities and potential for future use in automated diagnostic tools that will reduce the use of biopsy and speed up the whole process.

Keywords—brain tumour, pathology, artificial intelligence, machine learning (*key words****)***

I. Introduction

Gliomas are the 19th most common malignant neoplasm in the Czech Republic. In the Czech Republic alone, approximately 3~4 new cases are diagnosed per 100,000 people per year, and 341 patients died from them in 2011-2015 [1]. Glioblastoma multiforme (GBM) is the most common primary brain tumour, is histopathologically highly variable, tends to be locally invasive [2], and distinguishing it from e.g. cerebral abscess can be a complex and lengthy process. In contrast to abscess around GBM, which can be treated by resection, puncture or antibiotic targeting, GBM can currently only be treated palliatively by resection of pathological tissue and reduction of pressure on surrounding structures, followed by radiotherapy. The diagnosis of Glioblastoma multiforme relies on imaging methods (CT - computerized tomography, MRI - magnetic resonance imaging) and subsequent biopsy with histopathological examination of the tissue [3].

The aim of this article is to highlight the possibility of using machine learning models in non-invasive diagnostics in differentiating between different types of pathological tissues (here glioblastoma and cerebral abscess) using only imaging and artificial intelligence (machine learning) methods, which would bring relief to patients, possibilities of reallocation of staff time resources and automation of the process.

The paper is divided into 4 further sections - methodology (information about the method of work and methods of evaluation), results (summary of measured with calculated results), discussion and conclusion.

II. Methodology

For classification we used artificial intelligence (AI) methods, more precisely machine learning (ML). A dataset of 33 patients was utilized, divided into two groups - training and testing. For glioblastoma, the anonymized dataset for this study consisted of 16 patients for training, comprising 9 males with an average age of 58.3 years (min 22, max 78, median 60) and 7 females with an average age of 63 years (min 51, max 72, median 61), and 7 patients for testing, including 4 males with an average age of 64.25 years (min 45, max 80, median 66) and 3 females with an average age of 68.6 years (min 64, max 77, median 65). For abscess, data from eight patients were used for training and 2 for tests, however, no additional pieces of information about these patients were given.

For each patient, after collaboration with the neurosurgeons, 5 cuts were determined, for which the 'line of interest' (see Fig 1) from an ADC (Apparent Diffusion Coefficient) image of the diffusion MRI was acquired. More specifically – firstly, the start and end point are determined, i.e. a line segment is effectively created and projected onto the agreed MRI slice, from where diffusivity values (encoded by the brightness value) are exported to the corresponding location. The resulting waveform is manually divided into healthy tissue, tumour, and oedema sections. Transitions between segments will be referenced as "edges" (gradient 1 to 3 in Fig 1). From the other values of the segments themselves, the mean, median and standard deviation are calculated - thus forming a set of 9 learning parameters – mean, median and standard deviation for oedema, mean, median and standard deviation for tumour, gradient of the ascending and descending oedema edges, and gradient of the ascending tumour edge ("Gradient 1", 2 and 3in Fig 1 respectively), which reflect the histology of the tissue, were then, entered into Machine Learning algorithms. The KNN (k-of nearest neighbours) and SVM (support vector machine) models showed the best results in the initial search.

The recorded training was performed on the KNN (1 neighbour, Euclidean metric, equal distance weight, standardized data, PCA off) [4] and SVM (quadratic kernel function, automatic kernel scale, box constraint level 1, multiclass method One-vs-One, standardized data, PCA off) [4]

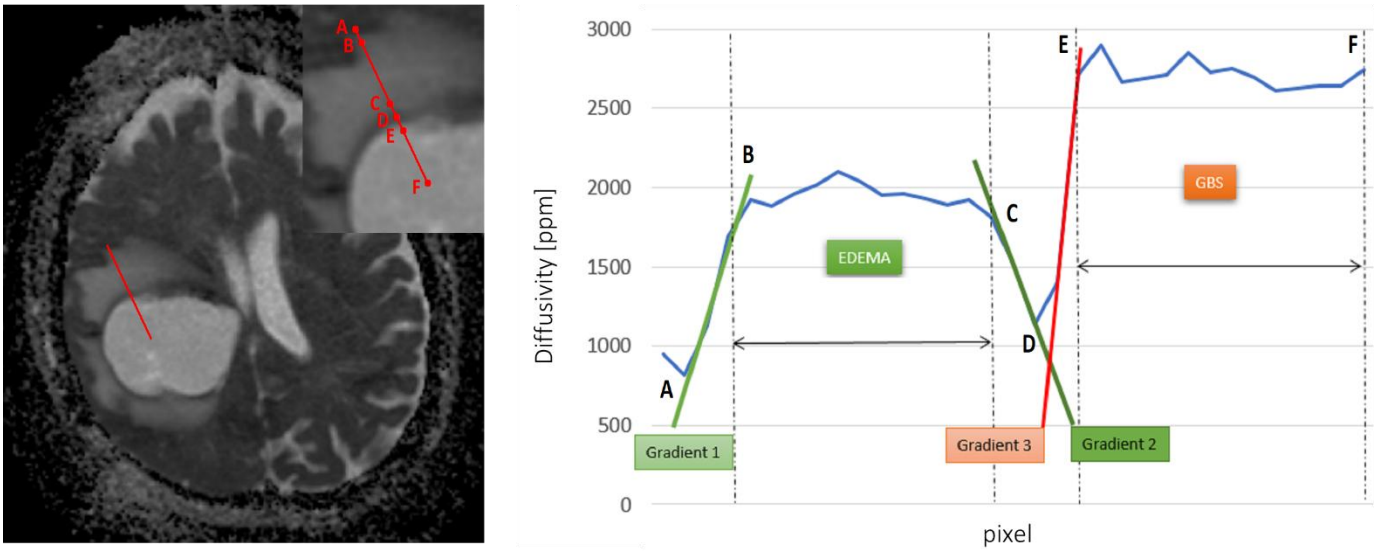


Figure 1 An ADC image with the area of interest enlarged and showing line of interest (left), diffusivity dependence over the course of the line of interest (right)

Cross Validation (sometimes called 'Cross Folds') was enabled, with a d efault value of distribution into 5 parts.

After testing, the values were processed into confusion matrices. From these values, the basic statistical description of the model ('its success rate') – TPR (true positive rate, sensitivity, recall), TNR (true negative rate, specificity), ACC (accuracy), bACC (balanced accuracy) and the predictive values - PV^+ (positive predictive value) and PV^- (negative predictive value) – were calculated. In addition, the ROC curve (receiver operating characteristic curve) is often used to express the ratio between TPR and FPR.

III. Results

The most important results are shown in the tables below - the classification (for both KNN and SVM) on the statistical parameters of tumour, edges and oedema parameters. For completeness (not in tables below) – the tumour + edge combination yielded TPR (82.5/60 – training [%] / test [%]), TNR (90/100), ACC (87.3/86.7), bACC (86.25/80), PV+ (82.5/100), PV- (90/83.3) for KNN, TPR (75/90), TNR (88.6/100), ACC (83.6/96.7), bACC (81.8/95), PV+ (78.9/100), PV- (86.1/95.2) for SVM.

IV. Discussion

From the results above, we can see that edge and tumour parameters are important for classification (between these two tissues) in both ML models, with which classification is successful, whereas oedema is insignificant, suitable probably only for refining the result - but in some cases doing the opposite. The different models have their own drawbacks - for example, for KNNs the computational complexity increases non-linearly with the input size or the increasing error rate when learning with irrelevant variables [5] [6]. The disadvantages of SVMs include e.g. higher computational complexity, propensity to overfitting or sensitivity to noise and outliers in the dataset [6]. The weakness of ADC images is their resolution (256×256 or 512×512 px) – for this case a resolution of at least 1024x1024 px would be more appropriate. In reality, acquisition of such resolution is limited by the storage capacity of the servers (single MRI sequence can take up to several gigabytes of space), longer acquisition time (approx. twice to four times) and computational power [7].

Table I. KNN and SVM results after learning only on tumour data

Tumour				
	KNN		*SVM*	
	training	*testing*	*training*	*testing*
TPR	82.5%	70.0%	82.5%	100.0%
TNR	85.7%	95.0%	91.4%	100.0%
ACC	84.5%	86.7%	88.2%	100.0%
bACC	84.1%	82.5%	87.0%	100.0%
PV+	76.7%	87.5%	84.6%	100.0%
PV-	89.6%	86.4%	90.1%	100.0%

Table II. KNN and SVM results after learning only on oedema data

Oedema				
	KNN		*SVM*	
	training	*testing*	*training*	*testing*
TPR	60.0%	30.0%	42.5%	20.0%
TNR	77.1%	75.0%	84.3%	100.0%
ACC	70.9%	60.0%	69.1%	73.3%
bACC	68.6%	52.5%	63.4%	60.0%
PV+	60.0%	37.5%	60.7%	100.0%
PV-	77.1%	68.2%	72.0%	71.4%

TABLE III. KNN AND SVM RESULTS AFTER LEARNING ONLY ON EDGES DATA

Edges				
	KNN		*SVM*	
	training	*testing*	*training*	*testing*
TPR	62.5%	90.0%	65.0%	60.0%
TNR	82.9%	100.0%	92.9%	90.0%
ACC	96.7%	96.7%	82.7%	80.0%
bACC	72.7%	95.0%	79.0%	75.0%
PV+	67.6%	100.0%	83.9%	75.0%
PV-	79.5%	95.2%	82.3%	81.8%

TABLE IV. KNN AND SVM RESULTS AFTER LEARNING ON ALL AVAILABLE DATA

All parameters				
	KNN		*SVM*	
	training	*testing*	*training*	*testing*
TPR	90.0%	80.0%	87.5%	70.0%
TNR	91.4%	100.0%	90.0%	100.0%
ACC	90.9%	93.3%	89.1%	90.0%
bACC	90.7%	90.0%	88.8%	85.0%
PV+	85.7%	100.0%	83.3%	100.0%
PV-	94.1%	90.9%	92.6%	87.0%

We are aware that a small dataset was used for the study, so the values cannot be taken as sufficiently conclusive, but rather as an outline and demonstration of the possibilities. Also, due to the small dataset available, the evaluation of the tests is not based on the ROC curve as it showed nonsensical results. The other parameters to evaluate are standard in the field, all taking the value $\langle 0,1 \rangle$ (higher is better) and together form a good basis for test assessment.

V. CONCLUSION

We have illustrated the potential of using machine learning to differentiate pathological tissues in the brain, from diffusion MRI images only. The results show that the aforementioned models - KNN and SVM - are robust and provide accurate results even with a limited amount of training data.

More extensive work is needed to confirm our results, especially in terms of the number and variability of patients involved, additional machine learning models, and diversity of pathologies.

We realize that the potential of AI methods (and also machine learning) in healthcare is huge, that they will facilitate the diagnosis (resolution) of tissues and speed up the whole process.

ACKNOWLEDGMENT

The research was carried out under support of the general student development project being executed at Brno University of Technology.

REFERENCES

[1] R. . Lakomý, T. . Kazda a P. . Šlampa, Gliomy: současná diagnostika a léčba, 2. vydání editor, sv. , Praha: Maxdorf, 2018.

[2] *Multiformní glioblastom: přehled nových poznatků o patogenezi, bio-markerech a perspektivách léčby,* editor, sv. 76/109, : , 2013.

[3] Q. T. Ostrom, N. . Patil, G. . Cioffi, K. . Waite, C. . Kruchko a J. S. Barnholtz-Sloan, „CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2013–2017,“ *Neuro-Oncology*, sv. vol. 22, č. Supplement_1, pp. iv1-iv96, 2020.

[4] W.-M. Lee, Python machine learning, 1st ed editor, Indianapolis, IN: John Wiley and Sons, 2019.

[5] A. C. Faul, A concise introduction to machine learning, 1st ed editor, Boca Raton: Chapman & Hall/CRC, 2019.

[6] E. Alpaydin, Introduction to machine learning, 2nd ed editor, Massachusetts: MIT Press, 2010.

[7] C. Westbrook, Handbook of MRI technique, 1st publ editor, Chichester: Wiley Blackwell, 2014.

Pathological tissue segmentation in brain MRI

1st Michael Malík
Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication,
Brno University of Technology, Brno, Czechia
Brno, Czech Republic
email: xmalik23@vutbr.cz

***Abstract*—This paper addresses the problem of image segmentation using a deep learning model. First, the stroke problem is described, followed by a brief description of MRI and also the Swin Factorizer model implemented to predict pathologies in brain MR images. Data preprocessing is performed using registration techniques and automatic segmentation prediction is performed for each patient from the ISLES22 dataset. In the end, the success of pathology segmentation is evaluated using Dice score, Hausdorff distance and F_1 score.**

***Index Terms*—stroke, deep-learning, segmentation, MRI, neural networks**

I. Introduction

This paper discusses the problem of deep learning for segmentation of selected pathological tissues, especially stroke in brain MR images. Its content belongs to the area of medical image processing and analysis. Deep learning is of great importance for the development of contemporary medicine and segmentation methods using artificial intelligence are constantly being improved. It leads to improvement in the diagnosis of brain pathologies, but also to the detection of other pathological tissues in other organs of the human body.

A stroke is an acute condition in which the blood supply to a specific part of the brain is disrupted, the damaged area does not receive adequate oxygen and nutrients, and therefore dies over time. In terms of etiology, stroke is distinguished into ischemic and hemorrhagic forms. The ischemic form of stroke accounts for about 85 percent of cases and is caused by blockage of a blood vessel, most commonly by a thrombus, resulting in insufficient blood in a given section of the brain. On the other hand, the remaining 15 percent of the cases are caused by rupture of a blood vessel, which causes bleeding directly into the brain or meninges, again resulting in insufficient blood supply to that area of the brain and is called hemorrhagic stroke [1]. There is now an increasing number of younger stroke patients for whom a faster diagnosis using automated segmentation methods based on artificial neural networks could speed up and facilitate the work of radiologists and physicians. Last but not least, more patients could be saved [2].

Magnetic resonance imaging (MRI) is used to diagnose stroke and the extent of brain damage after a stroke. It currently represents one of the best imaging modalities used for diagnosis. The advantage for the patient is that, unlike an X-ray machine or a computed tomography scan, the patient does not receive a dose of ionizing radiation. However, MRI also has its own specificities, in particular, its high price and low availability [3].

The gradient coils, radiofrequency coils, and the main magnet are essential parts of every MRI. The main magnet can be of permanent, resistive, superconducting, or electromagnet design. The most commonly used main magnet in practice is the superconducting coil as a closed system. The basic requirement for the main magnet is the homogeneity of the magnetic field. Another element is the radio frequency coils, which are receiving or transmitting. The last element is the gradient coils, which determine the selection of the tomographic plane [3] [4].

Diffusion-weighted imaging (DWI), fluid-attenuated inversion recovery (FLAIR) and apparent diffusion coefficient (ADC) sequences are used in imaging brain pathologies. Namely, DWI is used to display signal changes caused by the diffusion of water molecules in tissues. DWI is applied in the detection of brain pathologies, which may include, for example, stroke. FLAIR is based on an inversion recovery sequence and can suppress cerebrospinal fluid very well. ADC values reflect the degree of diffusion of water molecules in different tissues [4].

The SwinFactorizer model was implemented for stroke prediction and tested on the ISLES dataset22. The evaluation of the success rate was performed using the Dice score, the F_1 score, the Hausdorff distance, and the absolute volume difference. The procedures and methods are described below.

II. ISLES22 dataset

The freely available dataset is taken from the ongoing ISLES22 ischemic stroke lesion segmentation challenge [5]. It contains brain volume data acquired by MRI.

The ongoing challenge is expected to contribute to the identification of reference methods for lesion segmentation of acute and subacute ischemic stroke. The dataset includes a total of 250 patients with ischemic stroke. Each patient is imaged using MR sequences, namely DWI, ADC, and FLAIR sequences. In addition to these three sequences, the dataset also contains ground truth. The parameters for each sequence are as follows: height, width, and depth. For the DWI and ADC images, the height is 112 pixels, the width is 112 pixels, and the depth is 73 slices. In contrast, the FLAIR image has a height of 512 pixels, a width of 512 pixels, and a depth of 29 slices. The data shown are in NIfTI format for medical data.

It should be noted that in the images of the respective dataset, the background and skull are removed for each patient [5].

In Figure 1, you can see an example of display in DWI, FLAIR, and ADC sequences. And also the ground truth (or annotated mask) of a particular patient.

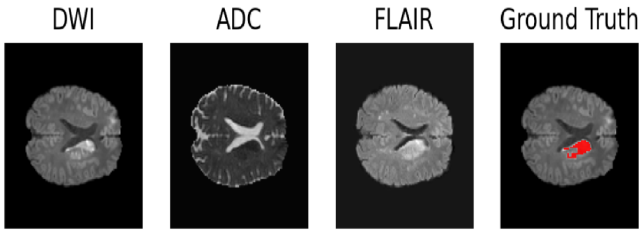


Fig. 1. Example of a section display in the DWI, ADC, and FLAIR sequence of a specific patient section and annotated mask.

III. Swin Factorizer

The Swin Factorizer [6] segmentation algorithm was chosen for its public availability. The author of the proposed network is Pooya Ashtari. In this paper, the trained model has been used to predict lesions in stroke patients. The Swin Factorizer model mentioned above is based on the Factorizer network, which uses low-rank MF factorization matrices that lead to reduced dimensionality and improved representational learning. Swin Factorizer differs from the Factorizer network only in that it uses negative factorization matrices (NFM), as they can reliably compress the data and extract other important parameters. The Factorizer architecture is similar in some ways to the U-Net architecture, also containing encoder and decoder blocks that are switched between within the same resolution, but enriched with Factorizer blocks. It uses a visual transformer, which has the advantage of being able to work with its entire memory. The structure of the visual transformer contains both NFM and multilayer perceptron (MLP) [6] that can perform the resampling operations necessary for the proper functioning of the network.

Figure 2 shows the architecture of the Factorizer network on which Swin Factorizer is based.

Furthermore, the Swin Factorizer is capable of segmenting images faster than traditional segmentation networks. This is because Swin Factorizer can efficiently utilize information from the image, thereby shortening the time required for segmentation. Another advantage of Swin Factorizer is its ability to work with a large amount of data. This allows Swin Factorizer to work with images of varying sizes and quality, which is particularly useful for applications that work with large data files. Thus, Swin Factorizer represents a significant step forward in the field of image segmentation. Its ability to efficiently work with large data files, speed, and accuracy makes Swin Factorizer a strong player in the field of image processing [6].

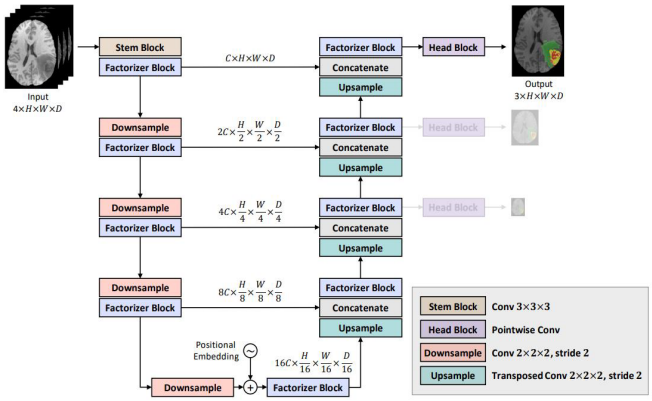


Fig. 2. Factorizer network architecture [6].

IV. Implementation of the Model

In the following Figure 3 you can see the block diagram to obtain the prediction.

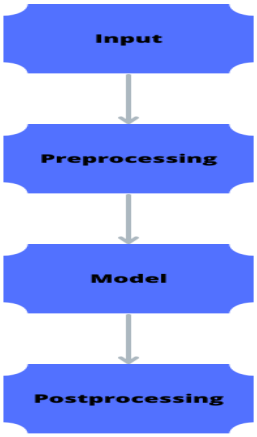


Fig. 3. Block diagram for obtaining a prediction

A. Data preprocessing

Images in ADC, FLAIR and DWI sequences are taken in NIfTI format. For further processing of the data, it is necessary first to register each image for each patient, since the images are not in the same coordinate system. The DWI and ADC images of the patients are aligned, while the FLAIR image in its native form is aligned to a single plane. The annotated mask is aligned in the same way as the ADC and DWI images, but the prediction will also be performed on the FLAIR image because of the greater amount of pathological information, plus the registration of the images allows for subsequent more accurate prediction.

The registration was done automatically for each patient. Elastic registration [7] was used, where the FLAIR image was set as the moving image and the DWI image as the fixed image. As a result, a new registered FLAIR image is merged and created as input to the Swin Factorizer prediction

model. In the elastic registration process, a model or function is usually used to describe the transformations between two images. For example, the model may be a transformation matrix that describes changes in position and angle, or it may be a set of deformations that describe how one image deforms to fit the other image. The model is usually adjusted and optimized using algorithms that minimize registration errors. In this case, the B-spline transform is used since it can deform the image with high accuracy.

The SimpleITK [8] library was used here to manipulate the image data. The library provides methods for loading and storing image files, as well as algorithms for image registration. In Figure 4 you can see the registered FLAIR image.

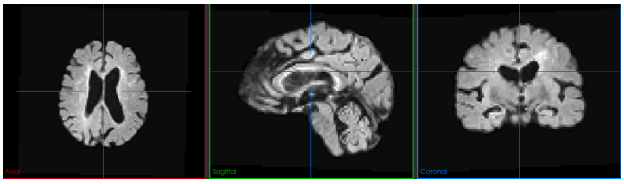


Fig. 4. Example of a registered image of a specific patient in axial, sagittal, and coronal sections.

B. Segmentation of the pathological lesion

An automated script was used to perform segmentation prediction using a trained Swin Factorizer network model by the author Pooya Ashtari and was adopted and modified. An automatic function was created to retrieve the data for each patient and perform the prediction.

The principle of the main code is to retrieve the ADC, DWI, and registered FLAIR images, which are used to predict the segmentation. All 3 images enter the network, but for prediction the main parameter is the registered FLAIR image, ADC and DWI are used to add information and improve prediction. The next part of the code is devoted to the inference and configuration of the segmentation model. The model consists of multiple trained Swin Factorizer networks that are weighted and used to predict the segmentation result. After segmentation, the results are stored in a file as the prediction result. The Swin Factorizer was trained by Pooya Ashtari using cross-validation, where out of 250 patients, 50 were always used as test patients and the remaining 200 were training patients. Thus, a total of five trained Swin Factorizer networks were used as an ensemble for the final prediction.

A total of 182 out of 250 patients were successfully predicted without problems. However, in 68 patients, the affinity matrix of the registered FLAIR image was not as good as the other ADC and DWI sequences. Therefore, an automated script was created to first read the registered FLAIR and ADC, DWI sequences, and then overwrite the affinity matrix values so that they are the same, and no prediction problem occurs. After that, the remaining 68 patients could be predicted. However, in these 68 patients, the prediction does not achieve as good results as in the previous 182 patients out of a total of 250 patients in the ISLES22 dataset, which is due to the change in the geometry of the affine matrices. Therefore, it is possible that some patients were imaged in poorer quality than the remaining ones.

C. Postprocessing

After prediction, postprocessing was performed in the form of an ensemble in which five Swin Factorizer models trained on five-folds were combined. The execution of the ensemble resulted in an overall improvement in prediction for all patients.

This Figure 5 shows a comparison of the prediction with the annotated mask for a specific patient.

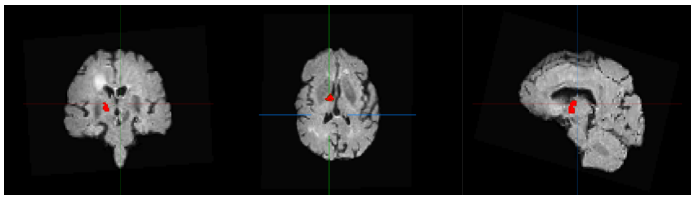


Fig. 5. Comparison of prediction (red) with annotated mask (green outline) in a specific patient in axial, sagittal, and coronal sections. The prediction result is shown here for the registered FLAIR image.

V. Results

Evaluation of prediction success was performed on the entire ISLES22 dataset, i.e. on 250 patients. The Dice score, F_1 score and Hausdorff distance were used to evaluate prediction success. These evaluation parameters were calculated from the obtained prediction, with all 3 modalities (ADC, FLAIR, DWI) involved in performing the prediction

A. Dice score

The Dice score is an evaluation metric commonly used in image segmentation tasks, particularly in medical image analysis. The Dice score ranges from 0 to 1, where a score of 1 indicates perfect overlap between predicted and ground truth segmentation, and a score of 0 indicates no overlap at all [7].

In the following Figure 6 you can see a boxplot evaluating the success rate depending on the number of patients.

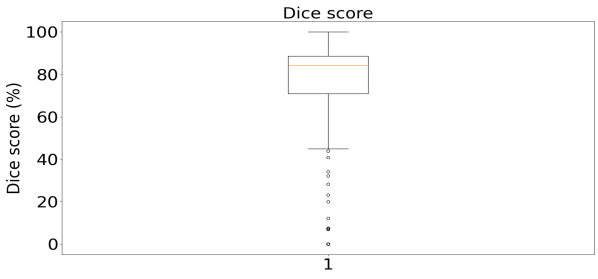


Fig. 6. Dice score boxplot shows the success rate of individual patients, with outliers indicating those 68 patients who experienced a change in affine matrix geometry after elastic registration.

B. Hausdorff distance

The Hausdorff distance is a metric used to measure the similarity between two sets of points in a metric space. In the context of image analysis and computer vision, the Hausdorff distance is often used to compare two segmented regions or contours, where each set of points represents the boundary of a segmented object [7].

In Figure 7 you can see a boxplot that evaluates Hausdorff distances for all patients in the dataset.

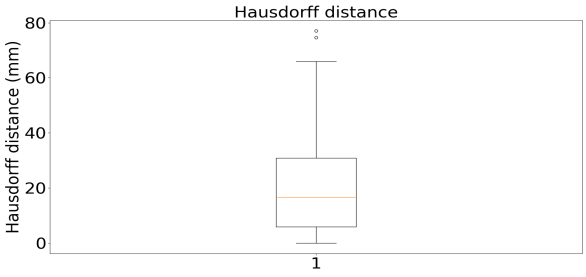


Fig. 7. The Hausdorff distance boxplot shows the Hausdorff distance for each patient, where a small value of hausdorff distance means a better prediction. The outliers are due to the change in the geometry of the affine matrices for 68 patients.

C. F_1 score

The F1 score is a commonly used evaluation metric in binary classification tasks, particularly in machine learning applications. It is a measure of the accuracy of a binary classifier and is calculated as the harmonic mean of the precision and recall of the classifier [7].

Precision measures the proportion of true positives (i.e., correctly classified positive samples) among all positive samples predicted by the classifier. Recall, on the other hand, measures the proportion of true positives among all actual positive samples in the dataset [7].

The F1 score ranges from 0 to 1, where a score of 1 indicates perfect precision and recall, and a score of 0 indicates that the classifier is unable to predict any true positive samples.

In Figure 8 you can see a boxplot that evaluates the F_1 score for all patients in the dataset.

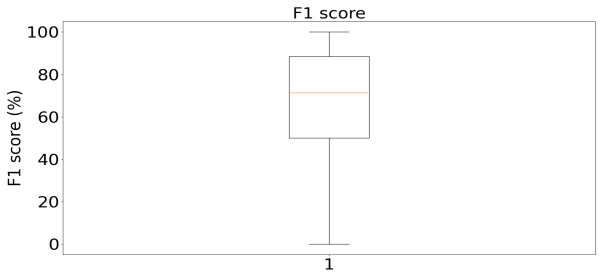


Fig. 8. F_1 score boxplot shows the success rate of individual patients, it is a classification metric that shows similar results to the dice score.

VI. Conclusion

Testing the Swin Factorizer algorithm on the selected dataset, ISLES22, showed that the trained model is able to predict pathological stroke lesions with high success. The prediction success was evaluated using segmentation and classification metrics. Namely, Dice score, Hausdorff distance and F_1 score. The following Table I shows the median for each evaluation metric. For Hausdorff distance, the smaller the number of the mentioned distance, the better the result. The results presented in Table I show that the prediction is accurate despite 68 patients for whom the prediction is almost zero due to distortions in the geometry of the affine matrices. The remaining 182 patients achieve very good results.

TABLE I
Comparison of different metrics

Prediction evaluation metrics	Results
Dice score	84,2 %
Hausdorff distance	16,7 mm
F_1 score	71,4 %

Acknowledgment

I would like to express my gratitude to the author of the Swin Factorizer network for providing the trained models. Additionally, I am deeply grateful to my supervisor, Jiri Chmelik, for his invaluable personal guidance throughout this project.

References

[1] Abdu, Hussen, Fentaw Tadese, Girma Seyoum, and Carol Milligan. 2021. “Comparison of Ischemic and Hemorrhagic Stroke in the Medical Ward of Dessie Referral Hospital, Northeast Ethiopia: A Retrospective Study”. Neurology Research International 2021 (June): 1-9. https://doi.org/10.1155/2021/9996958.

[2] Bártová, Jarmila. 2021. Přehled patologie. Vydání druhé. V Praze: Univerzita Karlova, nakladatelství Karolinum.

[3] Jan, Jiří. 2006. Medical image processing, reconstruction and restoration: concepts and methods. Boca Raton.

[4] Brown, Robert W., Yu-Chung N. Cheng, E. Mark Haacke, Michael R. Thompson, and Ramesh Venkatesan. [2014]. Magnetic resonance imaging: physical principles and sequence design. Second edition. Hoboken, New Jersey: Wiley-Blackwell.

[5] Hernandez Petzsche, Moritz R., Ezequiel de la Rosa, Uta Hanning, Roland Wiest, Waldo Valenzuela, Mauricio Reyes, Maria Meyer, et al. 2022. Scientific Data 9 (1). https://doi.org/10.1038/s41597-022-01875-5.

[6] Ashtari, Pooya, Diana M. Sima, Lieven De Lathauwer, Dominique Sappey-Marinier, Frederik Maes, and Sabine Van Huffel. 2023. “Factorizer: A scalable interpretable approach to context modeling for medical image segmentation”. Medical Image Analysis 84. https://doi.org/10.1016/j.media.2022.102706.

[7] Maier-Hein, Lena, Annika Reinke, Evangelia Christodoulou, and Ben Glocker. “Metrics reloaded: Pitfalls and recommendations for image analysis validation”. 2022. https://doi.org/10.48550/arXiv.2206.01653.

[8] Klein, S., M. Staring, K. Murphy, M.A. Viergever, and J. Pluim. 2010. “Elastix: A Toolbox for Intensity-Based Medical Image Registration”. IEEE Transactions on Medical Imaging 29 (1): 196-205. https://doi.org/10.1109/TMI.2009.2035616.

[9] Beare, Richard, Bradley Lowekamp, and Ziv Yaniv. 2018. “Image Segmentation, Registration and Characterization in R with SimpleITK”. Journal of Statistical Software 86 (8). https://doi.org/10.18637/jss.v086.i08.

Segmentation of spinal cord structures from MRI images using deep learning

Katerina Krejci
Department of Biomedical Engineering
FEEC, Brno University of Technology
Brno, Czech Republic
xkrejc78@vutbr.cz

Jan Valosek
Department of Neurology
Palacky University Olomouc
Olomouc, Czech Republic
0000-0002-7398-4990

Jiri Chmelik
Department of Biomedical Engineering
FEEC, Brno University of Technology
Brno, Czech Republic
0000-0001-9950-6279

***Abstract*—Morphometric measures obtained from magnetic resonance images (MRI) are commonly used to assess the severity of spinal cord compression. However, current automatic methods do not perform sufficiently when applied to T1-weighted MRI images. The purpose of this work is to develop automatic deep learning models for spinal cord and gray matter segmentation from T1-weighted images. The proposed models were trained on a dataset of 66 healthy subjects and compared with state-of-the-art methods. The preliminary results demonstrated a higher Dice coefficient and lower Hausdorff distance compared to the existing methods suggesting promising model performance.**

***Index Terms*—Magnetic Resonance Imaging, Deep Learning, Segmentation, Spinal Cord, Spinal Cord Compression**

I. Introduction

Degenerative spinal cord compression occurs as a result of pathological changes such as intervertebral disc degeneration or osteophyte formation. If not diagnosed on time, initially mild and asymptomatic degenerative spinal cord compression (i.e., non-myelopathic degenerative cervical spinal cord compression, NMDC) may progress over time into severe and irreversible damage to spinal cord tissue (i.e., degenerative cervical myelopathy, DCM) [1], [2]. The morphometric measures obtained from magnetic resonance images (MRI), such as the cross-sectional area, showed the potential to differentiate between these conditions [3]. Since manual annotation of MRI data is time consuming and affected by inter-rater variability, automatic deep-learning (DL) segmentation algorithms were proposed [4], [5]. However, these algorithms were trained mainly on T2-weighted and T2*-weighted (T2*w) images, and therefore they do not generalize well on other modalities like T1-weighted (T1w). The aim of this work is to develop robust DL-based algorithms for the spinal cord and gray matter segmentations from T1-w images.

II. Data and Methods

A. Dataset description

The MRI data were acquired on a 3T MRI scanner (Siemens, Prisma) at the CEITEC research center as part of a research project conducted by the University Hospital Brno. A total of 66 healthy subjects and 147 subjects with spinal cord compression were measured. This work uses T1w and T2*w data from 66 healthy subjects (42 females and 24 males; mean and standard deviation age of 56.2 $\pm$ 9.7 years). Data are organized according to the Brain Imaging Data Structure (BIDS) convention [7].

B. MRI data

Both T1w and T2*w data were acquired covering the same area of the cervical spinal cord (FOV 180 $\times$ 180 mm, 2.5 mm slice thickness, 42 MRI slices, voxel size 0.35 $\times$ 0.35 $\times$ 2.5 mm^3 after interpolation in Fourier space, original voxel size 0.70 $\times$ 0.70 $\times$ 2.5 mm^3).

C. MRI data pre-processing

Pre-processing of the MRI data was performed using publicly available Spinal Cord Toolbox (SCT) [8] and FSL (FMRIB Software Library v6.0) [9] libraries. Since we did not have manual ground truth labels available for T1w data, we decided to leverage labels from T2*w images. First, we registered the T2*w image to the T1w image. To improve the registration accuracy, we used a mask created around the automatically obtained centerline. Then we segmented the spinal cord and gray matter from a co-registered T2*w image using DL-based algorithms available within the SCT (`sct_deepseg_sc` [4] and `sct_deepseg_gm` [5]). All segmentations were visually inspected and corrected if necessary. The pre-processing pipeline is shown in Figure 1.

D. Deep-learning model training

Deep-learning neural network training was performed using the Ivadomed toolbox [10]. A 2D U-Net was implemented separately for spinal cord and gray matter segmentations. Dataset splitting was performed randomly in a ratio of 80 % for training and 20 % for testing. The inputs to the model were: the T1w data, the corresponding annotated ground truth segmentation, and a binary mask to refine the region of interest (obtained during the pre-processing). The used neural network hyper-parameters are listed in Table I for the spinal cord segmentation model and in Table II for the gray matter segmentation model. In both cases, extensive data augmentation was utilized. The augmentation parameters are summarized in Table III for spinal cord segmentation and Table IV for gray matter segmentation. The trained model was then applied to the test data that were not used for training (the rest of 20 % of dataset). The Dice coefficient and Hausdorff distance were used to evaluate the segmentation performance.

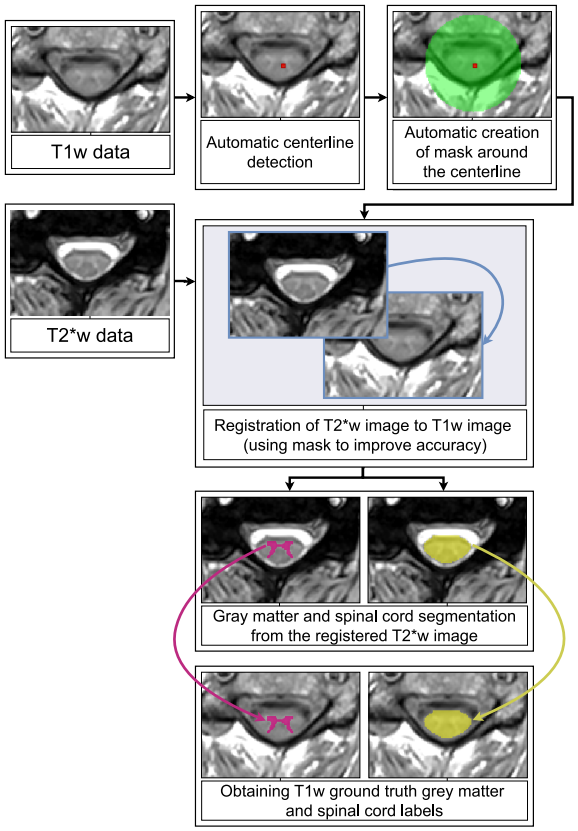


Fig. 1. Pre-processing pipeline scheme showing how the ground truth labels were obtained (red – spinal cord centerline, green – mask for registration refinement, yellow – spinal cord segmentation after registration of T2*w data to T1w data, purple – gray matter segmentation after T2*w data registration to T1w data)

TABLE I
HYPER-PARAMETERS FOR SPINAL CORD SEGMENTATION MODEL TRAINING

Name of hyper-parameter	Hyper-parameter setting
Batch size	32
Loss function	DiceLoss
Number of epochs	100
Learning rate	0.001
Learning rate scheduler	CosineAnnealingLR
BatchNorm momentum	0.1
Dropout rate	0.3
Optimization	ADAM optimizer

TABLE II
HYPER-PARAMETERS FOR TRAINING GRAY MATTER SEGMENTATION MODEL TRAINING

Name of hyper-parameter	Hyper-parameter setting
Batch size	11
Loss function	DiceLoss
Number of epochs	500
Learning rate	0.001
Learning rate scheduler	CosineAnnealingLR
BatchNorm momentum	0.1
Dropout rate	0.4
Optimization	ADAM optimizer

TABLE III
DATA AUGMENTATION PARAMETERS USED DURING THE TRAINING STAGE OF THE SPINAL CORD SEGMENTATION MODEL

Type of augmentation	Augmentation settings
Rotation	$\pm$ 5 °
Shift	$\pm$ 0.03
Scaling	$\pm$ 0.1
Elastic Deformation	α = [28.0, 30.0], σ = [3.5, 4.5], p = 0.1
Resampling	0.75 × 0.75 × 1.00 mm
Cropping	128 × 128 pixels

TABLE IV
DATA AUGMENTATION PARAMETERS USED DURING THE TRAINING STAGE OF THE GRAY MATTER SEGMENTATION MODEL

Type of augmentation	Augmentation settings
Rotation	$\pm$ 4.6 °
Shift	$\pm$ 0.03
Scaling	$\pm$ 0.02
Elastic Deformation	α = [29.0, 31.0], σ = [3.5, 4.5], p = 0.3
Resampling	0.25 × 0.25 × 1.00 mm
Cropping	150 × 150 pixels

III. RESULTS

A. Dataset

Four subjects were excluded from the spinal cord model training due to image artifacts. Similarly, nine subjects were excluded from the gray matter model training due to the low contrast between white and gray matter. Additionally, for the gray matter model training, we cropped the last eight slices for each subject, due to poor gray matter contrast caused by the lung's proximity. Thus, 34 slices were used for the gray matter model training for each subject.

B. Spinal cord segmentation

Figure 2 shows the comparison of spinal cord segmentations obtained by the existing methods (`sct_propseg` [6], `sct_deepseg_sc` [4]) and our model. The statistical evaluation is shown in Figure 3. Generally, the proposed model performs better than `sct_propseg` and `sct_deepseg_sc`, demonstrating a higher mean $\pm$ standard deviation Dice coefficient 0.941 $\pm$ 0.010 compared to `sct_propseg` (0.778 $\pm$ 0.075) and `sct_deepseg_sc` (0.863 $\pm$ 0.053). Similarly, the proposed model showed a lower Hausdorff distance (0.025 mm $\pm$ 0.006 mm) relative to `sct_propseg` (0.519 mm $\pm$ 1.050 mm) and `sct_deepseg_sc` (0.332 mm $\pm$ 0.455 mm). This suggests that our model has been correctly learned and is capable of accurate segmentation on data that it has not seen before.

C. Gray matter segmentation

Figure 4 visualizes the comparison of gray matter segmentation obtained by the existing method (`sct_deepseg_gm` [5]) and our model. The statistical evaluation is shown in Figure 5. Since the `sct_deepseg_gm` method was trained only on T2*w data, we can see that this method fails to segment T1w data. In some subjects, small parts have been segmented

(in blue in Figure 4), but this does not correspond at all to the physiological assumption of a typical gray matter shape. In contrast, our method provided plausible gray matter segmentations for almost all test subjects with the Dice coefficient 0.777 $\pm$ 0.187 and the Hausdorff distance 0.569 mm $\pm$ 1.671 mm (Figure 5).

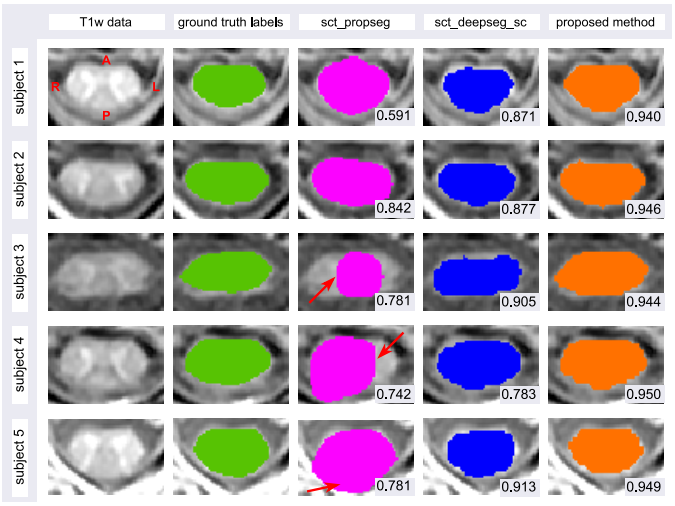


Fig. 2. Comparison of spinal cord segmentation methods (green – manually corrected annotated segmentations, pink – segmentations obtained using `sct_propseg`, blue – segmentations obtained using `sct_deepseg_sc`, orange – segmentations obtained using the proposed method. Numbers in right corners of images correspond to the average Dice score on a given subject relative to ground truth segmentations).

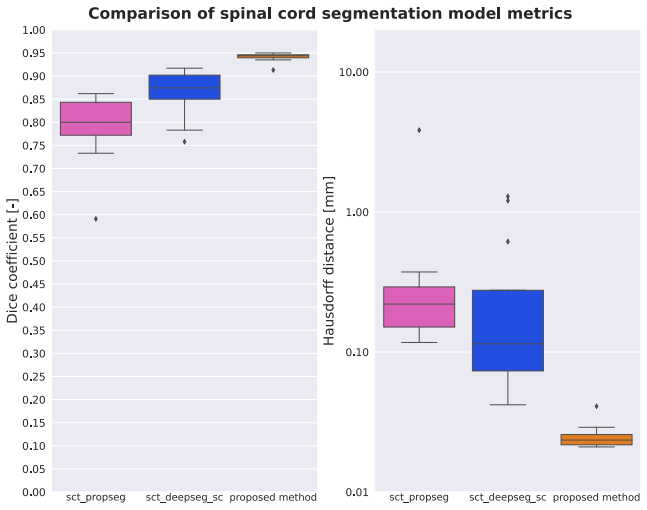


Fig. 3. Box plots showing the Dice coefficient and Hausdorff distance for individual spinal cord segmentation methods (the y-axis is in the graph of Hausdorff distance in logarithmic scale).

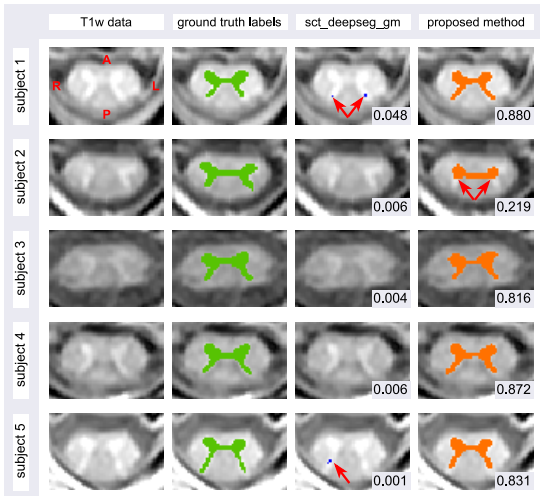


Fig. 4. Comparison of gray matter segmentation methods (green – manually corrected annotated segmentations, blue – segmentations obtained using `sct_deepseg_gm`, orange – segmentations obtained using the proposed method. The numbers in the right corners of the images correspond to the average Dice score on a given subject relative to ground truth segmentations).

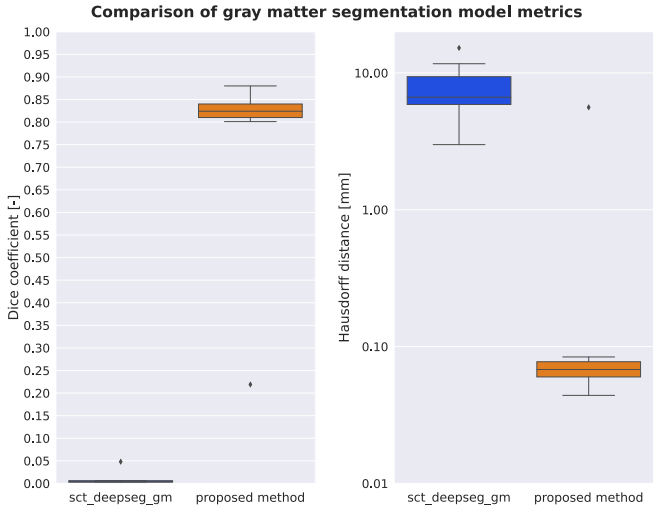


Fig. 5. Box plots showing the Dice coefficient and Hausdorff distance for individual gray matter segmentation methods (the y-axis is in the graph of Hausdorff distance in logarithmic scale).

IV. Discussion and Conclusion

In this work, we propose two deep learning models for automatic spinal cord and gray matter segmentation from T1w images. Since the original dataset did not contain ground truth labels, we leveraged automatically obtained labels from co-registered T2*w images to overcome time-consuming manual annotation.

Figure 2 shows that `sct_propseg` did not properly segment the spinal cord in some subjects. Namely, in subjects 3 and 4 the entire elliptical shape of the spinal cord was not captured correctly; and in subject 5, the cerebrospinal fluid was partially included in spinal cord segmentation (see red arrows

in Figure 2). Moreover, in subject 1, `sct_propseg` failed to segment several slices at all (see the low average Dice coefficient in Figure 2). These inaccuracies may have been caused by the low contrast between the spinal cord and cerebrospinal fluid [6]. Although the `sct_deepseg_sc` method outperformed `sct_propseg`, the resulting spinal cord segmentation was slightly under-segmented relative to the ground truth labels. This might have been because the `sct_deepseg_sc` method was trained mainly on T2w and T2*w data [4]. Since the proposed model was trained exclusively on T1w images, it achieves better performance for all test subjects compared to `sct_propseg` and `sct_deepseg_sc`. However, there is still the drawback that ground truth masks were obtained from only one rater, which may have introduced subjective influence into the training of the proposed method.

Figure 4 shows that the `sct_deepseg_gm` method failed to segment the gray matter on the T1w data completely. The small structures were segmented in some parts (see red arrows in Figure 4), but they did not correspond to the physiological assumption of the gray matter shape (butterfly-like shape). This can be attributed to the fact that the model of the `sct_deepseg_gm` method was trained only on T2*w data and therefore does not generalize well to the T1w data. In contrast, the proposed method provided plausible gray matter segmentations for most testing subjects. Only in subject 2 the proposed method segmented the gray matter inaccurately. This might have been caused by the different shapes of the gray matter of this subject. This points to the limitation that not all physiological gray matter shapes were completely included in the training dataset. A solution would be to include more such subjects with atypical gray matter physiological shapes in the training dataset to achieve better results.

In general, the preliminary results demonstrated promising model performance comparable or better relative to the existing methods available within the SCT toolbox. It is worth noting that the proposed methods were tested on data acquired with the same MRI protocol as the training data. As a result, the comparison with the methods available within the SCT might be biased because these methods were trained on different datasets. However, retraining the SCT methods on the same dataset would be challenging and out of the scope of this work.

Future work will include a comparison of 2D U-Net and 3D UNet architecture and model training with the self-configuring nnU-Net approach [11]. The proposed models will be applied to the dataset of subjects with spinal cord compression to automatically extract morphometric measures and evaluate spinal cord and gray matter atrophy between the groups of non-myelopathic degenerative cervical spinal cord compression and degenerative cervical myelopathy. If the models do not generalize well to the population with spinal cord compression, we will explore an active learning approach.

V. Acknowledgments

This research was supported by Ministry of Health of the Czech Republic, grants nr. NV18-04-00159 and NU22-04-00024.

The Titan Xp GPU used for implementation of the algorithms was donated by the NVIDIA Corporation as a support of academic research.

References

[1] BEDNARIK, Josef, Zdenek KADANKA, Ladislav DUSEK, Milos KERKOVSKY, Stanislav VOHANKA, Oldrich NOVOTNY, Igor URBANEK a Dagmar KRATOCHVILOVA. Presymptomatic spondylotic cervical myelopathy: an updated predictive model. *European Spine Journal* [online]. 2008, **17**, 421-431 [cit. 2023-03-08]. ISSN 0940-6719. Available from: <https://doi.org/10.1007/s00586-008-0585-1>

[2] VALOŠEK, Jan, Petr BEDNAŘÍK, Miloš KEŘKOVSKÝ, Petr HLUŠTÍK, Josef BEDNAŘÍK a Alena SVATKOVA. Quantitative MR Markers in Non-Myelopathic Spinal Cord Compression: A Narrative Review. *Journal of Clinical Medicine* [online]. 2022, **11**(9) [cit. 2023-03-11]. ISSN 2077-0383. Available from: <https://doi.org/10.3390/jcm11092301>

[3] KADANKA, Zdenek, Blanka ADAMOVA, Milos KERKOVSKY, Zdenek KADANKA, Ladislav DUSEK, Barbora JUROVA, Eva VLCKOVA a Josef BEDNARIK. Predictors of symptomatic myelopathy in degenerative cervical spinal cord compression. *Brain and Behavior* [online]. 2017, **7** [cit. 2023-03-08]. ISSN 21623279. Available from: <https://doi.org/10.1002/brb3.797>

[4] GROS, Charley, Benjamin DE LEENER, Atef BADJI, et al. Automatic segmentation of the spinal cord and intramedullary multiple sclerosis lesions with convolutional neural networks. *NeuroImage*. 2019, **184**, 901-915 [cit. 2022-10-16]. ISSN 10538119. Available from: <https://doi.org/10.1016/j.neuroimage.2018.09.081>

[5] PERONE, Christian S., Evan CALABRESE a Julien COHEN-ADAD. Spinal cord gray matter segmentation using deep dilated convolutions. *Scientific Reports*. 2018, **8**(1) [cit. 2022-11-05]. ISSN 2045-2322. Available from: <https://doi.org/10.1038/s41598-018-24304-3>

[6] DE LEENER, Benjamin, Samuel KADOURY a Julien COHEN-ADAD. Robust, accurate and fast automatic segmentation of the spinal cord. *NeuroImage*. 2014, **98**, 528-536 [cit. 2022-10-16]. ISSN 10538119. Available from: <https://doi.org/10.1016/j.neuroimage.2014.04.051>

[7] GORGOLEWSKI, Krzysztof J., Tibor AUER, Vince D. CALHOUN, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data* [online]. 2016, **3** [cit. 2023-03-08]. ISSN 2052-4463. Available from: <https://doi.org/10.1038/sdata.2016.44>

[8] DE LEENER, Benjamin, Simon LÉVY, Sara M. DUPONT, Vladimir S. FONOV, Nikola STIKOV, D. LOUIS COLLINS, Virginie CALLOT a Julien COHEN-ADAD. SCT: Spinal Cord Toolbox, an open-source software for processing spinal cord MRI data. *NeuroImage*. 2017, **145**, 24-43 [cit. 2022-10-19]. ISSN 10538119. Available from : <https://doi.org/10.1016/j.neuroimage.2016.10.009>

[9] JENKINSON, Mark, Christian F. BECKMANN, Timothy E.J. BEHRENS, Mark W. WOOLRICH a Stephen M. SMITH. FSL. *NeuroImage* [online]. 2012, **62**(2), 782-790 [cit. 2022-12-01]. ISSN 10538119. Available from: <https://doi.org/10.1016/j.neuroimage.2011.09.015>

[10] GROS, Charley, Andreanne LEMAY, Olivier VINCENT, Lucas ROUHIER, Marie-Helene BOURGET, Anthime BUCQUET, Joseph COHEN a Julien COHEN-ADAD. Ivadomed: A Medical Imaging Deep Learning Toolbox. *Journal of Open Source Software*. 2021, **6**(58) [cit. 2022-12-02]. ISSN 2475-9066. Available from: <https://doi.org/10.21105/joss.02868>

[11] ISENSEE, Fabian, Paul F. JAEGER, Simon A. A. KOHL, Jens PETERSEN a Klaus H. MAIER-HEIN. NnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods* [online]. 2021, **18**, 203-211 [cit. 2023-03-08]. ISSN 1548-7091. Available from: <https://doi.org/10.1038/s41592-020-01008-z>

The design of NB-IoT connected rain gauge

1st Jiří Jílek
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
xjilek21@vut.cz

2nd Radek Možný
Department of Telecommunications
Brno University of Technology
radek.mozny@vut.cz
Unit of Electrical Engineering
Tampere University
Tampere, Finland
radek.mozny@tuni.fi

3rd Petr Musil
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
musilp@vut.cz

***Abstract*—Estimation of rainfall volume by point-to-point microwave links is a promising way to increase the network of rain gauges. Nevertheless, this method of measurement is still in the research phase and has several problems. In this paper, we focus on one of the problems which is an insufficient number of existing rain gauges on the path of the link by designing a portable NB-IoT-connected rain gauge. To extend its capabilities, the device will also be equipped with thermometers and an illuminance sensor. Low power consumption is desirable as the device should be capable of measuring data even without the reach of the power grid. We also include a test measurement to verify the functionality of the device and its suitability for the intended use case.**

***Index Terms*—environmental sensors, ESP32, IoT, LPWAN, microwave link, NB-IoT, rain gauge, rainfall**

I. Introduction

Commercial microwave point-to-point links (CMLs) using frequency bands between 5 to 40 GHz, which are utilized for both last-mile and backhaul links, rely on line-of-sight between base stations. In case of any obstacle along the route, mainly caused by the weather, the connection quality degrades drastically. The current installation of microwave links can respond to these fluctuations, for example, by increasing the transmitted power, however with some delays. These delays could be prevented by implementing predictive control algorithms. These algorithms require realistic measurements of the environment close to the microwave link. There is a way of estimating rainfall amounts directly via CML, which looks promising. These estimations could help with microwave link signal stability and increase knowledge about rainfall distribution across areas. One of the current challenges in rainfall measurement using CML, whose simplified diagram is shown in the "Fig. 1", is the inability to verify output information. Meteorological rain gauge stations are usually located only in places that are meteorologically interesting and serve mainly to create weather forecasts. Their data is also often not publicly available. There is therefore a need for low-cost rain gauge stations that will combine low power consumption for the possibility of running measurements even without an available power grid and relatively accurate measurement of the rainfall volume. The purpose of this paper is to design such a device [1].

In this work, we first investigate the possibilities of automated measurement of the rainfall volume, then we proceed with selection of suitable technologies for the transmission of measured data using low-power wide-area networks (LPWAN). Based on the collected theoretical assumptions, we design a low-energy device that would measure the volume of precipitation with sufficient accuracy and use the NB-IoT network to transmit the measured data. The final revision of the device (fitted with Quectel BC660K-GL module and Espessif ESP32 microcontroller [MCU]) should be deployed along selected commercial microwave links and support ongoing research conducted at the Brno University of Technology.

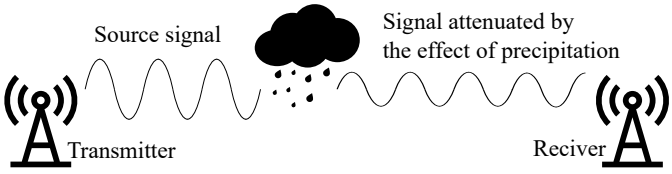


Fig. 1. Simplified principle of CML measurement.

II. Connected rain gauge basis

Rainfall volume measurements are very important in various areas. For example, in agriculture, it can be used to determine or predict a suitable period for sowing or harvesting crops. In hydrometeorological agencies data are used to validate weather prediction models or as input data in research. One of the methods for rainfall measurement is the use of manual (human-operated) or automated rain gauges [2].

There are several ways to automate rainfall volume measurement. The cheapest way is to use a tipping bucket rain gauge. This type of rain gauge is utilizing a rain-collecting funnel to direct rain to the tipping bucket assembly. When one part of this assembly is filled with a precisely defined amount of rainfall the tipping bucket is overturned. This event is recorded, and water is poured out from this part while rainfall is directed to the second part of the assembly. If the second part is filled with again precisely defined amount of rainfall the whole process is repeated [3]. The internal arrangement of the tipping bucket assembly is depicted in "Fig. 2". Tipping bucket

rain gauge can introduce some errors in the measurement but its simplicity, price and low energy consumption makes it very popular in most applications. To measure rainfall amounts more precisely we can use a weighting rain gauge, which is utilizing strain sensors to measure the weight of the water over time. Thanks to very small increments of weight we can measure current rainfall intensity as well. This type of measurement is usually used in permanent professional installations where it is possible to provide a stable flow of power to continuous measurement [3].

All ways of rainfall measurement mentioned above have one common disadvantage – they are measuring rainfall in the local area only. To measure rainfall in a larger area there is a method that is using weather radar to detect rainfall using reflected electromagnetic signals from water in any state of matter. With this method, we can distinguish all precipitation types but it's costly to build and operate. One of the newer ideas of rainfall measurement is using already build and commonly used point-to-point microwave links. These rainfall measurements are based on the fact that rainfall between two base stations measurably changes the signal properties of this link. The estimation of the precipitation amount can be obtained thanks to the equation for calculating rainfall attenuation according to ITU-R P.838 [4]. Rainfall intensity R in mm/h can be calculated as:

$$R = \sqrt[a]{\frac{A}{k \cdot d}} \tag{1}$$

where A represents attenuation caused by precipitation in dB, d is the distance between two base stations of the microwave link in m, and the coefficients a and k are dependent on the frequency and polarization of the microwave link [1].

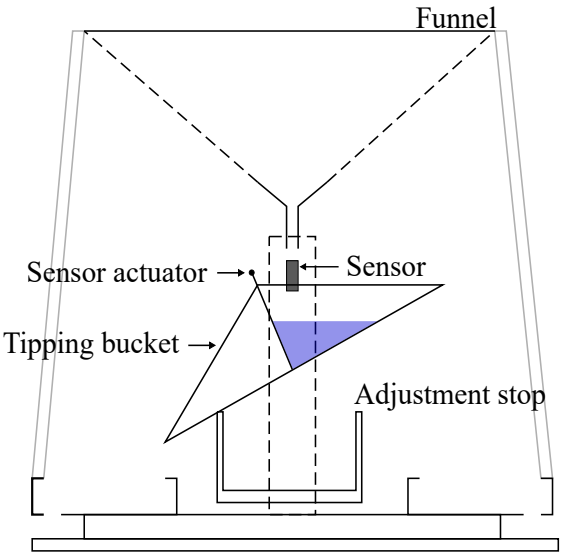


Fig. 2. Sketch of the tipping bucket rain gauge.

A. Utilizing LPWANs for remote data transmission

Using privately owned network infrastructure close to our device is usually not possible thus we need a way to send a small amount of measured data using as little power as possible even in remote locations. LPWANs are a great way how to achieve this goal. In this LPWAN field, there are three main candidates to choose from — long-range wide-area network (LoRaWAN), Sigfox, and narrow-band internet of things (NB-IoT).

LoRaWAN and Sigfox are technologies based around industrial, scientific, and medical (ISM) radio bands which is unlicensed. Specified conditions for broadcasting apply in unlicensed bands such as duty cycle in the range of 1%-10%. The community, ISM band, and an open-source protocol stack of LoRaWAN opened a way to no-fee communication using LPWAN technology but with some caveats. The biggest of them is coverage of these community-driven networks. Next in line, Sigfox is a company and operator offering the LPWA network of the same name. The disadvantage of this technology is its price and a maximum of 12 B uplink data and 8 B of downlink data with a very limited number of transmissions per day. Both LoRaWAN and Sigfox are not suitable for our use case mainly thanks to the 1% duty cycle. If there is a need for real-time measurement it would be impossible to comply with this limit.

Last elaborated technology, NB-IoT is the telecommunication industry answer on growing LPWAN market introduced in third generation partnership project (3GPP) release 13 with official name LTE cat NB1. NB-IoT is an LPWAN technology for machine-to-machine (M2M) communication over licensed bands without the need for new infrastructure. Thanks to good coverage (Vodafone is promising 100% coverage of the Czech Republic), very low power consumption (up to 10+ years of battery life) and no transmission limits it's a perfect fit for our use case [5].

III. Practical implementation of Rain Gauge Prototype

To measure rainfall volume precisely, with low power consumption, and to keep the price down we chose to design a device to which a tipping bucket rain gauge can be connected. If the right model of the tipping bucket rain gauge is selected, we can achieve a 0.1 mm resolution with an accuracy of ± 0.2 mm. The indication of the overturn is indicated using a reed switch inside of the tipping bucket assembly.

To gather more information about the environment around the microwave link base station we added two DS18B20 thermometers, which are located inside a designed 3D-printed active radiation shield. To accommodate the power consumption of the fan inside this active radiation shield 4.5 W photovoltaic panel was added. To measure the light illuminance of surroundings, the sensor of illuminance VEML7700 was added as well.

To utilize power from the photovoltaic panel efficiently, not only to power the fan inside an active radiation shield we have incorporated power management and protection circuits to charge and protect single-cell Li-ion or Li-Pol battery (we used a Li-ion battery with a stated capacity of 3000 mAh) and even power the whole device when the battery is charged. This

has been ensured by implementing the power path which cuts the battery out of the system and supplies power directly from the photovoltaic panel to the low drop-out regulator to supply steady 3.3 V.

Whole device is controlled by ESP32-WROOM-32E system on the chip (SoC) by Espressif systems. This SoC is responsible for getting data from the sensors, converting them to a suitable format, and including them in the appropriate AT command which is sent through a universal asynchronous receiver/transmitter (UART) line to the Quectel BC660K-GL NB-IoT module. Quectel BC660K-GL is one of the newest NB-IoT modules from the manufacturer. It supports both LTE cat NB1 and LTE cat NB2 standards and a wide variety of communication protocols such as UDP/TCP stack, MQTT, and Lightweight M2M (LwM2M) even in their secured versions (TLS on UDP/TCP stack, MQTTS) [6]. SoC module integrates deep sleep mode and the NB-IoT module regularly enters into power saving mode (PSM) for decreasing their power consumption while they are not in use.

Prototype of the device is fitted with Quectel's development kit of the NB-IoT module BC660K-GL-TE-B for quick and easy testing and development of the needed software based on the Arduino framework. For plug-and-play sensor connections, three RJ11 ports are available. One of them is used only for rain gauge connection, second one integrates the OneWire bus for connecting DS18B20 thermometers and power supply lines to power sensors and a fan inside an active radiation shield. The third connector is used for connecting I2C-enabled sensors such as the VEML7700 light illuminance sensor. Ready to test prototype is shown in the "Fig. 3".

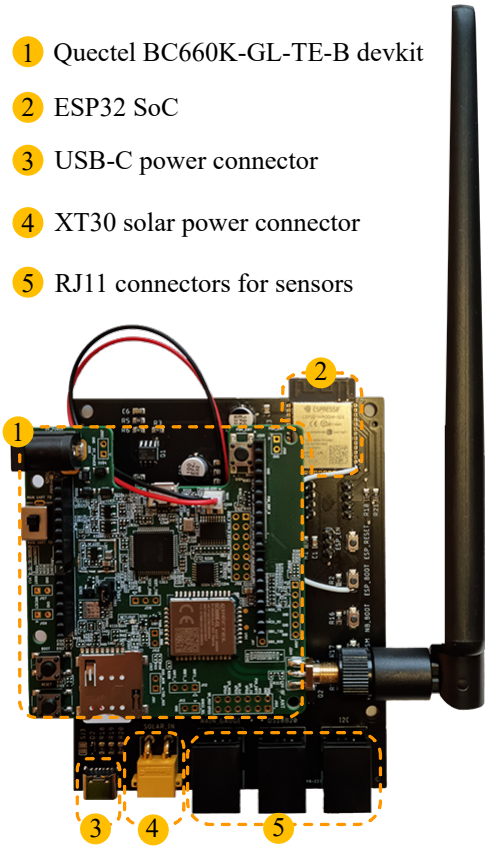


Fig. 3. Designed rain gauge prototype with a description of main device parts.

A. *Software implementation of rain gauge prototype*

Software which is running on ESP32 SoC is written in C++ using Arduino framework. To control the NB-IoT module, we wrote a library Quectel_BC660[1] that implements AT control commands into simple logic. The algorithm implemented in the software is shown in the "Fig. 4". The running of the program is governed by two conditions - "Did the device wake up due to the rain gauge tipping over?" and "Is it time to send the data?". If the device wakes up due to the tip of the rain gauge, the device stores this event in memory and compares the wake-up time with the time the data was sent. If the wake-up time is less than the next data send time, the device will adjust the sleep time to send data around the specified time. When it is time to send data, the device wakes up all the sensors, receives the data measured by them, and sends the data to the server using the NB-IoT module. Based on the set interval between sending data, the time of the next data sending is stored in the device's memory and the device goes into a deep sleep. All the data that the device stores for the needs of the program (the number of times the rain gauge is flipped and the time of the next data transmission) are stored in the random access memory (RAM) located in the real-time clock (RTC) part of the ESP32 SoC, this part remains powered even during deep sleep and thus this memory will not be reset. The RTC RAM is reset only when the device is restarted. If real-time rainfall data is needed, the data is sent every time the rain gauge flips over.

The data received by the server is processed and then stored in a time series database for later use. We chose NodeRED for data processing thanks to its simplicity and ease of use, and InfluxDB as our time series database, because there is already one instance, that is already deployed to collect data from CMLs.

IV. Device functionality verification results

To verify the full functionality of the developed prototype, we conducted a test measurement. Due to a cloudy but rainless day, we used a gardening hose with an attachment that could spray water in mist or drops to simulate both light and heavy rains. Results of measurement were sent every 5 minutes and data included the 5-minute sum of rain gauge overturns, one temperature, and one light illuminance measurement. "Fig. 5" shows two rainfall events, the first one which starts around 12:15 UTC represents simulated light rain, and the second one which starts at 15:30 UTC represents simulated heavy rain. Rainfall amount is not consistent due to wind which changed the amount of water drops hitting the rain gauge. "Fig. 6" shows the decrease in illumination during the afternoon and

[1]Library on Github: https://github.com/R4sp1/Quectel-BC660

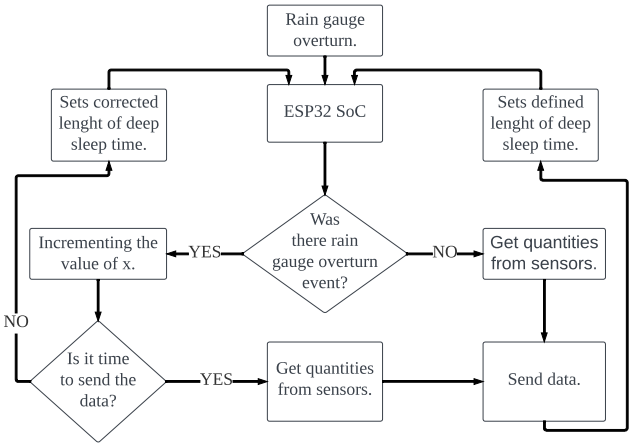


Fig. 4. Algorithm depicting developed prototype's SW implementation.

"Fig. 7" shows temperature which was calculated as average of two DS18B20 sensors located 2 m above the ground.

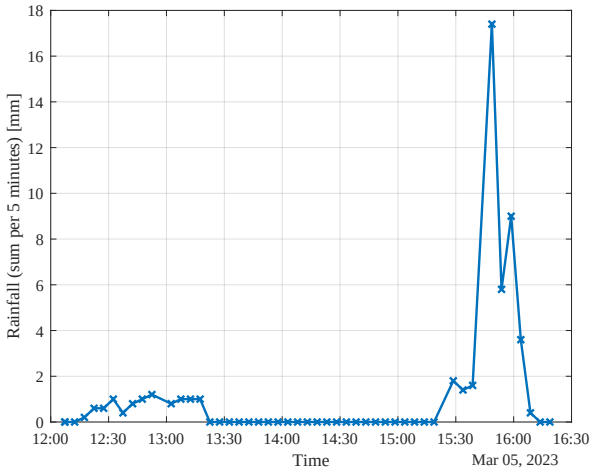


Fig. 5. Measurement of rainfall volume.

V. Conclusion

In this paper, we explored the possibilities of automated measurement of the volume of precipitation and presented the design of a low-energy device for collecting measured data. We used a tipping bucket rain gauge that combines sufficient measurement accuracy with minimal energy consumption and is thus suitable for use in devices that will be located outside the reach of the power grid. For the transmission of measured data, we used the Quectel BC660K-GL NB-IoT module, which supports LTE cat. NB1 and LTE cat. NB2. We chose the Espressif ESP32 SoC to control the device. In the end, we verified the functionality of the device using a simple test. The final version of the device will be deployed in the field to help with ongoing research on precipitation volume measurements using commercial microwave links.

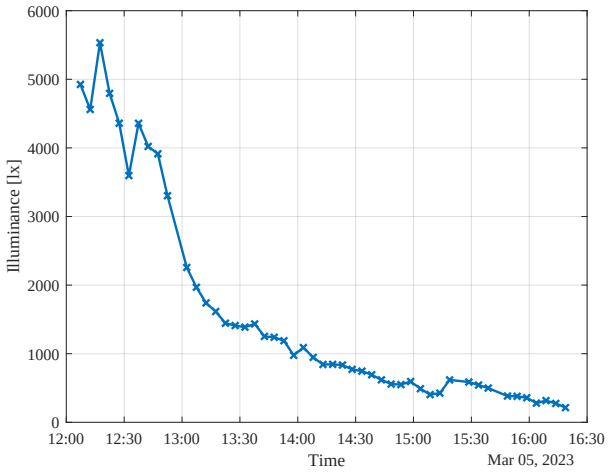


Fig. 6. Measurement of illuminance from VEML7700 illuminance sensor.

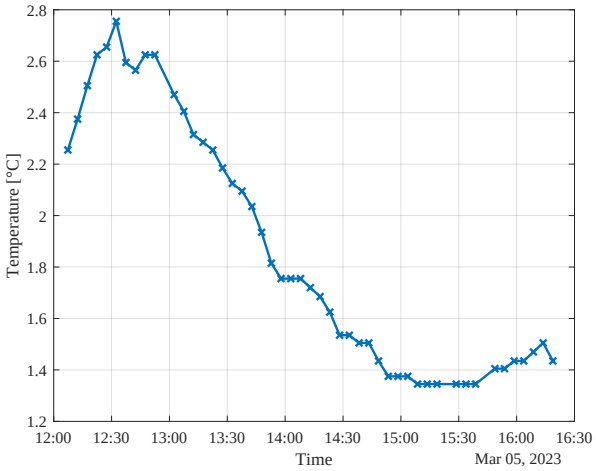


Fig. 7. Average of temperature measurement from two DS18B20 sensors.

References

[1] BUBNIAK, Milan, Petr MUSIL, Petr MLYNEK, Lukas BENESL a Martin RUSZ, "Design and Evaluation of the Application for a Measurement of Rainfall via Microwave Links." In: *2022 14th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)* [online]. IEEE, 2022, 2022-10-11, s. 8-13 [ref. 2022-12-02]. ISBN 979-8-3503-9866-3. Available from: doi:10.1109/ICUMT57764.2022.9943528.

[2] Tim R. Oke, "Instruments and Observing Methods." In: *Report No. 81*, World Meteorological Organization, 2006 [ref. 2023-03-04].

[3] B. Sevruk and S. Klemm, "Instruments and Observing Methods." In: *Report No. 39*, World Meteorological Organization, 1989 [ref. 2023-03-04].

[4] ITU-R P.838-3: "Specific attenuation model for rain for use in prediction methods." Geneva: International Telecommunication Union, 2005 [ref. 2023-03-04].

[5] R. Mozny, P. Masek, M. Stusek, K. Zeman, A. Ometov and J. Hosek, "On the Performance of Narrow-band Internet of Things (NB-IoT) for Delay-tolerant Services." In: *2019 42nd International Conference on Telecommunications and Signal Processing (TSP)*, Budapest, Hungary, 2019, pp. 637-642. Available from: doi:10.1109/TSP.2019.8768871.

[6] Quectel: "Quectel BC660K-GL: Compact NB-IoT Module with Ultra-low Power Consumption," [online]. Shanghai: Quectel Wireless Solutions, c2022 [ref. 2022-12-03]. Available from: https://www.quectel.com/wp-content/uploads/2021/03/Quectel_BC660K-GL_NB-IoT_Specification_V1.5.pdf

Dictionary Search Algorithm for Job Ads Profiling

1st Alex Wittner
Department of Telecommunication
Brno University of Technology
Brno, Czech Republic
230914@vut.cz

2nd Marek Sikora
Department of Telecommunication
Brno University of Technology
Brno, Czech Republic
marek.sikora@vut.cz

3rd Sara Ricci
Department of Telecommunication
Brno University of Technology
Brno, Czech Republic
ricci@vut.cz

***Abstract*—This article describes a novel way for developing an assisting module for analyzing cybersecurity skills in job advertisements. This module will be included in the Cybersecurity Job Ads Analyzer web application, which will be used to examine cybersecurity skill requirements in job opportunities. Our idea speeds up the development of the test database utilized in the analysis by eliminating the need for manual data entry. To that end, various dictionary search algorithms have been examined, and a comparison of four dictionary construction methods is offered.**

***Index Terms*—Dictionary search algorithm, cybersecurity skills analysis, ENISA framework.**

I. Introduction

Cybersecurity has become more prominent as a result of the Information Technology (IT) industry's rapid development and expansion. The lack of a uniform taxonomy of cybersecurity skills and abilities, however, brought a lack of acknowledgement of the skills needed in cybersecurity, and consequently, a lack of effective coordination of training activities. The recently published European Union Agency for Cybersecurity (ENISA) framework [1] describing 12 cybersecurity profiles gives a common ground for skills classification. In fact, the profiles are described through key skills and key knowledge which can be mapped to work roles.

Dictionary search algorithms [2] can help in the identification of cybersecurity skills in specific work roles by seeking specific words in the job advertisement description. Each work role described in an advertisement can be mapped to an ENISA profile that can be seen as a collection of skills. With the identification of which skills are suggested in work roles, it is possible to compare the job market needs and what is taught in education.

A. Contribution

In this article, we compare several dictionary search algorithms, select the one that gives the best results for our use case, and propose an extension that increases its accuracy. We then deploy the selected algorithm to enhance the functionality of a free web-based tool called the Cybersecurity Job Analyzer [3] developed in the Cybersecurity Skills Alliance – A New Vision for Europe (REWIRE) project [4]. The app allows cybersecurity skills need analysis, but currently requires manual insertion of new job advertisements where the user decides on the occurrence of each skill. The deployment of our algorithm allows to simplify this entry by automatic identification of present skills. Therefore, a user will only need to insert the Uniform Resource Locator (URL) of the job advert he would like to upload and fill in the basic information, and the rest of the information will be automatically filled by our module.

The rest of this article is organized as follows. Section II describes in more detail the ENISA framework, REWIRE Skills Groups, and the most appropriate dictionary search algorithms. Section III describes our implementation, i.e. the keyword dictionaries creation and the algorithm that is used within the application. Section IV presents the comparison of dictionary search algorithms and the experimental results with statistical analysis of 110 job advertisements. The final section contains our conclusions.

II. Preliminary

In this section, we discuss the theoretical background that is crucial for the understanding of our implementation. First, we review the ENISA framework describing the cybersecurity profiles. Second, the mapping of the ENISA framework and REWIRE Skills Groups, i.e., skills, is shown. This map is strictly necessary to understand how a dictionary is created. At last, the compared dictionary search algorithms are sketched.

A. ENISA Framework

The European Cyber Security Framework (ECSF) [1] is designed to provide individuals, employers and training providers in the European Union (EU) member states information and a common understanding of what is relevant for cybersecurity profiles. This is possible due to their identified titles, missions, tasks, skills, knowledge. 12 cybersecurity profiles are identified: Chief Information Security Officer (CISO), Cyber Incident Responder, Cyber Legal, Policy & Compliance Officer, Cyber Threat Intelligence Specialist, Cybersecurity Architect, Cybersecurity Auditor, Cybersecurity Educator, Cybersecurity Implementer, Cybersecurity Researcher, Cybersecurity Risk Manager, Digital Forensics Investigator, and Penetration Tester.

B. REWIRE Skills Groups

In [3], a total of 31 skills were selected with the help of detailed analyses and comparisons. These skills were created from the ENISA framework [1] by grouping the ENISA key skills and key knowledge into 31 Skills Groups. We refer to REWIRE Deliverable R3.4.1 [5] for more details.

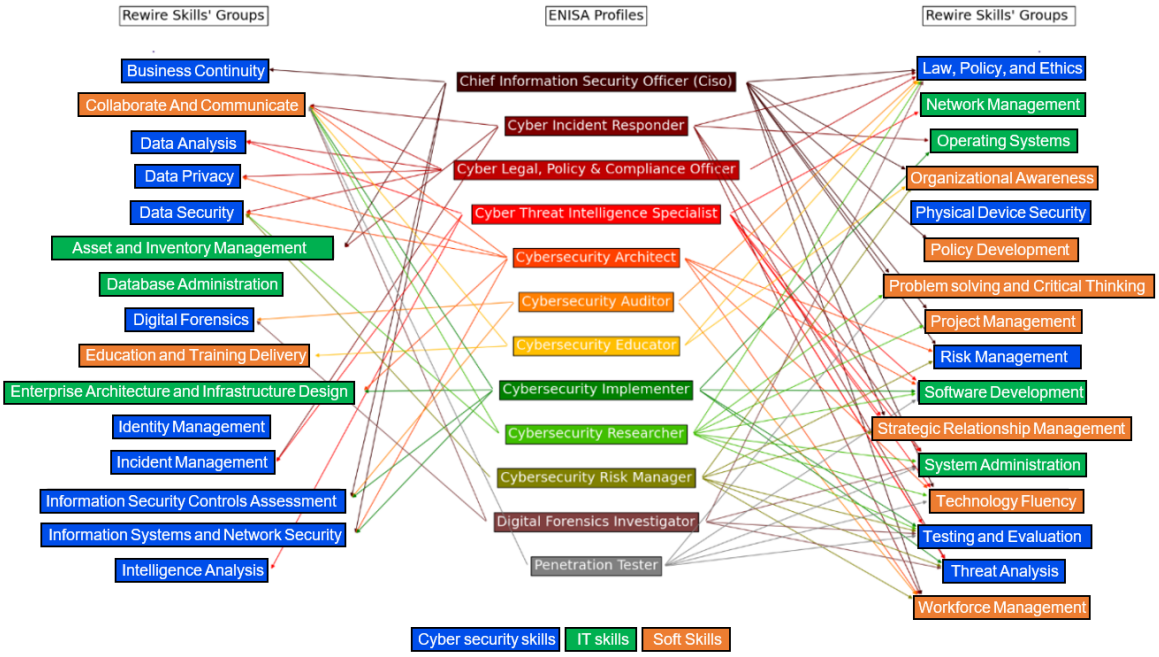


Figure 1: All 31 REWIRE Skills Groups and their connection with ENISA and the categories to which they belong

These skills can be split into 3 categories: 1) Cybersecurity Skills, 2) IT Skills, and 3) Soft Skills. IT Skills are more basic and core information technology expertise, whereas Cybersecurity Skills are those directly related to cybersecurity work roles. Non-technical knowledge is referred to as Soft Skills. Since they are vital for the successful performance of cybersecurity work duties, IT and Soft Skills should also be regarded as essential. It is important to keep in mind that cybersecurity is a multidisciplinary topic and calls for expertise outside of computer science. Figure 1 shows the 31 REWIRE Skills Groups and it is possible to see the interrelationships of the groups and which ENISA profiles.

C. Aho-Corasick Algorithm

Aho-Corasick (AC) algorithm [2] is a multiple-pattern searching algorithm based on pre-processing of patterns. Specifically, during the looping, all the patterns (finite set) for which the step is searched are worked on at the same time. The time complexity is in the worst case θ(N+M+Z), where 'N' is the length of the text, 'M' is the length of keywords, and 'Z' is the number of matches. Given a dictionary of words, the searching algorithm can be divided into 4 steps:

1) **Create the trie** – a data structure similar to the binary tree with different features.
2) **Create the suffix (failure) links** – links that are used by the matcher when a character that cannot follow a trie edge occurs.
3) **Create output links** – each node contains information whether it is final or not.
4) **Implement a matcher** – use a sequential trie traversal and output all matches while using the output and suffix links.

D. Commentz-Walter Algorithm

Commentz-Walter algorithm [6] consists of two very efficient algorithms, the aforementioned Aho-Corasick algorithm and the Boyer-Moore algorithm [7]. Commentz-Walter also works on the principle of trie and multiple pattern searches, supplemented by indexing. However, unlike Aho-Corasick, the trie is reversed. The matches in this algorithm are made using the principle on which the Boyer-Moore algorithm works, which is end-to-end traversal, hence the reverse trie. Commentz-Walter consists of two phases:

1) **Creating a reverse trie** – The procedure is the same as a trie, but the stamping works from "z" to "a" instead of from "a" to "z".
2) **Matching phase** – matching that works on the same principle as Aho-Corasick with a shifting technique derived from the Boyer-Moore algorithm.

III. Implementation

This section describes the methodology applied to create the individual dictionaries and the reasons of your choices. Moreover, our proposed extension of the dictionary search algorithms is presented and their usability in the web application is also shown.

A. Proposed Methods

A dictionary had to be created first, which was then used for the dictionary search algorithm. The accuracy of the search depends on the generated dictionary. 3 different dictionaries, where the words describing each REWIRE Skills Group were created, and then compared their performance. The elements of each dictionary were selected in 4 ways:

- **Expert input** – a preliminary dictionary was proposed by REWIRE partners with a few example words in every group.
- **The ENISA framework** – words are selected from ENISA key skills and knowledge descriptions. For instance, the ENISA Cybersecurity Risk Manager profile contains ”maturity models”, ”mitigate risks”, ”risk management tool”, and ”risk sharing options” that specifically identify the profile and can be added to the dictionary.
- **The National Initiative for Cybersecurity Education (NICE) Competencies framework** [9] – note that REWIRE Skills Groups are mapped to the NICE Competencies framework. As with the first method, the words were selected based on two parameters: uniqueness and expected occurrence in the competencies description. For example, for group Data Security we found the words ”confidentiality”, ”integrity”, and ”availability” in the related competencies definition.
- **Expert knowledge** – some words were chosen according to the author's experience with the given skill, for example in the Database administration the ”PostgreSQL” and ”phpmyadmin” words were added since they are key technologies used in this field. Other words were added based on reading existing job ads. And finally synonyms were used to already existing added words.

The dictionaries consist of:

- **One-word Dictionary** – consists of a single word only. For example, words like Device, Physical and Remote.
- **Two-word Dictionary** – contains objects consisting of two-words. This dictionary allows a more strict description of each Skill group. For example words: cluster analysis, anomaly detection, data mining.
- **Weighted Dictionary** – two sub-dictionaries were created: one with a specific word that unequivocally identifies the group, and one with one-word related to the specific group. For example words: Windows, Linux in group Operating systems with 100% weight and from same category words: system, systems, operating with not 100% weight.

Note that the Weighted Dictionary was generated after comparing the first two dictionaries. In fact, the testing showed that the One-word Dictionary permits recognizing more skills than the Two-word one. However, the One-word Dictionary found many groups that should not have been found. Therefore, the Weighted Dictionary 1^a (Table I) was created. The two sub-dictionaries were then used in a word search program with the following logic: if a word from the 100% dictionary appears in the text, then the group to which the word belongs is set to 1. If a word from the NOT 100% dictionary appears in the text, then a value of the $1/$number of words in that category is added to the auxiliary variable of each group. Then the auxiliary variables are examined and if their value is higher than 0.5 (50 %) the group number is set to 1.

This method of using two dictionaries was subsequently improved and that's why Weighted Dictionary 2^b (Table I) was created since if a certain group contained a lot of words, the chances of 50 % of the words appearing were very small. Therefore, the following logic was applied: if 3 words from a given group occur in the text, then the group to which the words belong is set to 1 if there are less than 5 words in the group, we work with 50 % accuracy.

B. Identifying Skills in Cybersecurity Job Ads

The previous steps of creating a dictionary and a comparison algorithm will then be used in the implementation of the existing web application. The web application provides the possibility of inserting new job ads and manual skills analysis is needed. After the final implementation of dictionaries and algorithms, the insertion of new job ads will be much easier. The authorized user will only need to insert the Hypertext Markup Language (HTML) link and confirm either the presence or not of a REWIRE Skills Groups. Figure 2 depicts a sketch of the analysis helper of the web application. As can be seen, the user will be given a preset value for the given group, according to the criteria that have been set. The user is able to view the words that have been found in the given group and at the same time they are highlighted, using the arrows it is possible to navigate to the next part of the text where others found words can be.

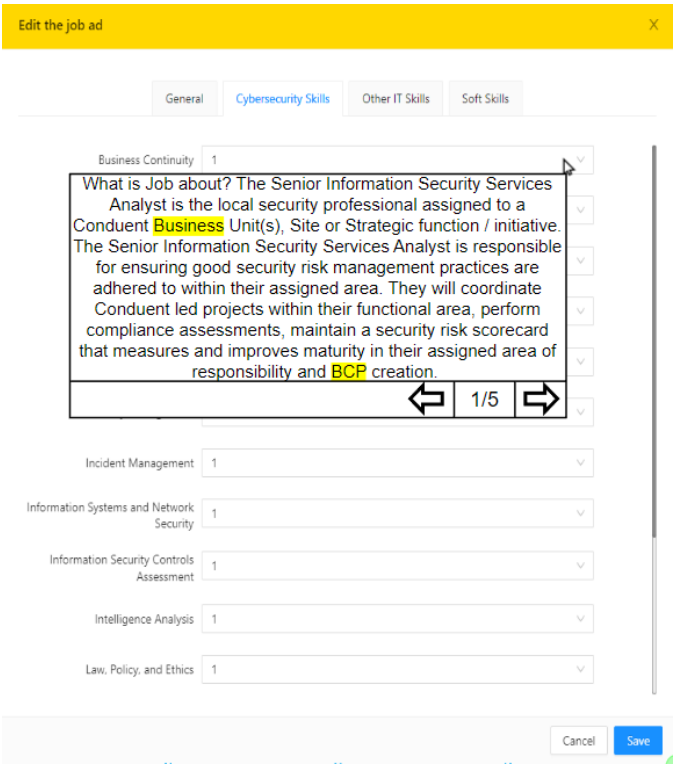


Figure 2: Sample of web application

IV. Experimental Results

This section describes the results of our benchmarking. It compares which dictionary algorithm was the most suitable for this work. It also shows which of the dictionaries used has

the best results in terms of accuracy and why accuracy cannot be evaluated only by positive matches.

A. Comparison of Dictionary Search Algorithms

Since it is necessary to search and find matches as quickly as possible (in order to work with the web application), algorithms were chosen for their speed and reliability. These algorithms belong to the group of multipattern searching algorithms. As the name suggests, these algorithms can search for several keywords (patterns) in a text at the same time.

Based on the study [8], Commentz-Walter is faster in cases where the length of the patterns is longer. The span of a long pattern cannot be determined directly, but patterns that contain more than 20 characters can be considered a kind of breakpoint. The range of 20-120 characters per pattern is given as the span of a long pattern. There is no guarantee that once the length of 20 characters is reached, Commentz-Walter is automatically better. Another important point is the number of patterns, as the Commentz-Walter algorithm becomes less efficient when working with more than 13 patterns. In our search, there are patterns with an average length of 8.2 characters and the current number of patterns is 515. According to the theory, the Aho-Corasick algorithm should be faster.

B. Comparison of Proposed Methods

In this section, we compare 4 generated dictionaries on their accuracy. At the test database, we ran the analysis on 110 job ads that were also analyzed manually. The results are shown in Table I. In the table, two kinds of accuracy have been considered: 1) the accuracy for binary classification:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN},$$

where ACC stays for accuracy, TP for true positives, TN for true negatives, FP for false positives, and FN for false negatives, and 2) since the application involves‘ to the user to confirm the group, the TP has more relevance than the other category, therefore, we consider this value alone.

When we check the ACC value, the Weighted Dictionary 2^b produces the higher matches, whereas, if we consider only the TP matches, One-word Dictionary performs considerably better than the others. Therefore, in case the dictionary search is the final output of the analysis, the Weighted Dictionary 2^b method needs to be deployed. In our case, where a user checks will confirm the selection, we are interested in the higher TP matches. In fact, a user can easily discard the FP values.

Name of Dictionary	TP (%)	ACC (%)
Two-word Dictionary	40.9	58.6
One-Word Dictionary	71.7	59.6
Weighted Dictionary 1^a	53.87	63.17
Weighted Dictionary 2^b	58.69	63.37

Table I: Dictionaries results
[a] match with half of the words in the group
[b] match with 3 words in the group

V. Conclusion

In this article, we presented a methodology for creating a helping module for the analysis of skills in job advertisements. In order to select the most suitable algorithm to identify skills, we ran a comparison of existing dictionary search algorithms. Accordingly to the presented comparison, the best pattern-searching algorithm for our purposes is the Aho-Corasick multipattern searching algorithm since this work is completely dominated by a larger number of short keywords, which are more advantageous for this algorithm than for the Commentz-Walter algorithm.

After selecting Aho-Corasick multipattern searching algorithm, 4 dictionaries were created where also weights were assigned to the word in each group. Note that our helping module allows the pre-filling of the skills that are found in the job advertisement, and it is up to the user to check the validity of the finding. The experimental results pointed out that the most convenient dictionary is the One-Word Dictionary. Possible future work is to improve the accuracy of the protocol by working on the weight and the dictionary and exploring new techniques.

Acknowledgment

The following funding source is gratefully acknowledged: the ERASMUS+ Sector Skills Alliance (SSA) programme of the European Union (grant 621701-EPP-1-2020-1-LT-EPPKA2-SSA-B 'REWIRE').

References

[1] The European Union Agency for Cybersecurity. (2022, Sep. 19). *European Cybersecurity Skills Framework Role Profiles* [online]. Available at: https://www.enisa.europa.eu/publications/european-cybersecurity-skills-framework-role-profiles.

[2] GeeksforGeeks. (2022, Jun. 14). *Aho-Corasick Algorithm for Pattern Searching* [online]. Available at: https://www.geeksforgeeks.org/aho-corasick-algorithm-pattern-searching/.

[3] Ricci S, Sikora M, Parker S, Lendak I, Danidou Y, Chatzopoulou A, Badonnel R, and Alksnys D. "Job Adverts Analyzer for Cybersecurity Skills Needs Evaluation," in *Proceedings of the 17th International Conference on Availability, Reliability and Security*, 2022 Aug 23, pp. 1-10.

[4] REWIRE - Cybersecurity Skills Alliance. (2022). *REWIRE Project Results* [online]. Available at: https://REWIREproject.eu/results/.

[5] REWIRE - Cybersecurity Skills Alliance: "R3.4.1 Mapping the framework to existing courses and schemes (Chapter 5.4)" [online], 2022, Available at: https://REWIREproject.eu/wp-content/uploads/2022/11/REWIRE_R3.4.1_Deliverable-v7-Final.pdf.

[6] Wikipedia contributors. (2023, Jan. 30). *Commentz-Walter algorithm* [online]. Available at: https://en.wikipedia.org/wiki/Commentz-Walter_algorithm.

[7] GeeksforGeeks. (2022, Nov. 9). *Boyer Moore Algorithm for Pattern Searching* [online]. Available at: https://www.geeksforgeeks.org/boyer-moore-algorithm-for-pattern-searching/.

[8] Dewasurendra S, Vidanagamachchi S. (2018). "Average time complexity analysis of Commentz-Walter algorithm". *Journal of the National Science Foundation of Sri Lanka* [online]. Sri Lanka, Available at: http://doi.org/10.4038/jnsfsr.v46i4.8630.

[9] National Initiative for Cybersecurity Education (NICE) — NIST. (2020). *The Workforce Framework for Cybersecurity (NICE Framework)* [online]. Available at: https://www.nist.gov/itl/applied-cybersecurity/nice/nice-framework-resource-center/workforce-framework-cybersecurity-nice#currentversion.

Real-Time Temperature Measurements in the Experimental Nuclear Reactor: A Data Analysis and Chart Creation

1st David Carboch
Brno University of Technology, FEEC
Dept. of Telecommunications
Technicka 12, 616 00 Brno, Czech Republic
xcarbo02@vutbr.cz

2nd Petr Dejdar
Brno University of Technology, FEEC
Dept. of Telecommunications
Technicka 12, 616 00 Brno, Czech Republic
Petr.Dejdar@vut.cz

***Abstract*—The paper deals with real-time temperature measurements in the experimental nuclear reactor VR-1. For temperature sensing, a commercial measuring instrument fibrisTerre fTB 2505 based on BOFDA (Brillouin Optical Frequency Domain Analyzer) is used, which measures Brillouin scattering. Brillouin scattering is temperature-dependent. A change in temperature is reflected by a frequency shift of the Brillouin scattering, which the measuring instrument continuously records and calculates the temperature change from the frequency shift. To protect the sensing optical fiber, a measurement probe was created, which is made of a plastic tube through which the fiber passes, with an arc formed at the end to protect the fiber from breakage. The data from the instrument are uploaded to the InfluxDB time-series database using a Python script and then displayed in real-time charts using the Grafana environment, which creates the charts based on SQL, a standard language for database creation and manipulation.**

***Index Terms*—BOFDA, Brillouin scattering, fiber optic sensor, Grafana, InfluxDB, nuclear, Python, temperature measurement**

I. Introduction

The application of fiber optic sensor systems has begun to thrive with the lower cost of optical components due to their widespread use in the telecommunications environment. They have certainly gained their popularity due to the resistant properties of optical fibers to electromagnetic interference and high temperatures. Therefore, they can be used in hostile environments where conventional electronic sensors can no longer be used or in places where conventional sensors do not fit [1], [2].

Fiber optic sensors can be divided into three groups according to their spatial arrangement, point sensors, semidistributed sensors, and distributed sensors. Point sensors sense the measured quantity at only one location. If multiple point sensors are inserted in series, it is a semi-distributed sensor. Distributed sensors measure the sensed quantity along the entire length of the sensor fiber. Furthermore, fiber optic sensors can be divided according to the way the optical signal is modulated and also according to the measurement method into direct and reflectometric [1]–[3].

The environments of nuclear power plants are very dangerous. The consequences of nuclear accidents are usually catastrophic. To prevent accidents from occurring, it is important to monitor all crisis areas of the plant, for example high reactor temperatures, leakage of fluids from reactor piping or pool levels in a water nuclear reactor. Humidity or radioactive radiation can also be sensed using optical fiber. They can be used, for example, where electronic sensors cannot fit. Many commercially available electronic sensors do not have sufficient characteristics to withstand long-term operation under nuclear power plant conditions. However, fiber optic sensors have greater resistance to hostile conditions in a nuclear power plant [4]–[7].

Distributed temperature and strain sensors - DSTS are also used in other industries. DSTS systems are applied, for example, in the oil and gas industry. Distributed sensors can obtain valuable information about the condition of an oil well's pipeline along its entire length and can monitor pipeline leaks that occur. Furthermore, DSTS systems are applied in geotechnical engineering. They are used for the risk assessment of infrastructure facilities during the excavation and operation of geotechnical projects such as tunnels. The sensors allow to optimize the construction process by creating a balance between increasing safety and reducing costs. DST systems are also used to monitor stresses in roof structures. Distributed sensors monitor the behavior of the structure to control safety [8]–[10].

II. Experimental setup

The DSTS system uses Brillouin scattering, which is frequency spaced at 11 GHz from the carrier signal. Brillouin scattering is caused by the interaction of light with propagating acoustic phonons in the fiber. The Stokes components are shifted to lower frequencies and the anti-Stokes components to higher frequencies. The DSTS system is sensitive to temperature and strain. Using proper techniques to differentiate between temperature and strain, both quantities can be measured simultaneously in a dynamic environment. If only one selected value is to be measured using the DSTS, the other value must be compensated for. For example, by using special fiber optic cables designed for this purpose. An optical fiber

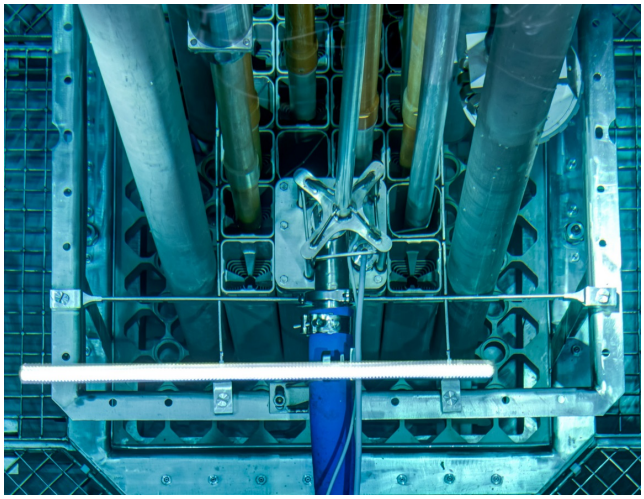


Fig. 1. VR-1 nuclear reactor core [11].

DSTS system can be achieved using a BOTDA - Brillouin Optical Time Domain Analyzer or BOFDA - Brillouin Optical Frequency Domain Analyzer. Both methods are based on the principle of Brillouin scattering [12].

Brillouin scattering can be described as the scattering of light in a diffraction grating moving at the speed of sound. In general, Brillouin scattering can take place in all directions. However, only the forward and reverse directions are allowed for transmission in a fiber. The frequency of the acoustic wave represents the frequency difference of the forward optical signal and the light scattering, which is also called the Brillouin frequency shift and is given by 1, where n is the effective refractive index of the fiber, V_{A} is the velocity of the acoustic wave, and λ_{P} is the wavelength of the forward optical signal. The effective index n and the acoustic wave velocity V_{A} are both temperature and strain dependent [13].

$$\nu_{\mathrm{B}} = \frac{2nV_{\mathrm{A}}}{\lambda_{\mathrm{P}}} \tag{1}$$

A. Calibration of temperature measurements

To calibrate the instrument, it was necessary to create a test sensor optical fiber. A 20 m long optical telecommunication fiber G.657 was used. The connectors were spliced to the optical fiber on both sides using a fiber optic splicer. To splice the optical fiber, it is necessary to remove approximately 20 mm of primary fiber protection from the end of the fiber using a pair of crimping pliers. The bare fiber is then cleaned using a special lint-free cloth soaked in isopropyl alcohol. The clean fiber is broken using a very precise breaker to about 10 mm. For a good splice is it important to break fiber as perpendicular as possible. The clean broken fiber and the equally treated fiber from the connector are inserted into the optical fiber splicer. The splicer aligns the fibers exactly opposite each other and splices them together using an electric arc. After splicing the fiber, a splice protection is put on the splice, which is then baked in a dedicated furnace on the splicer.

The calibration of the temperature measurement was carried out using a temperature chamber from Vötsch model VC^3 with temperature deviation of ±0,1 K to ±0,5 K. Approximately 1 m of test sensor optical fiber was inserted into the temperature chamber. As part of the calibration, several measurements were made with a temperature difference of 5°C each time. Individual measurements are shown by Figure 2.

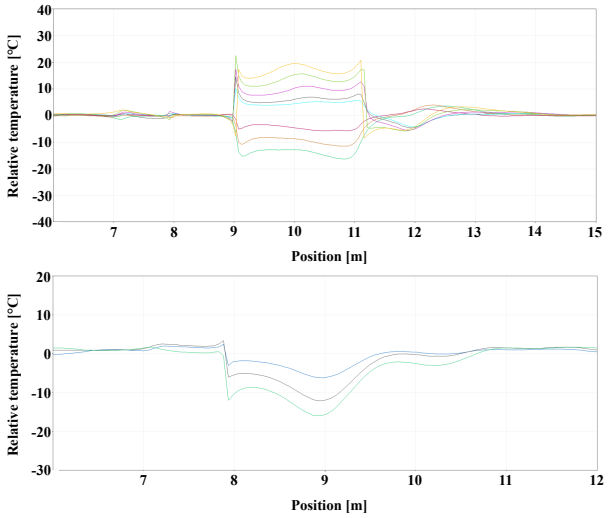


Fig. 2. Relative temperature calibration measurement charts.

B. Temperature measurement in VR-1

The commercial Brillouin scattering measuring instrument fibrisTerre fTB 2505 is connected to an Ethernet network in the hall where the experimental reactor VR1 is located for remote control of the instrument. The connection diagram of the measuring instrument can be seen in Figure 3.

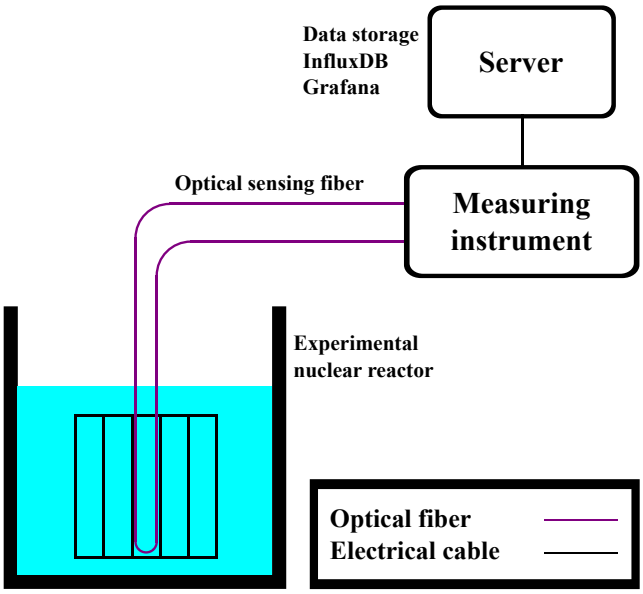


Fig. 3. Connection diagram of the measuring instrument.

Fiber optic connectors were spliced to the sensing fiber using a fiber optic splicer. Then the optic fiber was pulled through the fuel rod and placed with the rod into the reactor core. The connectors were cleaned with a cleaning pen and connected to the measuring instrument. Immediately after the initial measurement of the fiber length and attenuation value, it was found that the sensing fiber was not in good condition. A high attenuation value indicated that the fiber was very acutely bent, pinched, or even broken somewhere. This statement was confirmed by Brillouin scattering measurements. As can be seen in Figure 4, at about 6.5 m of the fiber length, a vertical line from yellow to yellow orange marks the bending point of the problem. Because of this situation, the relative temperature measurement was not possible. Until the entire fuel rod assembly in the reactor core was replaced, only the optic fiber attenuation could be measured with the instrument. It is shown in Figure 5.

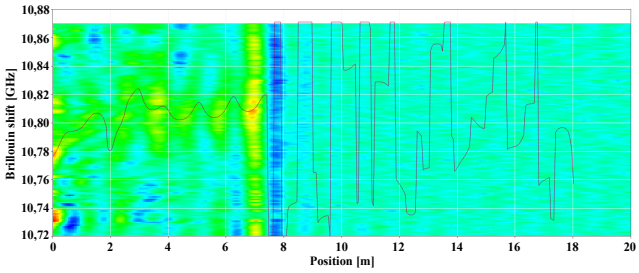


Fig. 4. Chart of Brillouin scatter measurement.

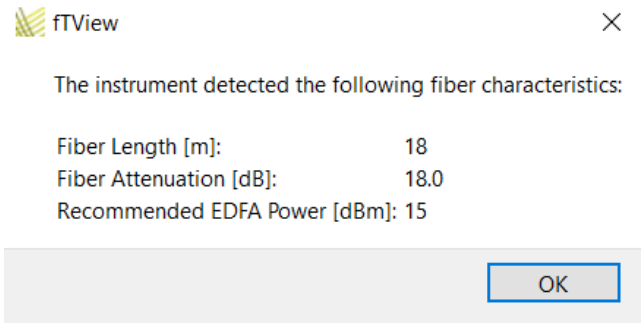


Fig. 5. Measuring the attenuation value of a sensing optical fiber.

The problem of a acute macrobend has been solved with a new probe. The new probe consists of an optical fiber threaded through a plastic tube about 1 m long. At the end of the probe, an arc is formed to prevent the sensor optical fiber from bending acutely, which could invalidate the resulting measurement. Subsequently, the probe with the optical fiber is already installed in the fuel rod in the reactor core, and the two ends of the fiber that lead from it are connected to the measuring instrument.

C. Data processing

The BOFDA measuring instrument continuously senses the Brillouin scattering over time along its entire fiber length. The detection is done in sequences. Each sequence contains fifty individual measurements. One measurement takes approximately twenty seconds, during which time the Brillouin scattering is scanned multiple times for maximum accuracy, and the resulting value is averaged. The sensing of the sensing optical fiber is carried out with a resolution of 5 cm. This means that for every 5 cm of fiber, a Brillouin scattering value is assigned, which is then converted to temperature and stored in a file. Chart of individual Brillouin scattering measurement is shown in Figure 6.

The measured values are stored in the server storage to which the measuring instrument is connected. The data are stored in CSV text format. The server runs a script in the background, written in Python, which records the newly created records of the individual Brillouin scattering measurements. As part of the script, the collected data are uploaded to the created time-series database InfluxDB. To keep the measurement time as accurate as possible, the script extracts the start time of an individual measurement from the name of the text file in which its values are stored. Thanks to this method of obtaining the measurement time, it is also possible to upload the measured data retrospectively.

The data obtained stored in the InfluxDB database can be easily displayed using the Grafana environment, which accesses the data in the database using SQL commands. Grafana allows you to create interactive charts of real-time applications. The created charts can be searched by value and time, and its various scales can be set. Using Grafana, it is possible to compare the measured temperature values with those obtained at midnight when the experimental reactor is in standby mode. From the resulting difference in values, a chart of the reactor reference temperature can be constructed [14], [15].

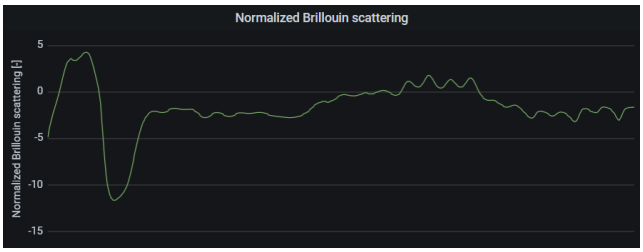


Fig. 6. Chart of individual Brillouin scattering measurement.

III. Results and Discussion

Initial measurements were unsuccessful due to a acutely macrobend of the optical fiber. Therefore, a new probe was created. The probe is made of a plastic tube through which the sensory optical fiber passes. At the end of the probe, a bend is created to protect the fiber from acute bending or breaking. The advantage of this probe is that it can be inserted into the reactor core after the entire fuel rod assembly is in place, and if the probe fails, the probe can be easily replaced, and it is not necessary to wait for the entire fuel rod assembly to be replaced to replace the sensing fiber, which is not as frequent.

The result of the paper is a real-time interactive chart in the Grafana environment. The chart shows the relative temperature of the fuel rods in the core of the nuclear reactor, which is based on the value recorded at midnight, when the experimental reactor is not operating, but is in standby mode. Relative temperature chart is shown in Figure 7. The real-time chart is based on a script written in the Python programming language. The script records the newly created text records from the instrument, processes them, and then takes them along with a timestamp it gets from the name of the text file and sends them to the time-series database InfluxDB.

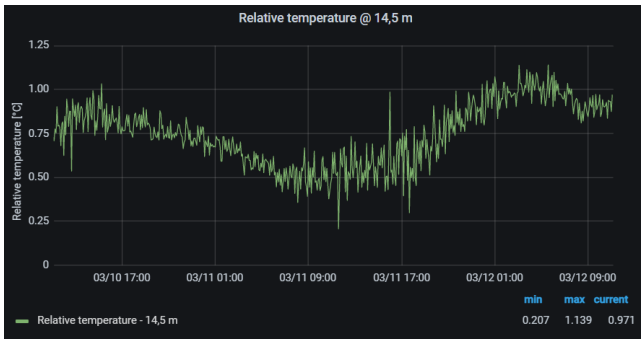


Fig. 7. Relative temperature chart.

IV. Conclusion

The first step to measure the relative temperature in the core of the nuclear reactor was to calibrate the commercial fibrisTerre fTB 2505 BOFDA-based measuring instrument. For the calibration of the instrument, about 20 m long sensing optical fiber was created in the laboratory. Telecommunication fiber G.657 was used, to which connectors were spliced using an optical splicer. Calibration of the sensing instrument was carried out with this sensing optical fiber by inserting part of the sensing fiber into a Vötsch VC3 temperature chamber. Several calibration measurements were performed and plotted on a chart.

Initial sensing of the relative temperature using the sensing optical fiber, which was first inserted into the reactor core together with the fuel rod and then connectors were spliced to it to allow the fiber to be connected to the measuring instrument, was unsuccessful. The failure was due to a acute bend of the fiber as it passed through the fuel rod, resulting in a high fiber attenuation value. Due to the high attenuation of the optical fiber, the measurement was impossible.

To solve the problem of acute bending of the optical fiber, a probe made of a plastic tube was created. The optical fiber is threaded through the plastic tube. At the end of the probe, an arc is formed with a sufficient large radius to prevent acute bending of the optical fiber. A low attenuation of the sensing fiber was achieved with the new probe. This probe also has the advantage that it can be inserted into the reactor core after the entire fuel rod assembly has been inserted.

The new probe made the measurement feasible. The sensor works continuously in time. Approximately every 20 seconds, a Brillouin scattering measurement is taken along the entire length of the sensor fiber with a resolution of 5 cm. A Brillouin scattering value is assigned to each piece of fiber, from which the temperature is calculated. The instrument saves the data in CSV format in a text file.

To process the data from the text files, a script was written in the Python programming language. With each new text file that the script is able to capture, the measurement values are uploaded to the time-series InfluxDB database. The record must include the start time of the temperature measurement, which is obtained from the file name.

The resulting measurement data, which are stored in the InfluxDB database, are processed and displayed using the Grafana environment. The environment allows the creation of interactive charts in real time. In this environment, a chart of the relative temperature in the core of the experimental reactor was created, which can be searched by value and time. This also compares the individual measurements with a reference measurement taken at midnight when the reactor was not loaded, but was in standby mode.

References

[1] R. Šifra, "Optovláknové senzorické systémy," *Automa: odborný časopis pro automatizační techniku*, vol. 2020, no. 5, 2020.

[2] M. Filka and M. Čučka, "Optické vlákno nejen jako přenosové médium," *Next generation telekomunikace [online]*, vol. 2018, no. 4, 2018.

[3] M. Filka, *Optoelektronika pro telekomunikace a informatiku*. druhé, rozšířené vydání ed., 2017. OCLC: 1043865639.

[4] M. Kasinathan, S. Sosamma, C. BabuRao, N. Murali, and T. Jayakumar, "Fiber optic sensors for nuclear power plant applications," *AIP Conference Proceedings*, vol. 1430, no. 1, pp. 1013–1020, 2012.

[5] M. Jelinek, O. Cip, J. Lazar, and B. Mikel, "Design and characterisation of an optical fibre dosimeter based on silica optical fibre and scintillation crystal," *Sensors*, vol. 22, no. 19, 2022. Cited by: 0; All Open Access, Gold Open Access, Green Open Access.

[6] H.-K. Lee, J. Choo, and J. Kim, "16 ch × 200 ghz dwdm-passive optical fiber sensor network based on a power measurement method for water-level monitoring of the spent fuel pool in a nuclear power plant," *Sensors*, vol. 21, no. 12, 2021. Cited by: 1; All Open Access, Gold Open Access, Green Open Access.

[7] H. J. Kim, H. Y. Shin, C. H. Pyeon, S. Kim, and B. Lee, "Fiber-optic humidity sensor system for the monitoring and detection of coolant leakage in nuclear power plants," *Nuclear Engineering and Technology*, vol. 52, no. 8, p. 1689 – 1696, 2020. Cited by: 9; All Open Access, Gold Open Access.

[8] I. Ashry, Y. Mao, B. Wang, F. Hveding, A. Bukhamsin, T. K. Ng, and B. S. Ooi, "A review of distributed fiber-optic sensing in the oil and gas industry," *Journal of Lightwave Technology*, vol. 40, no. 5, p. 1407 – 1431, 2022. Cited by: 5; All Open Access, Green Open Access.

[9] S. Großwig and M. . Schaller, "Monitoring system for previous convergence measurements in tunnel construction," pp. 1149–1151, 2018.

[10] M. Zhao and J. Liao, "Structural safety monitoring during retrofit of shanghai expo puxi entertainment hall," vol. 226-228, pp. 1203–1208, 2012.

[11] "Atominfo.cz aktuálně o jádru," *AtomInfo.cz Aktuálně o jádru [online]*, vol. 2017, no. 11, 2017.

[12] P. Clément, R. Gabet, V. Lanticq, and Y. Jaouën, "B-otdr solution for independent temperature and strain measurement in a single acquisition," *Journal of Lightwave Technology*, vol. 39, no. 18, pp. 6013–6020, 2021.

[13] C. A. Galindez-Jamioy and J. M. López-Higuera, "Brillouin Distributed Fiber Sensors: An Overview and Applications," *Journal of Sensors*, vol. 2012, p. 204121, Oct. 2012.

[14] Grafana Labs, "Grafana documentation," 2023.

[15] InfluxData, "Influxdb documentation," 2023.

ABB v Brně

Lepší svět začíná s vámi

Jsme součástí globálního technologického lídra působícího v oblasti 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 ke mimo jiné v nejvyšší budově světa Burdž Chalífa, v londýnském metru nebo v datovém centru Facebooku.

Přibližně 2100 zaměstnanců

Tržby 7,5 mld. Kč

Významné centrum výzkumu a vývoje

Nejnáročnější projekty ve 100 zemích světa

4 výrobobní závody, centrum digitálních řešení, R&D a servis

Studentům a absolventům nabízíme:
- Placené brigády
- Odborné praxe
- Stáže
- Trainee program
- Vedení balakářské 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 největši jednotku v rámci brněnského závodu. Své produkty dodává do přibližně 100 zemí světa.

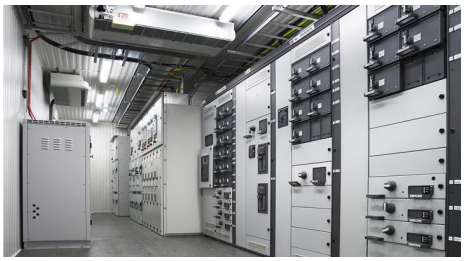




Rozváděče nízkého napětí
Rozváděče nízkého napětí fungují v režimu tzv. virtuálního závodu a sdílí svou kapacitu se závodem v polské Bielsko-Białe. Opět se jedná o jeden z největších závodů svého druhu v Evropě. Jeho produkty jsou dodávány do přibližně 50 zemí světa.

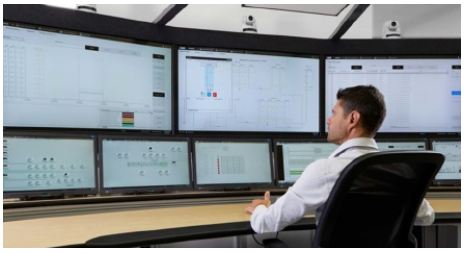
Přístrojové transformátory a senzory
Jednotka přístrojových transformátorů a senzorů je 2. největší v rámci brněnského závodu. Zároveň se jedná o největší závod svého druhu na celém světě. Vůbec první transformátor byl v závodě vyroben již v roce 1919 a jejich výroba je tak založena na více než stoletém 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.



OBD Diagnostic tool

Miroslav Babeľa
Department of Telecomunications
Brno University of Technology
Czech Republic
xbabel02@vutbr.cz

Ing. et Ing. Petr Musil
Department of Telecomunications
Brno University of Technology
Czech Republic
musilp@vut.cz

Ing. Radek Možný
Department of Telecomunications
Brno University of Technology
Czech Republic
xmozny03@vut.cz

***Abstract*—Cars are getting more and more complicated as both technology and government regulations evolve. This fact is not helped by the current automotive trends of hiding everything from the driver, and it is of concern for owners and buyers of cars as they are getting much more expensive to maintain. To reduce maintenance costs to a minimum, owners should have access to an easy to use diagnostic tool that would help them to anticipate potential problems and fix them before they cause more damage. This paper aims to create a device that would provide a way for the driver to see essential engine parameters.**

***Index Terms*—car, diagnostic, STM32, OBD, CAN, K Line**

I. Introduction

In recent years cars have become a lot more complicated. This complexity comes from technological advancements in the comfort and performance departments and governmental safety and environmental protection regulations. However, this fact brings a new set of potential problems for the owners of said cars because as is the norm with anything mechanical or electrical, the question is not if something will break but when it will break.

In order to stay on top of vehicle maintenance, the owners should keep an eye on key engine parameters such as battery voltage, coolant temperature, etc. Monitoring these values can signal potential faults and therefore issues can be rectified before they cause more damage or leave the driver stranded. Gauges for this kind of variables used to be a common sight for older cars, but the modern trend in the automotive industry is to reduce the number of values shown directly on the dashboard to make it less cluttered. This fact leads to critical gauges like coolant temperature being obscured or removed altogether, which prevents the driver from becoming aware of potential faults before they become a more serious problem.

A prime example of this problem is a BMW 3er E90 from 2005 onwards, where the designers have decided to hide the coolant temperature gauge into a service menu inside the board computer, which is not easily accessible to the driver. This is especially a big problem for the owners of the 320d variant, which is notorious for failed thermostats. Without a temperature gauge, this problem is hardly noticeable, especially during warmer seasons, and it can cause further damage to the engine and its accessories due to overcooling the engine.

The device proposed by this paper aims to solve this problem not only for the owners of the BMW E90 generation of vehicles but also to provide an easy way to read out engine parameters on most cars currently present on European roads.

II. On-board Diagnostics

Since the eighties, governments have become aware of the environmental impact of cars and therefore started introducing emission regulations. To comply, manufacturers began to produce more complicated engine designs and soon transitioned to using computers for controlling the engine and its emissions. For the best performance and the least amount of released pollutants, engines must operate within tight tolerances on variables like air-fuel ratio or ignition timing. As mentioned before, due to the added complexity, faults are bound to happen, and therefore there is a need from manufacturers, technicians and government agencies for a way to diagnose these faults, primarily related to emission systems and ensure that the cars on the road are producing pollutants as little as possible. First such diagnostic tools were produced by the vehicle manufacturers themselves, and because of it, they were specific to the cars from that manufacturer. These tools only became standardized, and their support in cars became mandatory in the nineties with the introduction of OBD II set of rules. [1], [2]

These rules combine multiple standards defining communication protocols, physical interfaces for cars and tools, diagnostic functions and trouble codes. Every OBD II compliant car or diagnostic tool must use a specific 16-pin connector, which is specified in ISO 15031-3 / SAE J1962 [4] and can be seen in Figure 1.

Diagnostic functions comprise both a way to read out engine parameters and diagnostic trouble codes. These functions are defined in ISO 15031-5 / SAE J1979 [5]. Each function and parameter available in these standards has its identification number - Service Mode or Parameter ID (PID). Retrieval of any parameter is done by a request from the diagnostic tool to the vehicle's engine control unit (ECU) and reception of the response. For example, a request for coolant temperature would be made with Service Mode 01 and PID 05 at the ECU address. Values inside these PIDs are encoded into four bytes with hexadecimal values. The decoding has to be done with the software of the diagnostic tool while adhering to the said standards.

Manufacturers are required to implement only relevant PIDs to their engines concerning emission systems. Outside of emission systems, car manufacturers can implement their own proprietary PIDs for use with their in-house diagnostic tools.

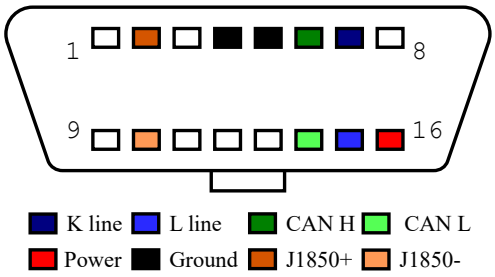


Fig. 1. Female OBD II connector according to SAE J1962

III. Communication protocols

Modern cars use multiple communication protocols inside their electrical system for both communication with the modules inside the car and for communication with an external diagnostic tool. To be compliant with OBD II, the car has to use one of four protocols for communicating with the diagnostic tool - ISO 9141, ISO 14230, ISO 15765 or SAE J1850. The first three protocols have been or are still used in cars from around the globe, while the SAE J1850 was used mainly in North America by GM and Ford.

A. ISO 9141 K Line

This protocol is the oldest one to be used with OBD II dating back to 1989, and it was specifically created for the on-board diagnostic systems. It is based around UART communication with the speed of 10400 Baud, and it can be used in single (K Line) and two-wire (K/L Line) configurations, where the K line is used for both the initialization and subsequent communication. In contrast the L line is only used during the initialization. [6]

Every communication between ECU and the diagnostic tool starts with a specific initialization sequence. During this sequence, ECU and diagnostic tool exchange their addresses and agree on bus speed and timings. Successful initialization is then followed by requests specified by the diagnostic tool. [7]

B. ISO 14230 KWP2000

ISO 14230, or Keyword Protocol 2000, is the newest protocol specified for use with OBD II related communication. It can be seen as an improved version of the ISO 9141 K line which provides enhanced vehicle diagnostic outside the scope of OBD II. Its application layer and network layer can be used on top of another communication protocol. [9]

Similarly to ISO 9141, KWP2000 requires initialization before transmitting a diagnostic request. Besides ISO 9141, ISO 14230 defines 2 types of initialization sequences - fast and slow. The slow sequence is nearly identical to the sequence from ISO 9141, the only difference being the bytes for configuring the bus speed and timings, while the fast sequence is entirely different. The manufacturer can choose one of these sequences for its vehicle, while the diagnostic tool is required to support both of them to ensure universality. [8]

C. ISO 11898 CAN BUS

Besides the previously mentioned protocols, CAN or Controlled area network protocol was originally designed for communication between micro-controllers and other devices on bus without needing a centralized computer. Nowadays is primarily used for in-vehicle communication networks, and with the extension from ISO 15765, it can be used for diagnostic purposes, including OBD II diagnostics.

CAN bus utilizes differential bus for connection to every module on the network [11]. Because CAN Bus is a decentralized network, it uses arbitration for flow control purposes. Every module on the CAN bus has its own identification number, which is used during arbitration, when two or more modules are trying to send a message simultaneously. During the arbitration, IDs are checked bit by bit until exactly one of them has a logic one at position where others have logic zero, then the ID with logic one loses arbitration and has to wait for the bus to be idle. In reality, this translates to lower IDs having higher priority. [10], [12]

IV. Hardware Design

As mentioned before, this paper focuses on the design of a device that will provide a way to conveniently present live engine parameters to the driver. Requirements for this device were:

- Plug & Play - no need for additional setup by the end-user.
- Compact size - OBD interfaces are often located in the drivers foot well and therefore the device has to be as compact as possible in order to prevent accidental damage to it.
- Easily readable from the drivers point of view.
- Low power consumption - user has to be able to leave the device plugged in without worrying about it quickly draining the battery.
- Low cost

There are already devices on the market that can provide such services, but none meet all of the criteria, mainly in the Plug & Play department. The competitors use an app on a smartphone to control the device, which is not suitable for live data monitoring because the user needs to start it up each time they want to check it. It can also be troublesome when the user wants to use other apps simultaneously, such as navigation or music player, and needs to switch between them while driving.

The proposed device uses only a simple button and a display to interact with the user and provide the requested information without the need for additional setup at the start.

Figure 2 shows the block diagram for the proposed device.

A. Micro-controller

As a basis for this device, several options were researched, and a micro-controller (MCU) STM32L341 from ST Microelectronics has been chosen, particularly for its integrated CAN Bus controller, a plethora of other communication interfaces, favorable pricing and good availability.

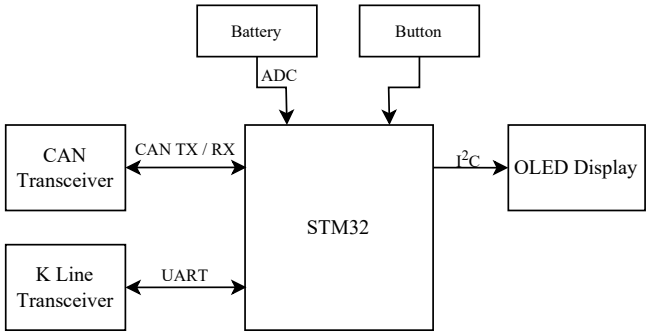


Fig. 2. Block diagram for the proposed device

B. CAN BUS & K Line Transceivers

Although a CAN-BUS controller is already present in the chosen MCU, a dedicated CAN BUS transceiver is required to translate both CAN RX and CAN TX lines to CAN H and CAN L signals. This task is accomplished by CAN BUS Transceiver SN65HVD231 from Texas Instruments. The perk of this particular transceiver is its ability to be turned off when not in use and therefore lowers the power consumption.

K Line uses a modified UART interface with half-duplex communication, but due to different voltage levels, a dedicated K line transceiver is combined with a full duplex interface on the MCU. The K-line transceiver is constructed with discreet components, where a simple voltage divider is used on the RX side and a simple bipolar transistor wired in as a switch on the TX side.

C. Power supply

As the working voltage of the chosen MCU is 3.3 V, a voltage regulator is needed for powering up the circuit inside the car. To not waste energy and create unnecessary heat, a switching voltage regulator TPS562201 from Texas Instruments has been chosen in particular for its relatively small footprint, good efficiency and availability.

D. Display & other peripherals

As mentioned before, the device has to be easily readable and this fact reflected in the selection of the display. The selected display is a cheap 0.96" I^2C OLED display with a resolution of 128x64 and a SSD1306 OLED controller. These displays are readily available and therefore have great support for multiple platforms, including STM32. Also, a big factor in the display selection has been its low power consumption compared to LCDs.

While the device is primarily meant for measuring coolant temperature, it can also be used for measuring all kinds of OBD II defined PIDs and therefore a separate button was added to provide a simple way for the end user to switch between the values.

E. Board & case design

For the final board design, a four layer stackup has been used in order to keep the device compact with single-side assembly for lower manufacturing costs. The board layout can be seen in Figure 3.

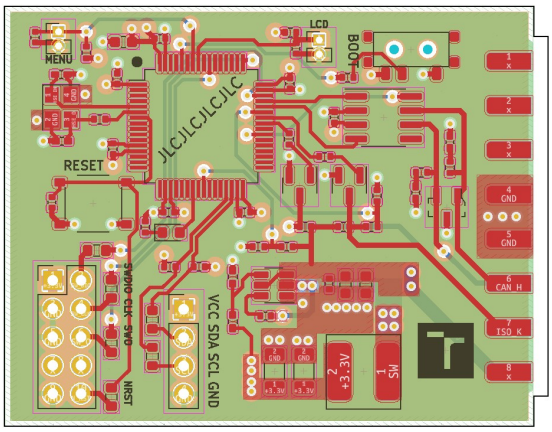


Fig. 3. Board layout

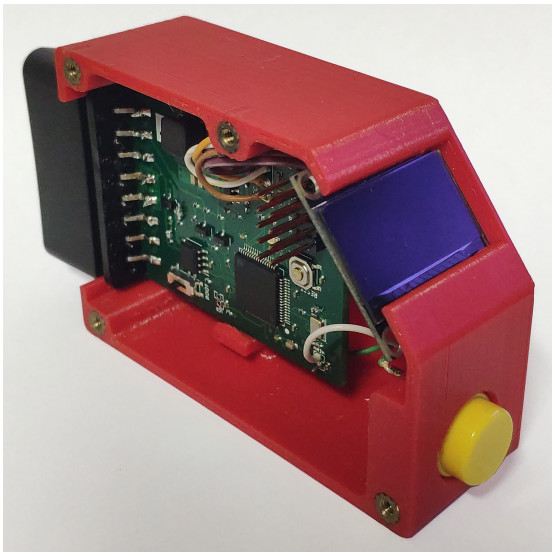


Fig. 4. Final device

The case was specifically designed for use in BMW 3er E90 series and it has display angled towards the driver. Final assembly can be seen in Figure 4. To produce a specifically adapted case for this device, a 3D print technology has been utilized, mainly due to its speed and costs.

V. Software Design

The software side of this device is written in C, mostly using built-in STM32 HAL libraries. The flow chart of the software can be seen in Figure 5.

The program cycle begins with measuring the battery voltage, and if it is higher than 12.9 V, the program continues. Otherwise it will wait for the voltage to increase. After passing the voltage test, the program initializes the required peripherals and tries to initialize communication with the car via one of the three described protocols.

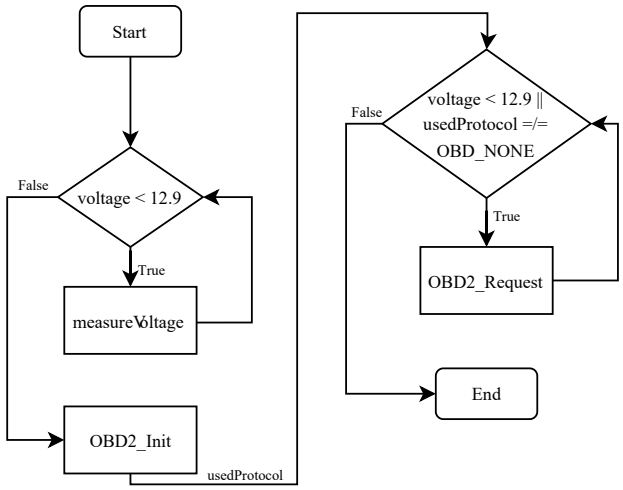


Fig. 5. Software flowchart

Repeated requests for selected PIDs follow successful initialization while the battery voltage is above 12.9 V. If the initialization is unsuccessful or the battery voltage has dropped below the set threshold, the software resets the device after 5 s and the cycle is repeated.

The current version of the software only supports requesting OBD II related PIDs in service mode 01 and battery voltage measurement. The future version of the proposed device will also include fault code read-out function so that the device can be more competitive with existing products on the market.

VI. Cost analysis

One of the proposed device's main goals was the final product's low costs. This is necessary for the final product to be competitive in the current market. Table I shows the final cost summary for one assembly.

The main cost item is the PCB, which comes at 11.10 € for a single board. The board has the most significant potential for cost reduction, where components like switches and debug LEDs could be omitted. For commercial use, economies of scale also come to play because the highest cost of any PCB production is setup and component fees which are fixed costs regardless of the number of ordered boards.

The economies of scale also could help with additionally sourced components like connectors and displays, where a wholesaler could be used for sourcing the components in bulk.

The only item where further reduction of costs is unlikely is the case because economies of scale don't work well for 3D-printed components.

VII. Conclusion

With the current trends in automotive design and the ever-escalating emission regulation, the topic of easy car diagnostics for car owners will become much more prominent. Therefore it can be argued that the proposed device has a huge potential not only for owners of the BMW 3er E90 series of cars but also for owners of other brands.

TABLE I
Cost summary for one device.

Item	**Supplier**	**Price [€]**
PCB	JLCPCB	0.404
SMT Assembly		0.28
Stencil		0.304
Components		4.882
Extended parts fee		2.904
Setup fee		1.618
Components for Hand Soldering	Farnell	0.69
OBD II connector	Aliexpress	3.06
OLED display		1.81
3D Printed Case		1.62
Shipping	DHL	4.422
Taxes		2.588
Summary		**24.582**

The main advantage of the proposed device is its ease of use. Drivers can leave it plugged in, and the device will start displaying the set parameters after the engine has been started, without interaction with the user, unlike currently available solutions on the market. At the time of writing this paper, the fault code read-out function is not available, but with its future implementation, the proposed device could easily compete with the aforementioned devices on the market.

References

[1] NEDELCHEV, Alek, 2022. AUTOMOTIVE DIAGNOSTICS — A BRIEF HISTORY IN HOWEVER MANY CHAPTERS. *Abrites.com* [online]. 2022. [Accessed 9 March 2023]. Available from: https://abrites.com/blog/automotive-diagnostics-a-brief-history-in-however-many-chapters

[2] *DIRECTIVE 98/69/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL*, 1998. [online]. 1998. Available from: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1998L0069:19981228:EN:PDF

[3] DENTON, Tom, 2006. *Advanced automotive fault diagnosis*. 2nd ed. Burlington: Elsevier Butterworth-Heinemann. ISBN 978-075-0669-917.

[4] *Diagnostic Connector Equivalent to ISO/DIS 15031*, 2002. . USA: Society of Automotive Engineers. SAE J1962

[5] *Road vehicles - Communication between vehicle and external equipment for emissions-related diagnostics - Part 5: Emissions-related diagnostic services*, 2006. . Switzerland: International Organization for Standardization. ISO 15031-5

[6] *Road vehicles - Diagnostic systems: Part 2: CARB requirements for interchange of digital information*, 1994. . Switzerland: International Organization for Standardization. ISO 9141-2

[7] MAHAJAN, Gauri, PARCHANDEKAR, S.K. and TAHIR, Mohammad, 2017. Implementation and Validation of K Line (ISO 9141) Protocol for Diagnostic Application. In: *International Research Journal of Engineering and Technology* [online]. Maharashtra, India. 2017. p. 708-713. [Accessed 10 March 2023]. Available from: https://www.irjet.net/archives/V4/i7/IRJET-V4I7181.pdf

[8] *Road vehicles - Diagnostic communication over K-Line (DoK-Line) - Part 2: Data link layer*, 1999. . Switzerland: International Organization for Standardization. ISO 14230-2

[9] FALCH, Martin, 2022. OBD2 Explained: A Simple Intro [2022]. *CSS Electronics* [online]. 2022. [Accessed 10 March 2023]. Available from: https://www.csselectronics.com/pages/obd2-explained-simple-intro

[10] *CAN Specification: Version 2.0*, 1991. . Stuttgart. Robert Bosch GmbH

[11] *Self-study programme 238: Data Exchange On The CAN Bus I*, 2001. . Wolfsburg. Volkswagen AG

[12] *Self-study programme 269: Data transfer on CAN data bus II*, 2003. . Wolfsburg. Volkswagen AG

Multi-Sensor Data Fusion Methods for Localization in Complex Environments

Leida Ioani*, Radovan Juran†
Department of Telecommunications, Brno University of Technology
Brno University of Technology
Brno, Czech Republic
*xioani00@vutbr.cz, †radovan.juran@vut.cz

***Abstract*—This paper brings an insight into RTLS systems and techniques for leveraging advantages of combining several types of motion digitization, in this case UWB TDoA and IMU sensors in cooperation to deal with problems like non/line/of sight (NLOS) or multipath effects, causing errors in wireless-based systems. The fundamental principles of positioning methods are explained and a brief overview of literature is introduced.**

***Index Terms*—ultra-wide band, UWB, inertial measurement unit, IMU, real time location system, RTLS**

I. Introduction

As the digitization of automotive and access control industries increases rapidly, getting implemented in real storage and production facilities, there's need for reliability, precision and service availability for the maximum amount of time.

In these industries not only accuracy is needed but also protection from deliberate attacks as there is a necessity to deal with problems such as NLOS, multipath etc.

Sensor fusion deals with shortcomings of RTLS by combining multiple aspects of data fusion including optimization filters.

In the last couple of years the popularity of UWB significantly increased since it provides very precise and secure positioning. By combining UWB data with data obtained from sensors based on self-contained navigation technique we are able to improve compensation of errors.

II. Ultra-wide Band

UWB is a short-range radio frequency wireless technology. Thanks to its short radio impulses this technology is able to provide fast, accurate and secure positioning. The pulses are transmitted within nanosecond intervals across a wide spectrum in range of 3.1–10.6 GHz, with very low transmission power, bandwidth of hundreds of MHz (typically 500 MHz), and spectral density of -41.3 dBm/MHz, which prevents an interference with other radio transmissions [1].

The UWB technology has evolved in the second half of the 20th century, formerly restricted only to military applications. In early 2000's the FCC (Federal Communication Commission) allocated an unlicensed radio spectrum for the use of UWB as a data transfer technology such as Wi-Fi. However, power restriction rules were set extremely strict which led to delay in IEEE standardization [2].

Eventually standardized IEEE 802.15.4 and IEEE 802.15.4a as an impulse radio based ranging and localization technology mainly used for indoor tracking and as an access control system. The goal in the last couple of years was to introduce more secure access control technology which led to the latest IEEE standard of UWB technology 802.15.4z finalized in 2020 [3].

A. Principle of positioning using TDoA

Time difference of arrival (TDoA) is one of the positioning methods based on multiple fixed time-synchronized anchors. The initiator (tag) sends a signal and anchors receive the signal at different times based on their distance. The position of the device can be estimated using the difference between ToAs (time of arrival) of radio signals sent to the anchors of known location. This method is well used since it requires very low energy consumption [4]. The principle is based on multilateration. Positions can be described with a hyperbola equation (1) with the knowledge of the time difference of arrival. As seen in the Fig. (1), at least three anchors are necessary to determine position of an object in 2D space [5].

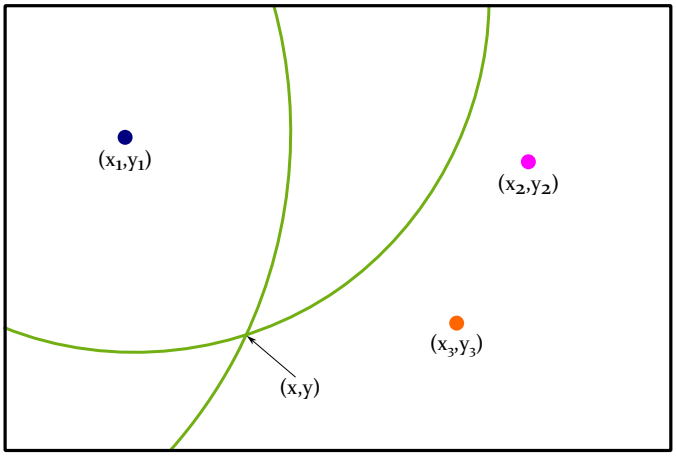


Fig. 1. Hyperbolic TDoA [6]

$$\sqrt{(x-x_i)^2+(y-y_i)^2}-\sqrt{(x-x_j)^2+(y-y_j)^2}=d_i-d_j \quad (1)$$

x_i, y_i, x_j, y_j – position of the anchors
x, y – position of the initiator of the signal
d – distance between the anchors (can be calculated by multiplying time of flight with speed of light)

III. Inertial navigation

Systems utilized in inertial navigation don't rely on any radio signals or satellites. Inertial navigation systems usually contain three orthogonal IMU (inertial measurement unit) sensors like accelerometer and gyroscope. These sensors are typically silicone based miniaturized mechanical and electrical systems created with techniques of microfabrication. The three axes allow to determine pitch, yaw and roll angles including acceleration (a) in every direction while gyroscope allows the system to align accelerometer horizontally to the Earth since the system is not aligned to horizon as a result of Earth's rotation. Measurements provided by these sensors are used to track the position and orientation of an object in real time.

A. Principle of positioning

The first integration of acceleration determines velocity with respect to time.

$$\begin{aligned} a &= \frac{dv}{dt} \\ dv &= a\ dt \\ \int_{v_0}^{v} dv &= \int_{0}^{t} a\,dt \\ v &= v_0 \pm at \end{aligned}$$

With the second integration of acceleration we are able to determine the distance which allows us to plot trajectory [7].

$$\begin{aligned} v &= \frac{ds}{dt} \\ ds &= v\ dt \\ ds &= (v_0 + at)\ dt \\ \int_{s_0}^{s} ds &= \int_{0}^{t} (v_0 \pm at)\,dt \\ s &= s_0 + v_0 t \pm \frac{1}{2} a t^2 \end{aligned}$$

IV. Sensor fusion

There are many issues that can disrupt the precision of the wireless transmission technologies. And when it comes to radio-based localization technology, the precision is very much needed since time reception must be measured as accurately as possible.

One of the situations that is possible to encounter during radio-based localization is multipath which is a reflection of a signal and a result of NLOS (Non Line of Sight) scenario.

Line of sight (LOS) is a vital factor in radio-based localization. It is the direct path with shortest distance between a transmitter and a receiver of a signal. Since radio waves travel in straight lines, the space between devices should be clear for effortless transmission. That is really hard to achieve with all of the possible obstacles, such as furniture, buildings, walls, etc.

In the situations where the transmitter is in non line of sight to the receiver, it has to be dealt with multipath (2). The problem with multipath is that the reflected signals can be received as well as the direct signal. Since the direct signal travels the shortest distance, it is received sooner than the reflected ones. This can lead to several situations. The multiple signals received by the device can have different phase thanks to the delay. Worst possible scenario is the opposite phase of signals which results into mutual disrupt, meaning the signal is received as a null.

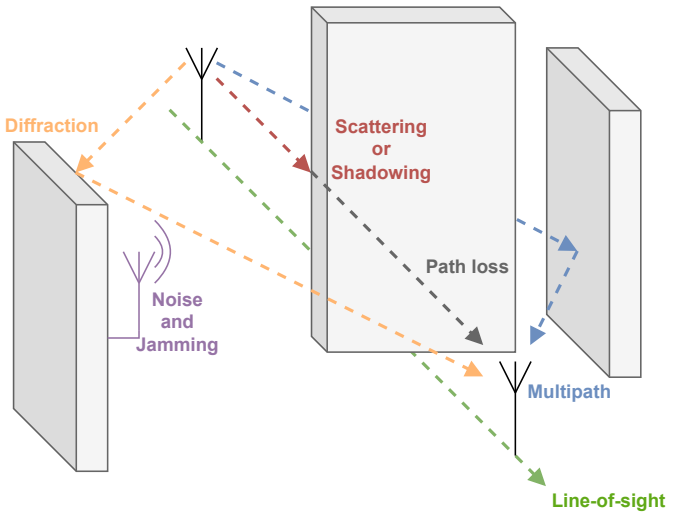


Fig. 2. General problems of wireless systems could become critical.

The goal of sensor fusion is to overcome the limitations of an individual sensor, thus obtain more accurate results. Gathering data from several sensors allows to produce more robust information so the reliability grows with less uncertainty [8].

A. Kalman filter

Kalman filter is an optimal estimation algorithm that predicts location, speed and direction. It is basically mathematical tool useful for statistical analysis and predictive design of sensors. Kalman filter provides techniques to reckon missing information from indirect or noisy measurement and predict future courses of dynamic systems. Due to this reason, Kalman filter is one the most used estimation algorithms in navigation.

The execution of the algorithm is based on linear quadratic estimation and it is divided into two main steps. During the fist step the estimated values are created with predicted uncertainty. The second step is the correction. The estimated values are corrected with the use of average factor. The algorithm is recursive which makes it able to work in real time only needing the current value and the previous one.

There are other variants of Kalman filter like EKF (Extended Kalman Filter) or UKF (Unscented Kalman filter) [9].

B. Complementary filter

Complementary filter is another very effective method for sensor fusion. It uses low computational and processing power and provides very quick estimation of angles. This filter is highly used to merge measurement received from sensors like accelerometer and gyroscope since it is not as complex to implement as Kalman filter. The gyroscope provides precise

values in the short-term, but over certain period of time the values become less accurate as a result of the sensor drifting. The accelerometer provides long-term stability, however the sensor suffers from vibration which means that returned values are corrupted with noise.

A complementary filter provides precise measurement by high-passing the gyroscope measurements and low-passing the accelerometer measurements, which allows to bring up the benefits of each sensor – gyroscopes short-term accuracy and the accelerometers stability over a long time periods [10].

V. Results

In some instances the UWB data from some of the anchors may not be received, due to the environmental difficulties.

Raw data obtained from the accelerometer need to be calibrated to linear acceleration by reducing gravity force.

Optimization filters provide techniques to reckon missing information from indirect or noisy measurement. Here are presented the most popular filters for data processing.

Paper	Sensors	Optimization filter
[10]	IMU	KF, CF
[11]	UWB, IMU	EKF
[12]	UWB, IMU	FK
[13]	RSSI, IMU	EKF
[14]	UWB, IMU	EKF
[15]	UWB, IMU	EKF, IEKF
[16]	IMU	EKF, CF

filters: KF – Kalman, EKF – Extended Kalman, IEKF – Iterative Extended Kalman, CF – Complementary

VI. Discussion

In future we would like to focus our work on implementation of the optimization algorithms and test them on a one calibrated, real-world dataset to compare their performance. As we could see in the comparison table, filters that are used the most for UWB and IMU fusion are various versions of the Kalman filter. Selected programming language was Python, specifically version 3.9 as Python has many libraries which are useful for optimization algorithms and data visualization in plots.

Necessary data for our topic are measurements about linear acceleration and UWB. To be able to hold a suitable dataset that meets our requirements for indoor localization systems, we also plan to collect our own data for the performance comparison tests.

VII. Conclusion

We bring basic principles of UWB and IMU trajectory determination, positioning and their implementation for raw data preprocessing. It was described how UWB uses time of the flight of a signal to determine position with the TDoA ranging method. The issues and possible attacks in wireless positioning were evaluated as well as possible problems that can be encountered during inertial navigation. Inertial sensors and their use in inertial navigation were also depicted.

References

[1] Eliko. In *The Story Behind Ultra-Wideband Technology and Indoor Positioning*, 2018. II

[2] L. B. Michael a R. Kohno Ghavami M. In *Ultra Wideband Signals and Systems in Communication Engineering*, England, 2004. John Wiley & Sons. II

[3] Zignani Andrew a Stephanie Tomsett. Abi research. In *ULTRA-WIDEBAND (UWB) FOR THE IOT– A FINE RANGING REVOLUTION?*, USA, 2021. II

[4] Inpixon indoor intelligence. In *Precise, Reliable and Scalable Positioning Calculations: What is Time Difference of Arrival (TDoA)?*, 2022. II-A

[5] Mauricio A. Caceres Duran, Antonio A. D'Amico, Davide Dardari, Mats Rydström, Francesco Sottile, Erik G. Ström, and Lorenzo Taponecco. Chapter 3 - terrestrial network-based positioning and navigation. In Davide Dardari, Emanuela Falletti, and Marco Luise, editors, *Satellite and Terrestrial Radio Positioning Techniques*, pages 75–153. Academic Press, Oxford, 2012. II-A

[6] Tarvis Mick. In *A demonstration of TDOA multilateration*, 2022. 1

[7] Myron Kayton. In *Navigation Land, Sea, Air and Space*, New York, 1990. The Institute of Electrical and Electronics Engineers. III-A

[8] Yan Zhang and Linfu Duan. A phase-difference-of-arrival assisted ultra-wideband positioning method for elderly care. *Measurement*, 170:108689, 2021. IV

[9] Mohinder S. a Angus P. ANDREWS GREWAL. In *Kalman Filtering: Theory and Practice Using MATLAB 3st ed*, Canada, 2008. John Wiley & Sons. IV-A

[10] Robottini. In *Kalman filter vs Complementary filter*, 2011. IV-B, V

[11] Giovanni Miraglia, K. Niki Maleki, and Loyd R. Hook. Comparison of two sensor data fusion methods in a tightly coupled uwb/imu 3-d localization system. In *2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, pages 611–618, 2017. V

[12] Fengbo Wu and Zhaoqi Liu. Research on uwb / imu fusion positioning technology in mine. In *2020 International Conference on Intelligent Transportation, Big Data Smart City (ICITBS)*, pages 934–937, 2020. V

[13] Veerachai Malyavej, Warapon Kumkeaw, and Manop Aorpimai. Indoor robot localization by rssi/imu sensor fusion. In *2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, pages 1–6, 2013. V

[14] B. Venkata Krishnaveni, K. Suresh Reddy, and P. Ramana Reddy. Indoor positioning and tracking by coupling imu and uwb with the extended kalman filter. *IETE Journal of Research*, 0(0):1–10, 2022. V

[15] Haifeng Zhang, Zhitian Li, Shuaikang Zheng, Yunfei Liu, Pengcheng Zheng, and Xudong Zou. Uwb/ins-based robust anchor-free relative positioning scheme for ugvs. *Measurement Science and Technology*, 33(12):125007, sep 2022. V

[16] Brandon McCarron. In *Low-Cost IMU Implementation via Sensor Fusion Algorithms in the Arduino Environment*, USA, 2013. California Polytechnic State University, San Luis Obispo. V

Principle Behavior of the Compensation Algorithm with Respect to the Input Data

P. Repa, M. Bubniak, and P. Musil
Department of Telecommunications
Brno University of Technology
Brno, Czechia
230632@vut.cz, xbubni03@vutbr.cz, xmusil56@vutbr.cz

***Abstract*—This article shows how CMLs (commercial microwave links) can be used to estimate precipitation activity. This is a relatively new technique for obtaining information about atmospheric precipitation. The article describes the whole procedure from obtaining raw data from the telecommunications service provider to the calculation of the rainfall fields represented by maps, which is done by software Telcorain. Special focus is on the temperature compensation algorithm, which can correct fluctuations in a received signal level of a CML unit. Finally, how to use a non-conforming link with a high correlation value in the calculation of the average rainfall intensity and the generation of rainfall maps is explained.**

***Keywords*—commercial microwave links, correlation, received signal level, rainfall, temperature compensation**

I. Introduction

CML networks are increasingly being used for rainfall observation due to their wide coverage and cost-effectiveness. These networks use existing microwave communication infrastructure to measure the attenuation of signals caused by rainfall. They have the potential to provide valuable rainfall data for hydrological and weather forecasting applications. Fig. 1 shows a line-of-sight radio link between two locations, typically a transmitter and receiver. CMLs are generally used to interconnect mobile telephone base stations. These links operate with a very small wavelength, which allows the antennas to direct the transmission of information into narrow beams that can be directly pointed at the receiver antenna.

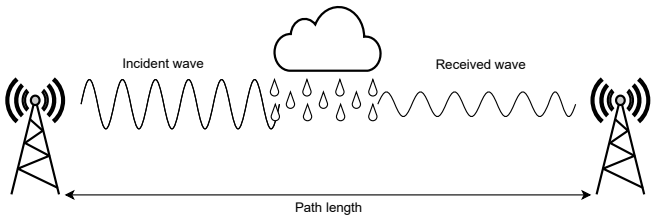


Fig. 1. Rainfall estimation using microwave links.

The use of microwave links in precipitation monitoring was extensively researched by Dutch researchers R. Uijlenhoet, A. Overeem, and H. Leijnse. They wrote the summary paper [1], which aims at using microwave links as an alternative or additional means for area-based rainfall monitoring. They were the first to write and publish code (in language "R") for processing microwave link data and rainfall mapping, which is publicly available [2]. The members of German team led by C. Chwala and H. Kunstmann are the next researchers who are significantly active in this field. In 2019, they published a review article [3], in which they introduce the technique of precipitation estimation using CML networks, describe an overview of current results, and critically assess the possible limitations and future challenges of the technique. The first experiment providing empirical evidence that it is possible to estimate precipitation activity using microwave links was conducted in 2006 by researchers from Israel (H. Messer, A. Zinevich, and P. Alpert) [4].

II. Principle of precipitation estimation

The estimation of precipitation using microwave links is based on a decrease of the power of the received electromagnetic radiation signal. From the decrease in the received signal level relative to a reference received signal level, which is determined during the dry season, the attenuation rate due to precipitation can be calculated. Using the amount of attenuation, it is possible to determine the average rainfall intensity along the path of the link. The attenuation rate increases significantly at frequencies above 15 GHz.

The estimation of precipitation using microwave links can be used to create (real-time) precipitation maps. This means that microwave links can complement precipitation information where rain gauges and weather radars are also present. The relationship between the attenuation caused by rain and the intensity of precipitation is expressed by the following equation:

$$R = ak^b \tag{1}$$

where R is the rain intensity in (mm/h), k is attenuation due to rain in (dB) (attenuation can be caused by absorption, scattering and multipath propagation), the coefficients a and b are empirically derived and depend on the frequency of the microwave link. The values of a and b are given by the tables of ITU-R Recommendation P.838 [5].

The whole process of precipitation estimation consists of the following steps: 1. Raw data processing, 2. Classification of dry and wet seasons, 3. Determination of the reference signal level, 4. Calculation of average rainfall intensities, 5. Generation of precipitation maps [2].

III. From raw data to the calculation of rainfall intensity

A. Raw data processing

The first and often most important step is to obtain the raw data of the CMLs. This data typically consists of the RSL (received signal level) and TSL (transmitted signal level) of a given CML channel. Temperature data of units are also desirable, but not always available. In most cases, microwave link data are not made available to the public by service providers, and access to this data is often time-consuming.

A database export of the parameters of microwave links located in the territory of the city of Prague for a period of 3 years (2020, 2021, and 2022) was received from the provider of CMLs. As a storage of these data the InfluxDB database is being used. The individual CML channel data are indexed using the IP addresses of the units and have a time resolution of 30 seconds or 5 minutes (some periods).

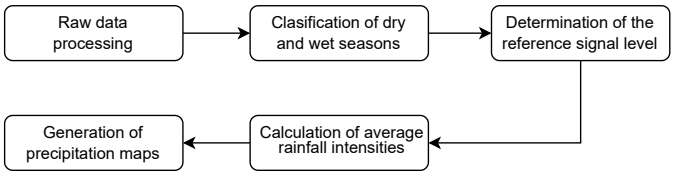


Fig. 2. Process of precipitation estimation.

B. Classification of dry and wet seasons

A wet season occurs when rainfall totals are greater than the amount of water evaporated into the atmosphere over a number of consecutive days, resulting in an excess of rainwater. A dry season occurs when rainfall totals are less than the amount of water evaporated into the atmosphere over several consecutive days, leading to a deficit of rainwater. As can be seen in Fig. 3, the rainfall curve (pink) shows the rainfall activity that was recorded by the rain gauge from Czech Hydrometeorological Institute (CHMI). At nearly the same moments, the microwave link recorded an increase in attenuation of TRSL (transmitted *minus* received signal level). These sections are evaluated as rain events. Classification is necessary for the correct determination of the reference signal level because variations in the RSL of a CML occur not only during precipitation, but also during changes in air temperature, humidity, etc.

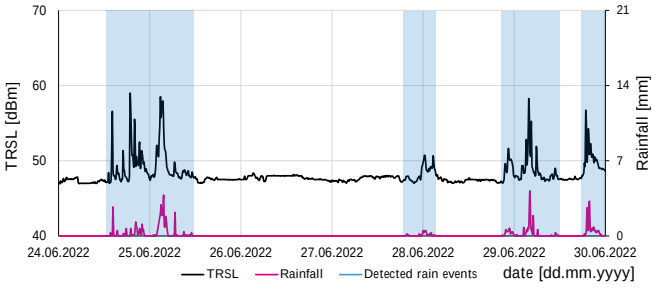


Fig. 3. Identification of rain events in raw data.

C. Determination of the reference signal level

In this step, a reference value of the received signal is established against which the attenuation due to the precipitation activity is calculated. Ideally, this should be a value of the signal just before the start of a rain event.

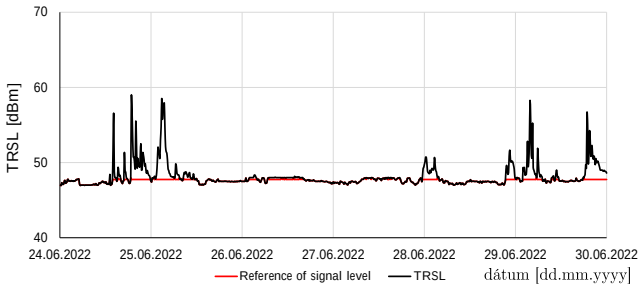


Fig. 4. Determination of the reference signal level.

D. Calculation of average rainfall intensities

By subtracting the current TRSL value from the reference signal level the attenuation k in (dB) is obtained, from which we get the average rainfall intensity R in (mm/h), Equation (1). At this point any compensation for parasitic attenuation, e.g. wet antenna attenuation (WAA) is performed. The precipitation monitoring by CMLs in connection with problem of WAA has been studied by a Czech group of researchers led by V. Bares and M. Fencl. They have published several papers about the measuring of WAA and proposed several methods for its quantification [6], [7].

E. Generation of precipitation maps

The last and final step is to convert the average rainfall intensity into a two-dimensional domain. Several interpolation and reconstruction techniques can be used. The interpolation of the data is performed not only for points along the path of the links, but also for points lying in the interspace between the paths of the individual links. By generating graphical outputs, we obtain a precipitation map, which is shown in the Fig. 5. The output rainfall fields are generated by software Telcorain created by M. Bubniak in Python language, using pycomlink library of C. Chwala et al. [8]

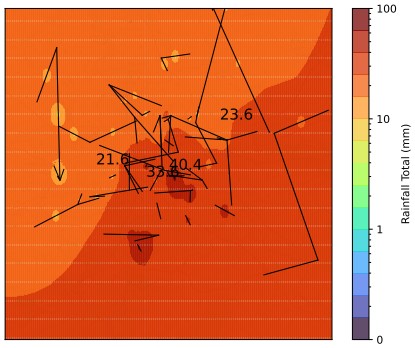


Fig. 5. Generated precipitation map using microwave links.

IV. Compensation for fluctuations in the received signal levels

For future needs to compensate for variations in the received signal level, it is necessary to design an algorithm that performs correlation analysis using a suitable method. In addition to the design of the compensation algorithm, an analysis of the correlation coefficient between the received signal level and the temperature of the microwave unit has been performed. When the specified value of the correlation coefficient is exceeded, the compensation of the given microwave link will be performed. This will ensure that the microwave link can still be used in the calculation of the precipitation estimation using microwave links. This compensation is achieved by mathematical calculation, based on the reference value of the microwave link temperature. The reference value was derived during the winter months, when the temperature dependence is not influenced to such an extent by the temperature of the microwave unit. Using linear regression, the values are corrected to produce a new value P_{corig}.

$$P_{corig} = RSL_{actual} - a(T_{actual} - T_{ref}) \tag{2}$$

where P_{corig} is the corrected RSL received signal level in (dBm), RSL_{actual} is the actual level of the RSL signal at the given time instant in (dBm), a is the component from the linear regression, which is of the form `y = ax + b`, T_{actual} is the actual temperature of the microwave link at a given time instant in (°C), T_{ref} is the selected reference temperature of the microwave link in (°C).

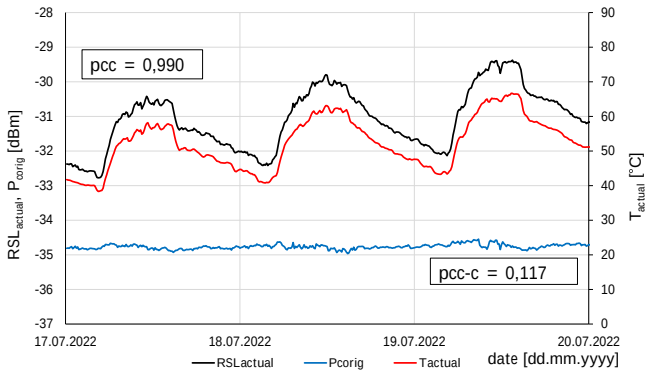


Fig. 6. Compensation for fluctuations in the received signal levels.

Fig. 6 shows the compensation for fluctuations in the received signal level. This is the identified dry session from 17th July − 20th July 2022. As mentioned above, fluctuations in the received signal level occur not only during precipitation activity, but also during changes in air temperature. This is the case when the fluctuation is caused by the air temperature, respectively the temperature of the microwave link. It is necessary to compensate for these fluctuations as far as possible.

For the case displayed in Fig. 6, the fluctuation of the received signal level (black curve) is around 3.5 dBm, which is a very high value proven also by Person's correlation coefficient (*pcc*), which takes on a value of 0.990. In other words, this means that the received signal level has a very strong positive dependence. This significant fluctuation is undesirable for both researchers who use microwave links to estimate precipitation activity and for telecommunications service providers. For them it means a possible degradation of transmission speed and also threat of a link outage, which is highly undesirable for a CML user as well. Thanks to the compensation algorithm, the compensation of the microwave connection was made possible, which means that the fluctuation of the received signal level is around 0.5 dBm (blue curve). The corresponds to a value of Peason's correlation coefficient (*pcc_c*) of 0.117. The fluctuation value of the received signal level around 0.5 dBm is taken as a sample behavior of the microwave link. It can also be said that for this value, we could characterize the received signal level of the microwave link as a reference value of signal level.

V. Flow chart

Flow chart of the compensation algorithm, which is used in compensating for the temperature instability of the received signal level of the microwave link is shown in Fig. 7.

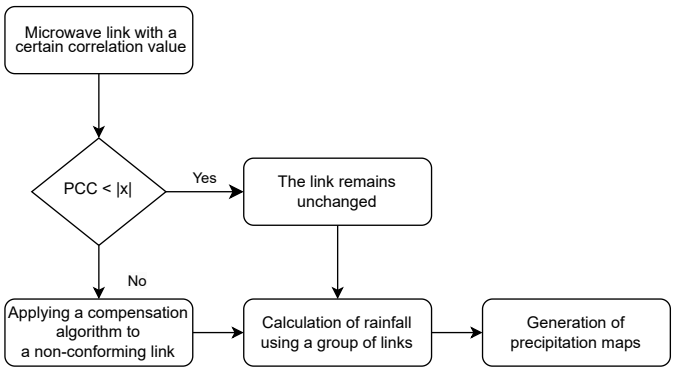


Fig. 7. Flow chart of compensation algorithm.

The basis of the algorithm is the retrieval of data from InfluxDB. These are then processed and stored into multi-dimensional arrays by software Telcorain. From there, the data are loaded and work begins. The first step is to split the microwave link into side A and side B. The second step is to check the correlation value. If the correlation value is greater than a specified threshold of the correlation measure PCC (Pearson correlation coefficient), a compensation algorithm will be applied to the given non-compliant microwave link to compensate for the received signal level. The compensated received signal level is then used to compute the precipitation map. If the correlation value is less than the specified threshold of the correlation measure PCC, the compensation algorithm will not be applied to the microwave link and will be left unchanged. If at least one side of the microwave link, side A or side B, has a PCC greater than the specified value of the correlation rate, the compensation algorithm will be applied to both sides of the microwave link. This link is then used in the calculation of rainfall fields in form of precipitation map.

VI. Conclusion

Precipitation estimation using microwave links is a reliable and efficient method for creating precipitation maps, possibly in real-time. This technique provides valuable information on rainfall intensity, which is crucial for various applications, such as weather forecasting, hydrological modeling, and flood warning systems. The process of precipitation estimation involves several steps, including raw data processing, classification of dry and wet seasons, determination of the reference signal level, calculation of average rainfall intensities, and generation of precipitation maps. The accuracy of precipitation estimation using microwave links can be improved by compensating for parasitic attenuation and designing an algorithm that performs correlation analysis. The use of microwave links as a complementary method to rain gauges and weather radars can enhance our understanding of precipitation patterns and their impacts on the environment and society.

Acknowledgment

This project was supported by company CBL Communication by light s.r.o. which provided access to the recorded parameters of its Prague commercial microwave links for the purposes of this work.

References

[1] R. Uijlenhoet, A. Overeem, and H. Leijnse, "Opportunistic remote sensing of rainfall using microwave links from cellular communication networks", WIREs Water, vol. 5, no. 4, pp. 713-713, 2018, doi: 10.1002/wat2.1289.

[2] A. Overeem, H. Leijnse, and R. Uijlenhoet, "Retrieval algorithm for rainfall mapping from microwave links in a cellular communication network", Atmospheric Measurement Techniques, vol. 9, no. 5, pp. 2425-2444, 2016, doi: 10.5194/amt-9-2425-2016.

[3] C. Chwala and H. Kunstmann, "Commercial microwave link networks for rainfall observation: Assessment of the current status and future challenges", WIREs Water, vol. 6, no. 2, 2019, doi: 10.1002/wat2.1337.

[4] H. Messer, A. Zinevich, and P. Alpert, "Environmental Monitoring by Wireless Communication Networks", Science, vol. 312, no. 5774, pp. 713-713, May 2006, doi: 10.1126/science.1120034.

[5] ITU-R P.838-3: Specific attenuation model for rain for use in prediction methods. Geneva: International Telecommunication Union, 2005

[6] M. Fencl, P. Valtr, M. Kvicera, and V. Bares, "Quantifying Wet Antenna Attenuation in 38-GHz Commercial Microwave Links of Cellular Backhaul", IEEE Geoscience and Remote Sensing Letters, vol. 16, no. 4, pp. 514-518, 2019, doi: 10.1109/LGRS.2018.2876696

[7] P. Valtr, M. Fencl, and V. Bares, "Excess Attenuation Caused by Antenna Wetting of Terrestrial Microwave Links at 32 GHz", IEEE Antennas and Wireless Propagation Letters, vol. 18, no. 8, pp. 1636-1640, 2019, doi: 10.1109/LAWP.2019.2925455

[8] M. Bubniak, P. Musil, P. Mlynek, L. Benesl, and M. Rusz, "Design and evaluation of the application for a measurement of rainfall via microwave links," in 2022 14th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2022, pp. 8–13.

Model of hot water exchanger station

1st Martin Major
Brno University of Technology
Faculty of Electrical Engineering and Communication
Department of Control and Instrumentation
616 00 Brno, Technicka 3082/12, Czech Republic
Email: 221002@vut.cz

2nd Tomáš Sýkora
Brno University of Technology
Faculty of Electrical Engineering and Communication
Department of Control and Instrumentation
616 00 Brno, Technicka 3082/12, Czech Republic
Email: 173753@vut.cz

***Abstract*—This article focuses on a custom-built heat exchanger station for teaching in the laboratories of the Department of Control and Instrumentation at the Faculty of Electrical Engineering and Communication in Brno. It deals with the identification of the components already installed in the station model and the programming of the system for its control. A heat exchange station is a device that is used to safely transfer heat energy between heat transfer fluids. The entire station and the heat exchange process are controlled by a PLC from Siemens and operated via an operator panel. The control program and the visualization of the operator panel were created in the TIA Portal development environment from Siemens. The used programming language is FBD.**

***Index Terms*—Exchanger, Heat exchanger station, Siemens, Tiaportal, HMI**

I. Introduction

This paper describes the individual components involved in the station but also describes in detail the software used for the control program. The blocks used in the program and the screen layout of the control panel are described, including their options. The model station is used in school laboratories to provide students with hands-on experience in controlling and managing industrial processes, such as just fluid temperature control.

II. Exchanger stations

A heat exchanger is a device that transfers heat energy from the primary circuit to the secondary circuit by means of the flow of heat transfer media. Heat exchangers can be divided into 4 main types according to the principle of heat transfer:

- recuperative
- regenerative
- contact
- mixing

Another divide is according to the medium used:

- water-water
- steam-water
- oil-water

An important divide is also according to the mutual movement of the heat transfer media:

- concurrent
- countercurrent
- cross

Exchangers can also be divided according to their design into plate and tube exchangers, which differ from each other not only in construction but also in use and parameters. [1]

III. Exchanger station model

The model created for the purpose of this work will be placed in the laboratories at the Department of Control and Instrumentation. It will be used for teaching purposes in the field of control and regulation, and verification of theoretical knowledge of control in practice. The model can also serve as a demonstration of the control of the technological process of temperature control or the use of PID controllers.

A. Heat exchanger

This station uses a soldered plate heat exchanger from ALFA LAVAL. The plate heat exchanger has up to ten times higher conversion coefficient than the tubular heat exchanger. It is made up of a set of metal plates that are clamped together and fitted with a rubber seal around the perimeter to create channels for water flow. There are two openings on the sides of the plate where one acts as an input and the other as an output. So on one side of the plate, the hot medium flows in, and the cold medium flow out. The heat permeates through the walls of each plate and thus heat is transferred between the media. The design of the heat exchanger and the principle of media flow are shown in Figure 1, which is taken directly from the manufacturer Alfa Laval. [2]

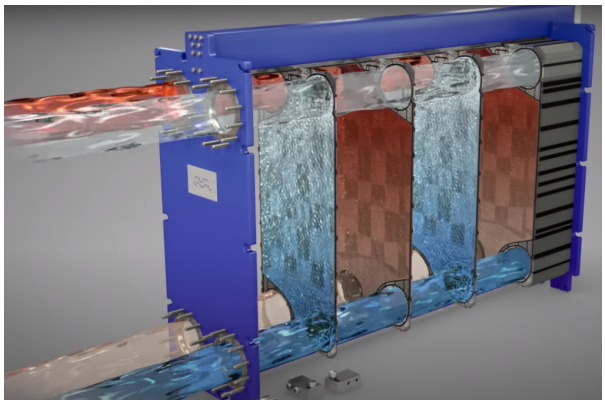


Fig. 1. The principle of media flow in plate heat exchanger [2]

In the picture above, the primary (red) fluid enters the upper left flange of the exchanger, then flows through each even channel downwards, where it then leaves the exchanger through the lower left flange. The secondary (blue) fluid, on the other hand, enters the exchanger through the lower right flange and flows up through each odd channel, where it exits the exchanger through the upper right flange. Importantly, the cold fluid flows in the opposite direction to the warm fluid, so it is a counter-current exchanger. This results in a more efficient heat transfer than a countercurrent or cross-current exchanger.

B. Pipeline

The model of the exchanger station contains the already mentioned two mutually separated circuits connected by metal pipes. In each circuit, there is a circulating pump to drive the liquid, thermometers, flow meters, analog pressure gauges, differential pressure gauges, valves operated by servo motors, and expansion tanks for pressure equalization. Thermometers include both electrical temperature sensors for switching valves and supplying information to the control PLC, as well as simple analog display thermometers for temperature control in the line.

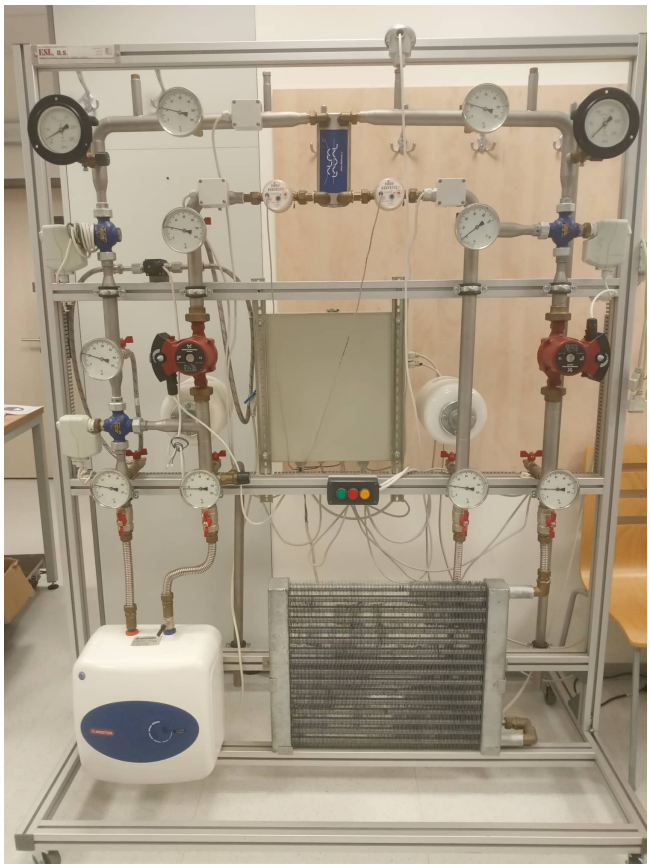


Fig. 2. Assembled exchanger station

In the primary circuit, there is a flow heater as a source of hot water, and in the secondary circuit, there is a heat appliance, i.e. a radiator.

Water temperature regulation is provided by valves with actuators, which open and close according to the requirement, thus controlling the flow of hot water to cold water. In the primary circuit, the valve regulates the temperature of the water returned from the heat exchanger, and in the secondary circuit, it regulates the temperature of the water entering the heater.

C. Pumps

The pumps that drive the water are from Grundfos, designation Alpha+ 25-40 180. It is a pump that can be controlled according to proportional pressure, i.e. maintaining pressure according to the flow rate, or according to a set constant pressure. It is also possible to set the pump manually, using a switch on the terminal block. In the station model the pumps are marked as C [3]

D. Valves and actuators

For water flow control, valves type RV 111 R from LBM are used, which are designed directly for heating equipment due to their high tightness and flow characteristics. Three three-way valves are used at the station. The valves are controlled by a servo drive type SSC61 from Siemens. The actuators are powered by 24V and controlled by a DC 0 - 10V signal, where 0 Volts means fully closed and 10 Volts means fully open. The changeover time from one stop to the other is given as 30 seconds. [4]

E. Thermometers

The model uses 4 resistance thermometers from Rawet. The T3 thermometer is designed for contact temperature measurement and installation directly into the pipe. The remaining thermometers T1, T2, T4 are designed for measuring the surface temperature of the pipe. All thermometers include a PT100 transmitter with a 4-20 mA current output.

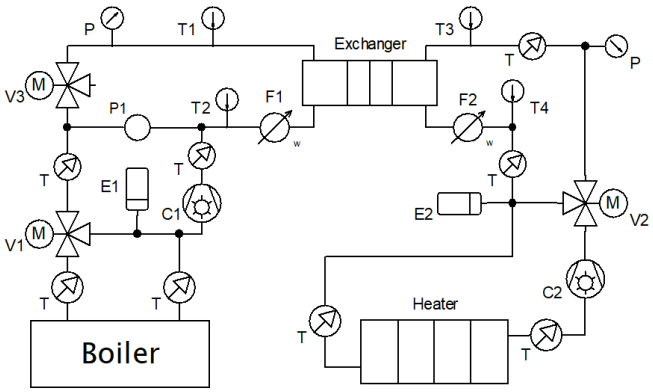


Fig. 3. Exchanger station model

F. PLC

A PLC from Siemens was selected to control the entire station. The PLC was selected based on availability and the

company's cooperation with the Faculty of Electrical Engineering and Communication. Specifically, the model 1512C-1 PN from the S7-1500 series was selected. This series is considered by the manufacturer to be one of the fastest control systems in the world. It has high performance and system bus speed, where it can process commands in nanoseconds. The PLC supports the Profinet interface. 1x AI/AO card, which expands the PLC by 5 analog inputs and 2 analog outputs, and 2x DI/DO card, which expands the PLC by 16 digital inputs and 16 digital outputs. The process flow will be monitored and controlled from the TP 700 Comfort touch panel. [5]

IV. Software equipment

A. Development environment

The control system for the operation and control of the exchanger station was developed in the TIA Portal development environment, which combines a number of engineering tools into one platform. This environment allows writing the control program in 4 programming languages (Ladder Diagram, Function Block Diagram, Statement List, and Structured Control Language). The environment also includes, for example, simulation tools and functions for efficient diagnostics. TIA Portal is the official development environment from Siemens. The environment is not freely downloadable, a license is required to use it.

B. Control program

A library with function blocks for individual elements such as actuators, analog inputs, and analog outputs has been created. A function block always contains the main logic. For example, standardizing the analog value for analog input and output using the SCALEX instruction. The analog input block also includes the ability to set a simulation value and compare it to users' limits.

The programmed block of the action item is always used in the function for the given action item, its own function data block is created and then the function is used in the Main program.

Control of the actuators is possible in two modes. Automatic mode for automatic control of the heating process, and manual mode for service purposes. It is therefore possible to switch on, for example, only one pump independently. At the same time, it is always necessary to switch the MOD switch to the desired mode to block the other mode. The function block for the actuators also contains inputs for a safe switch-on and protective switch-off.

The function block for controlling actuators such as pumps and boiler includes counting motor hours and the number of switching. This monitoring is set up via ADD instructions and writing to static variables that are stored in the created Param structure. These two variables are set as retain variables. This means that the PLC remembers its value even after a power failure. The motor hours are calculated from the time of the previous cycle, so there is no need to use a timer.

C. Visualization

The process visualization can be designed directly in the TIA Portal environment. The seven-inch TP700 Comfort touch panel is used for display. In designing the visualization, emphasis was placed on simplicity and clarity so that students could easily navigate it.

The main screen of the operator panel shows the station model with the individual display fields. There is a display of the temperatures in the primary and secondary circuits of the station, both at the input and output of the heat exchanger. It also displays the opening of the control valves indicated in percentages, where 0% means a fully closed valve - i.e. no water flows from the boiler or to the appliance, and 100% means a fully open valve. Values from pulse flow meters are also shown, to monitor the amount of water flowing, and from this the subsequent calculation of the flow rate. As a last numerical indication, the pressure in the primary circuit is shown.

The indication of the operation of the actuators, both pumps, and boiler, is shown in green. If the elements are active, the icons of the elements are highlighted in green. Otherwise, when they are inactive, they are colored red.

The desired temperature of the thermometer T3 can be entered on the operator panel as a temperature input for heating.

At the bottom of the panel, there is a bar with buttons to select automatic or manual mode and switch between screens.

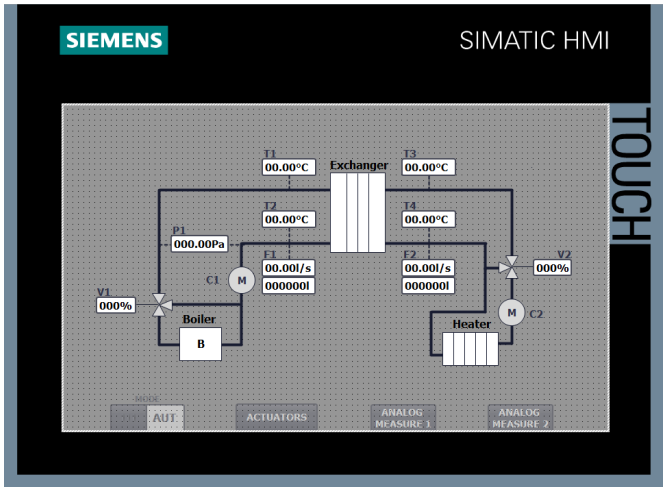


Fig. 4. The appearance of system visualization

The actuator button switches the screen to the actuator overview, where individual elements can be individually switched on or off. This control only works in manual mode. The motor hours and the number of starts of each actuator are also shown.

When the analog measurement button is pressed, the options for setting user limits for comparing and evaluating conditions such as overheating and so on are displayed. There is also a display box on the analog measurement screen where you can enter simulation values for the thermometers.

V. Conclusion

The result of this work is a functional model of a heat exchanger station, which will be used in teaching courses focused on industrial technology. The students will be able to experience the functionality of individual actuators and sensors on a real device.

The work itself is devoted to the revival of an already-made system, for which it was necessary to make new wiring and change the control system to the standard Siemens system used in the laboratories of the Institute of Industrial Automation.

Thanks to detailed documentation and verification of the functionality of the entire system, it is possible to start using the system fully. It will be possible to identify the system and create an assignment and instructions for a laboratory task to be included in the tutorial.

Acknowledgment

The completion of this paper was made possible by the grant No. FEKT-S-23-8451 - ”Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence, automation, and measurement” financially supported by the Internal science fund of Brno University of Technology.

References

[1] M. Baláš, Kotle a výměníky tepla, Fourth edition. Brno: Academic publishing house CERM, 2022
[2] “The principle of heat exchange”, Alfalaval, 0c2015-2023. .
[3] “ALPHA+ 25-40 180”, Grundfos. .
[4] “Siemens servo drive”, Domintex, 0c2011. .
[5] “The fastest control system for automation”, Siemens, 0c1996 - 2023. .

Design of Asset Administration Shell testbed for Industry 4.0

Tomáš Pečinka
Department of Control and Instrumentation
FEEC, Brno University of Technology
Brno, Czech Republic
230154@vut.cz

Petr Marcoň
Department of Theoretical and Experimental Electrical Engineering
FEEC, Brno University of Technology
Brno, Czech Republic
marcon@vut.cz

***Abstract*—This paper deals with an decentralized industrial control using Asset Administration Shell (AAS). It introduces reader to the concepts of Industry 4.0 and explains the structure and methods of AAS development using available software tools. This paper also illustrate decentralized control in practice with the use of a demonstration testbed. The testbed consists of an online eshop and a group of 3D printing units whose AAS are implemented in the programming language Python.**

***Index Terms*—AAS, Asset Administration Shell, Decentralized control, Industry 4.0, MQTT, 3D printing**

I. Introduction

In recent years, an increasing number of companies have invested in the modernization of their manufacturing according to the Fourth Industrial Revolution. This is leading to new processes and faster and more efficient manufacturing. An increasing number of businesses are also demanding fully automated production lines without the need for human intervention. One way to meet this demand is to establish a new, decentralised network with individual components making decisions independently of each other and communicating across all layers of the manufacturing hierarchy. However, this requires a suitable "digital shell", that stores data about its own components and provides the decision-making and communication algorithm. One of the possible versions of this digital shell is an Asset Administration Shell (AAS). The AAS is a hybrid of the digital twin that has precisely defined internal structure in which stores digital data and its active part provides all the algorithms needed for decision making and communication in a decentralized network.

The aim of this article is to familiarize the reader with these digital shells and their practical use in the field of Industry 4.0. In the second and third chapter, the necessary theoretical background for the orientation in Industry 4.0 and further information on AAS are presented. The introductory chapters are followed by a demonstration testbed, one of the possible ways to implement AAS into an production line.

II. Industry 4.0

Industry 4.0 can be defined by the following principles:

- Production processes are interconnected within the entire corporate pyramid by vertically and horizontally integrated information systems.
- Control and interaction with industrial elements remotely via internet connectivity and inter-connectivity between devices.
- Separate production units are replaced by automated production lines.
- Virtual prototypes and designs are preferred over physical ones.
- Flexible manufacturing where several companies are involved in the product development process.
- Machines making autonomous decisions in real time based on sensor data and requests from the control system, increasing production efficiency.
- Logistics optimization using algorithms that build an efficient route and cater the required materials.
- Efficient use of natural resources and the recycling of materials can be taken into account in the design of the product and, if necessary, revised before manufacturing begins. [1]

III. Asset Administration Shell (AAS)

A. Introduction to the concept of AAS

The Asset Administration Shell (AAS) is one of the key components of Industry 4.0, that provides a standardized representation of a digital or physical asset. It is a description that is defined according to strict rules and is not dependent on technology or on the particular device. The motivation for developing AAS was the fact that currently there is no existing standard defining the description in the field of industrial automation, that is independent of both manufacturer and technology. [3] The AAS must be composed of a passive part that contains information about the assembly and the AAS. Passive part may be extended by an active part allowing communication with other AAS. Detail of the AAS with both passive and active part is shown in Fig. 1

B. Identifiers

Identifiers are a necessary and effective tool to uniquely identify Industry 4.0 elements. Identification is required at least for [2]:

- Asset Administration Shell
- Asset (identifier specified in the encapsulating AAS)
- Submodel (both for template and instance)

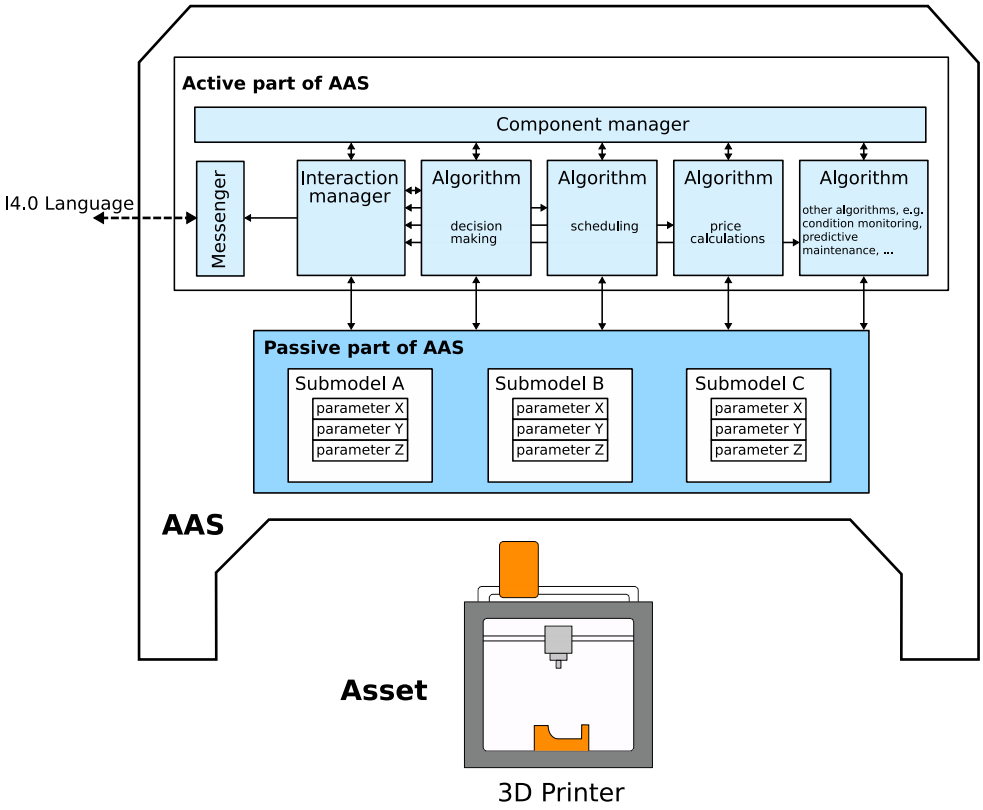


Fig. 1. Asset Administration Shell designed for 3D printer (inspired by [4])

There are three types of identifiers available for identification:

- International registration data identifier (IRDI) - defined by ISO29002-5, ISO IEC 6523 and ISO IEC 11179-6
- Internationalized resource identifier (IRI) - defined by RFC 39862
- Custom identifiers such as Universal unique identifier (UUID) or Globally unique identifier (GUID), which are defined by the manufacturer according to their internal needs.

Identification serves not only as a means to uniquely distinguish elements in the AAS, but also to establish links (references) between templates and instances of submodels, metamodels and AAS elements. [2]

C. Types of AAS

We currently classify AAS into three categories based on the activity with which they interact in the Industry 4.0 network.

1) Type 1 AAS (passive): Type 1 AAS, sometimes referred to AAS as a file, contains only the passive part, which consists of serialized JSON or XML files containing static information about the asset they encapsulate. These files are shared across companies to exchange information about the AAS. [3]

2) Type 2 AAS (reactive): A reactive AAS is an AAS containing a passive part that is hosted on a server as a runtime instance containing a communication API. A reactive AAS contains static data stored in the passive part of the AAS, but in contrast to a passive AAS, it is able to interact with other reactive and proactive AAS. Using reactive AAS, it is possible to display sensor data, product availability, production efficiency etc. in real time. [3]

3) Type 3 AAS (proactive): The proactive AAS, referred to only as active in some of the literature, extends the passive AAS to include active behavior, which consists of the ability to communicate, decide and negotiate independently. [3] [4]

D. Passive part of AAS

The passive part of the AAS contains all the information about the asset it encapsulates. This is information from the assett manufacturer, information from the operator, technical documentation, service manuals, etc. The requirements for the creation of the passive part are strictly given and can be found at [5].

The internal structure of the passive part of the AAS consists of a header and a body. The header contains the identifier of the AAS itself, but also of the asset it encapsulates. According to the German association ZVEI, the header should contain the required security level and a reference to the body of the AAS, where other security attributes (security keys, authentication, etc.) are listed. Using the identifiers listed in the header, other entities can access information about the type and instance of the AAS. [6]

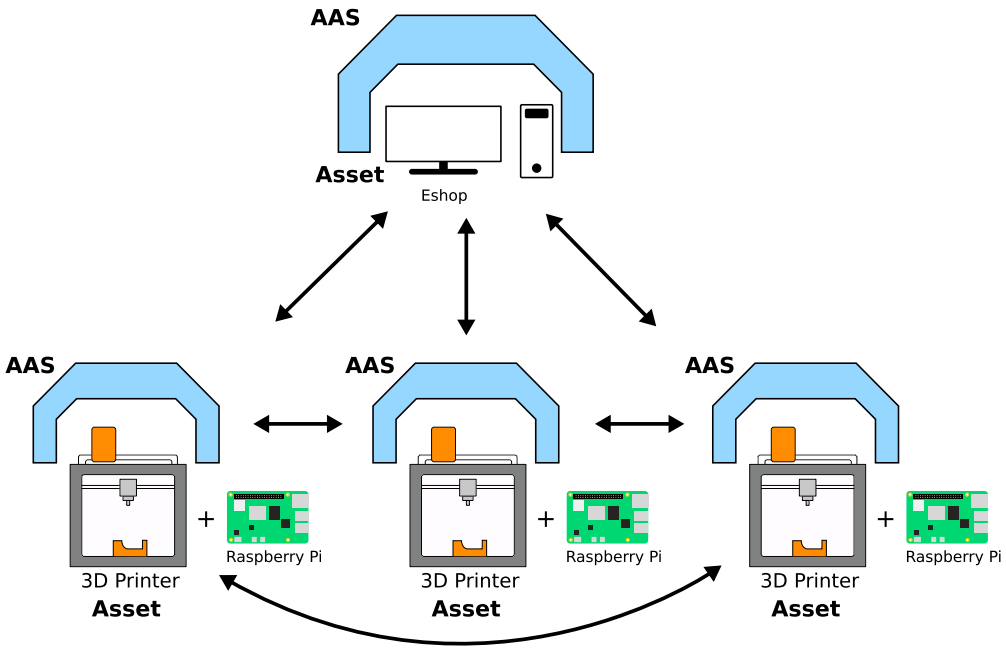


Fig. 2. Testbed design

The body of the passive part is made of submodels. Submodels are structured blocks that describe different aspects of the AAS.Each submodel refers to a clearly defined domain. The most common submodels describe the technology that the asset possesses, security level, energy efficiency, identification of AAS and assets, and condition monitoring. [7]

A submodel is further broken down into individual elements, e.g., an electric motor with submodel called ”operating parameters” will contain the elements ”speed rotation” and ”torque”. As already mentioned, each submodel must contain its own identifier. The use of identifiers for submodel elements is recommended. [2]

E. Active part of AAS

The active part of the AAS provides the ability to communicate with other AAS, make decisions, negotiate, plan, perform predictive maintenance and other activities that make the AAS capable of real-time responsiveness to changes and streamline production. The active part divides the AAS into two basic types, service requesters and service providers. A service requester can be, for example, a semi-finished product that requests information from machines about whether they meet certain parameters for its processing. However, it can also be the machine itself that requests the operator to provide maintenance or to replace a work tool. The service provider may be the same machine that provides finished products to other machines. From the above example, it can be seen that the division between service provider and service requester can vary depending on who the AAS is currently communicating with. One AAS could be both a service provider and a requester. [4]

IV. Testbed

This chapter focuses on the design of a demonstration testbed using AAS. In the following subsections are described different parts of the demonstration task and the design of the practical implementation.

A. Design

The aim of the testbed is to create a number of sample AAS that encapsulates physical or digital elements and interact with each other in a suitable way. The AAS thus created will form a universal production unit that will be functional without operator intervention.

The testbed design can be seen in the Fig. 2. Testbed will consist of one ordering server (customer eshop) and several 3D printers. The customer will have access to the web interface of the e-shop and will be offered multiple 3D printing options with different models, colors, dimensions, materials, etc. After configuring the order, the customer confirms the print. Here the negotiation process starts between the AAS implemented on the backend of the e-shop and the AAS running on the Raspberry pi and connected to the 3D printers.

B. Communication between AAS

For the purpose of testbed the MQ Telemetry Transport (MQTT) protocol was chosen, mainly due to the availability of development libraries and straightforward implementation using the paho-mqtt 1.6.1 library written in Python. MQTT was invented in 1999 as a protocol with minimal bandwidth requirements and simple implementation. The protocol was originally used in proprietary embedded systems, but today finds frequent use for Internet of Things (IoT) devices. The protocol has been upgraded several times since 1999 and currently a standardized ISO/IEC 20922:2016 version of MQTT 5

is available. The MQTT protocol is based on the publisher-subscriber principle, i.e. publishing and subscribing data under certain topics using a central element called a broker. Devices connected via MQTT can subscribe and share to an unlimited number of topics. [8]

C. Bidding process

The semantics of negotiation and the structure of transmitted messages are based on the I4.0 language, which is defined according to VDI/VDE 2193. Negotiation between provider and requester services is carried out using a specific interaction set, which is contained in the interaction manager and is known to both negotiating AAS. This is most often negotiated through a bidding process. The type and content of the messages exchanged between the participants in the bidding process are predefined according to whether they are service providers or service requesters. A typical bidding process takes place between one service applicant and one or usually several service providers. The communication is initiated by the service requester with the sending out of a service request. If the message is in the correct format, the service provider must decide whether to accept or reject the offer and send its response back to the requester. The requester evaluates the offer and sends a message to the service provider whether to accept or reject it. If the offer is accepted, the service provider shall initiate the requested action and inform the applicant of its success. [4]

D. Autonomous production changes

The most significant symbol of Industry 4.0 is definitely autonomous control, i.e. control of production without operator intervention. However, this also entails maintaining this type of control in the event of a sudden breakdown of a production line or component, material shortage or dynamic changes in products priority. Programming such a control system by hand would be enormously time and material consuming.

This problem could be solved by deploying production lines with AAS. The centrality of the control system, which is usually the Manufacturing Execution System (MES) nowadays, would be transferred to the individual AAS, spreading the computing power over several smaller units that would make their own decisions based on received data.

E. Hardware and software implementation

The hardware realization of the testbed consists in the creation of an industrial line that contains three 3D printers. Each printer is connected to a RaspberryPi Zero. The Raspberry has two activities, firstly operating the printer and preparing the models for the 3D printer, and secondly it runs the AAS which communicates with other AAS in same network and the printing software. This enables the printing to be started remotely and its status to be monitored at the same time.

The software part of the testbed is primarily written in the python programming language. The Eclipse BaSyx Python SDK library is used to create the passive part of the AAS. The active part of the AAS, i.e., the communication, negotiation and decision-making algorithms are also written in python and utilize the data from the passive part for their operations.

V. Conclusion

This work introduced a new approach in the field of industrial control using AAS. The deployment of AAS in current industry would significantly increase the efficiency of production and would be able to respond to sudden material shortages or production changes. The described implementation of the algorithms of the active part of the AAS was successfully implemented and tested on mutual communication and negotiation between the service requester and the service provider. Future work will be to incorporate the already created AAS into a real testbed, including operating 3D printers using microcontrollers, a customer ordering portal and monitoring production data.

Acknowledgment

The research was carried out under support of the general student development project being executed at Brno University of Technology.

References

[1] Federal Ministry for Economic Affairs and Climate Action of Germany, “What is Industrie 4.0 ?”, *Platform Industrie 4.0*, 2016.

[2] Plattform Industrie 4.0, “Details of the Asset Administration Shell - Part 1: The exchange of information between partners in the value chain of Industrie 4.0 (Version 3.0RC02)”, *Platform Industrie 4.0*, p. 235, May 2022.

[3] A. Belyaev, C. Block, B. Boss, C. Diedrich, P. Juhel, W. Hartmann, O. Hillermeier, N. Ondracek, S. Pfeifer, F. Scherenschlich, and J. Schmelter, “Modelling the Semantics of Data of an Asset Administration Shell with Elements of ECLASS”, p. 54, Jun. 2021.

[4] A. Belyaev and C. Diedrich, “Specification ”Demonstrator I4.0-Language” v3.0”, p. 38, Jul. 2019.

[5] ZVEI - Zentralverband Elektrotechnik- und Elektroindustrie e.V. Automation Division, “ZVEI Recommendation 2020.01 The Digital Nameplate: Consistent, sustainable, future-proof, networked”, p. 8, Oct. 2020.

[6] Plattform Industrie 4.0, “Industrie 4.0 Plug-and-Produce for Adaptable Factories: Example Use Case Definition, Models, and Implementation”, p. 69, Jun. 2017.

[7] Ministry of Economy and Finances, Alliance Industrie du Futur, Federal Ministry for Economic Affairs and Energy (BMWi), Plattform Industrie 4.0, and Ministero dello Sviluppo Economico, “Structure of the Administration Shell: Trilateral perspectives from France, Italy and Germany”, p. 64, Mar. 2018.

[8] “MQTT Essentials: The Ultimate Guide to MQTT for Beginners and Experts”, *HiveMQ*, 01-Oct.-2019.

Graphical user interface for embedded applications

1st Peter Bránecký
Department of Control and Instrumentation
BUT - FEEC
Brno, Czech Republic
xbrane01@vutbr.cz

2nd Petr Petyovský
Department of Control and Instrumentation
BUT - FEEC
Brno, Czech Republic
petyovsky@vut.cz

3rd Matej Pončák
Department of Control and Instrumentation
BUT - FEEC
Brno, Czech Republic
xponca00@vutbr.cz

***Abstract*—This work deals with the design and implementation of the graphical user interface for development board RPi Pico Kit [1]. This development board is based on the Raspberry Pi Pico board. Furthermore, this work deals with the user inputs from mouse and keyboard through USB interface. Resulting GUI framework allows programmer to create applications with a graphical user interface, which can be controlled by mouse and keyboard.**

***Index Terms*—Raspberry Pi Pico, education kit RPi Pico Kit, RP2040, embedded systems, Graphic User Interface, keyboard, mouse**

I. Introduction

Most embedded systems don't meet hardware specifications for utilizing general libraries for work with the graphical user interfaces (GUI). Some of them cannot support video output whatsoever. This means that, we cannot find universal library that supports creation of a GUI for embedded systems. RPi Pico Kit with Raspberry Pi Pico onboard supports generation of video output within its hardware capabilities. However, as mentioned before, we don't have access to a tool that would provide us with the capabilities to create applications with a GUI. This work aims to create tolling for programmer, that will allow him to create these types of applications.

In this article, we will look at solution for generating video output on Raspberry Pi Pico, handling user input from mouse and keyboard through USB interface and implementation of framework that provides ways to create GUI applications for RPi Pico Kit.

II. Developement board

The RPi Pico Kit is development board created as an innovation of the laboratory exercises for course Embedded Systems [1]. In this section we will mention some capabilities of RPi Pico Kit, which provides Raspberry Pi Pico board access to various peripherals.

A. Raspberry Pi Pico

The Rapberry Pi Pico family is range of tiny, yet fast and multipurpose boards build using the RP2040 microcontroller chip. RP2040 was created by the Raspberry Pi Foundation itself. This 32-bit chip is built on the Arm Cortex-M0+ processor, which has very low consumption for speed that it provides. This dual-core processor with a flexible core speed up to 133 MHz has access to a 256 kB of internal SRAM memory and supports a wide range of flexible I/O options, such as I2C, SPI and uniquely programmable I/O. Wireless version, Rapberry Pi Pico W, features the 24 GHz 802.11n WLAN module. This makes it a perfect solution for IoT applications with a price as low as 8€.

For our purposes is important that this board comes with programmable I/O and adequate core speed up to 133 MHz. We will find out that 256 kB of SRAM is our main bottleneck when creating applications with GUI.

B. RPi Pico Kit

The RPi Pico Kit provides developer of the Raspberry Pi Pico board, with ability to use various peripherals which this kit provides and extends methods for powering the board. The kit hosts second RP2040 chip, which is used for uploading programs to Raspberry Pi Pico board and provides programmer ability to debug applications.

In figure 1 we can see all the peripherals and accessories that RPi Pico Kit provides. For our work is important that, there is VGA connector fitted on the kit. There is also a DVI connector, but we are not currently using it in our work. The VGA interface is connected to board on GPIO11 - GPIO20 pins. First 8 pins are used for colour representation, this means that 8-bit colour depth is being used. Last 2 pins are used for synchronization signals VSYNC and HSYNC. The VGA interface is connected to I2C bus which allows reading extended display information data of connected monitor.

III. Video signal generation

For generating suitable VGA signal, we have chosen the PicoQVGA library [3]. This library support resolution of 320x240 pixels and 8-bit colour depth with the R3G3B2 format. The library uses one of two programmable I/O (PIO)

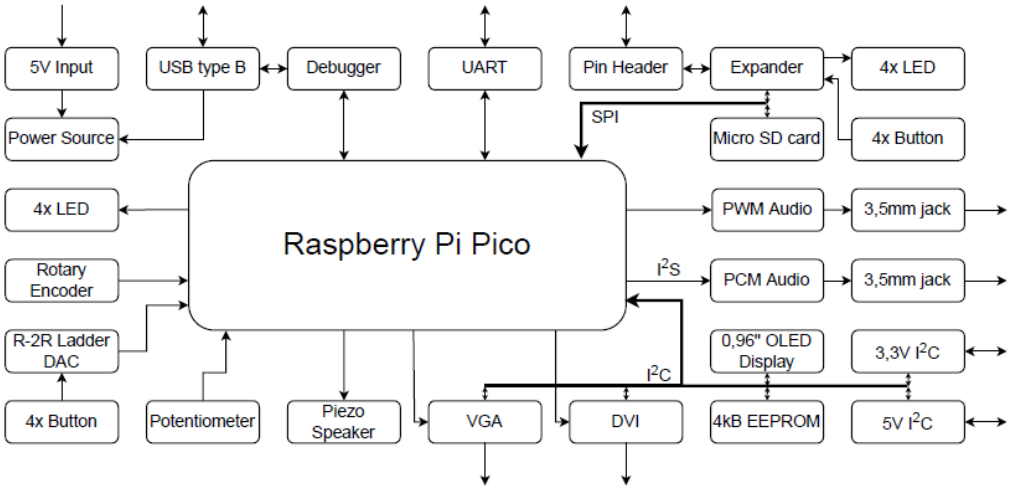


Fig. 1. RPi Pico Kit HW diagram [1]

blocks and two direct memory access (DMA) channels out of twelve available. The DMA channels are used for transferring data from the frame buffer, which is located inside RAM, to the PIO block. This allows data transfer between the program memory and the PIO block without direct processor involvement, thus increasing program efficiency.

The PIO block is programmed similarly to a processor. However, it has its own small instruction set. Main purpose of the PIO block is that it allows for creation of additional communication interfaces. Each PIO block has a program memory off 32 instructions and 4 state machines. These state machines share program memory, but they are functioning independent of each other. Every state machine has access to two FIFO queues. First is used for input data, second for output data. The PIO block can create interrupt request IRQ0 and IRQ1.

State machines inside the PIO block, which is used by PicoQVGA library, are responsible for generating VSYNC, HSYNC signals and outputting pixel colour data to GPIO pins. These GPIO pins are then connected to monitor through VGA interface on the kit. The DMA module creates IRQ0 interrupt request, after every finished scanline on monitor.

IV. Handling user input

To handle user input from mouse and keyboard we have chosen to use the TinyUSB library [4]. The TinyUSB library is free, open source and supports wide range of microcontrollers. It was designed for usage with microcontrollers, so its thread-safe with no dynamic memory allocations. The library allows microcontroller to act as USB Host or USB Device. In our case we need to microcontroller to act as USB Host. In this mode we can handle Human Interface Devices (HID), Mass Storage Class (MSC) devices or USB hub with one level support.

A. Data handling

The TinyUSB library will process all incoming data from our devices into form of a report. Each device category has its own report type which developer then uses to handle incoming information. From mouse, we can receive information of 5 pressed keys and information about relative mouse movements. From keyboard, we can receive scancodes from 6 pressed keys at once. These scancodes need to be translated into their corresponding ASCII representation. This is done using lookup table inside our code.

V. Framework for GUI creation

In this section we will look at implementation of RP GUI framework, which can be used to create applications with GUI for the RPi Pico Kit.

A. Visual elements

Drawing of the visual elements on screen is handled through interface created for usage with PicoQVGA library. This interface substitutes native data types for data structures. Using the interface prevents making logical mistakes and typos while working with this library.

There are currently 10 visual elements accessible to developer:

- Button
- CheckBox
- Label
- Line
- PolyLine
- ProgressBar
- RadioButton
- Rectangle
- Stepper
- Switch

Developer can define width and height of each element type and its displayed colour. For clickable elements, such as a button, developer can assign function which is executed when the element is clicked.

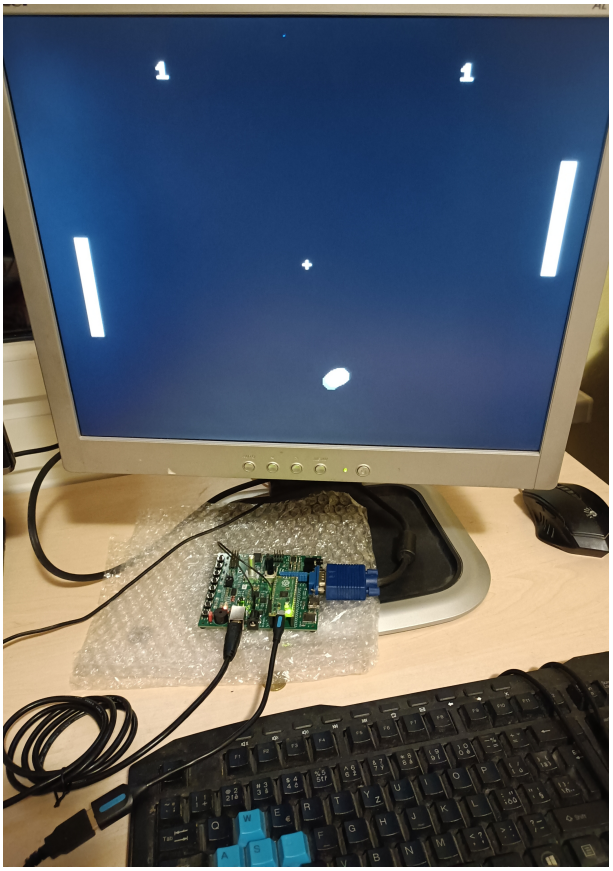


Fig. 2. RPi Pico Kit setup with example of use 1



Fig. 3. RPi Pico Kit setup with example of use 2

B. Layouts

All created visual elements needs to be added to some kind of layout. Currently available layouts are StackLayout and AbslouteLayout. In AbsoluteLayout developer defines exact position on screen for each created visual element. In StackLayout, position of element is determined automatically based on root position of given StackLayout. Elements in StackLayout are then drawn from top to bottom, in order which there were added inside layout. Layouts then need to be grouped inside pages. Pages allow adding of multiple layouts into single page. Developer can then add multiple pages to core part of framework and choose which page will be then drawn onto the screen.

C. Data bindings

The data displayed by visual elements can be stored directly inside the element or, it can be stored as a reference to data elsewhere in code. This means that, we only have to change the value referenced by the element instead of changing it directly thought the element itself. So, we only need to keep track of the referenced data instead of the whole element.

D. Core of framework

The core part of the framework consists of functions which allows developer to group crated pages and then draw one selected page to screen. This part of the framework checks data created by mouse and then, creates events based on the keys being pressed or mouse being moved. When the mouse moves, cursor position on the screen is adjusted accordingly. When a mouse key is pressed, framework fires event that click occurred. Three types of events are supported: on mouse pressed, on mouse released, on mouse clicked. After event happening, the framework will find out if cursor is indeed above some clickable element, e.g. Button, and call the handle function assigned to the element.

Framework can draw elements on screen in two modes: CORE0 and CORE1. In CORE0 mode, all draw functions are called from core0 of processor. This means that frame buffer is filled with data from draw functions running on core0. Similarly, in CORE1 mode, frame buffer is filled with data from draw functions which are running on core1 of processor.

Developer can with use of the mouse and keyboard run arbitrary functions defined in his code and choose which core to use for filling the frame buffer. Framework waits for VSYNC signal automatically when drawing pages.

VI. Example of use

As an example of use, we have created a simple game for two players called Pong. In this game, players are trying to score goals against each other using their pedals. Players can use keyboard, connected to the Raspberry, to move pedals across the screen, thus play the game. All visual elements of

the game; pedals, ball and counters, are created and managed by our framework. Figure 2 shows setup of RPi Pico Kit, with connected keyboard, and running Pong game on the VGA monitor.

Another example is shown at the figure 3. In this example, the framework is used to control four LEDs on the kit, display current value from potentiometer and counter which is being incremented by rotary encoder. The switch element is used to change the state of each individual LED. The progress bar element is used to show the value from AD converter, which can be changed by the potentiometer on the kit. The label element is used to show current value of counter. The switch elements and label are using the data binding feature to access and update data defined in data structure inside our code.

VII. Conclusion

The work deals with implementation of a tooling that allows developer to create applications with a graphical user interface (GUI) for the Rpi Pico Kit development board or the Raspberry Pi Pico itself. Created solution builds on the PicoQVGA library which is used for generating VGA video output for monitors and the TinyUSB library which handles user input from a mouse and keyboard.

Developer can then create applications with a GUI using provided RPGUI framework. By using this framework, developer enables end user with the option to control flow of running application on Raspberry. For this the user can use keyboard with mouse connected to Raspberry via USB hub.

Goals for the future of this work are: implementing SD card support for logging and data storing, implementing support for sprites which can be loaded from SD card, testing and performance improvements of the whole framework and finally, implementing more visual elements such as slider or entry field into framework.

Acknowledgment

The completion of this paper was made possible by the grant No. FEKT-S-23-8451 - ”Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence, automation and measurement” financially supported by the Internal science fund of Brno University of Technology.

References

[1] M. Ponč ák, “Innovation of the laboratory exercises for course Embedded Systems,” VUT FEKT Brno: 2022 [cit. 2. 3. 2023]. Available at: http://hdl.handle.net/11012/204867.

[2] Raspberry Pi, ”Raspnerry Pi Pico product brief”, Rev. 7/2022. Available at: https://datasheets.raspberrypi.com/pico/pico-product-brief.pdf.

[3] M. Němeček, “PicoQVGA - minimalist display QVGA for Raspberry Pico.” Available at: http://www.breatharian.eu/hw/picoqvga/index_en.html.

[4] “TinyUSB: open-source cross-platform USB Host/Device stack for embedded systems,” Available at: https://docs.tinyusb.org/en/latest/.

Greenhouse manipulator

1st Given Martin Sukdolák
dept. of control and instrumentation
Brno University of Technology
Brno, Czech Republic
xsukdo00@vut.cz

2nd Ing. Tomáš Benešl
dept. of control and instrumentation
Brno University of Technology
Brno, Czech Republic
xbenes23@vut.cz

***Abstract*—This paper focuses on the automation of the greenhouse process using a SCARA manipulator in order to make the greenhouse work more efficient with minimum water consumption. The article discusses the actual implementation of the manipulator in the greenhouse and discusses the advantages and disadvantages of the use.**

***Index Terms*—Greenhouse, manipulator, SCARA, water consumption**

I. Introduction

As the population continues to grow, the global consumption of food, namely vegetables, fruit and livestock are growing too. The grow has resulted in a expansion of the agricultural economy and It has followed by extension of use more water. Agriculture is the most water demanding sector from all. One liter of a water is needed to grow one kCals of food. The increase of food demand results to higher water consumption that cause to lack of drink water to survive in some parts of the countries and it forces the people to migrate to the other parts of country where is still enough drinking water. [1]

Due to development in electrification and automatizing has appeared idea to make the farming process more efficient. The first efforts to implement "smart farming" were already in the 1970-1980s. Today, improvement of automatizing has brought the opportunity to develop many flexible farming devices that reduce unwanted waste. It can reduce environmental impact and increase food sustainability. Robotic platforms on the field can also use the collected data about soil, seeds, crops and use them for effective irrigation in the future.The main use of smart farming is on repetitive tasks, where automated robots are more efficient than humans or large machines. [2]

II. Farming robots

A few robots are now developed specifically for agriculture. The most robots are adopted from other industries and adapted to a specific purpose. These robots can be divided into two main groups. The first group of robots can do only a specific task and the second group is robots that can use different kind of tools and they can work in wide range of different jobs. The main problem of these systems is that the most robots are not still enough robust to withstand in real-world conditions such as dust, rain, fog and high/low temperatures. For example, the most manipulators can not withstand high humidity in the greenhouses. [2]

A. Market Comparison

The main inspiration for the project is the startup project Farmbot. However, the project can not be placed in a greenhouse as the whole system rides above the flower seedbed. Another disadvantage of this design is that the plants are accessed from the top. It results to poor access to the plants and next to an inability to do tasks such as weeding, fruit harvesting or soil moisture measurement.

On the other hand, Root-ai and Certhon Harvest Robots are manipulators that are directly developed to work in the greenhouse. Certhon Harvest Robots have taken already developed 5-axis robotic arm and they have redesigned it. It is able to harvest tomatoes in a hydroponic greenhouse. Root-ai is a manipulator heavily developed for hydroponic greenhouse tomato harvesting, but its manipulator is a SCARA type. Both of these systems have machine vision to evaluate the ripeness of the tomatoes. The disadvantage of these systems is that they are developed for large industrial greenhouses and they are not so accessible to the general user.

This paper seeks to combine the benefits of all these systems. From the Farmbot project uses the accessibility for the general user, additionally, from the Certhon Harvest Robots project takes an existing manipulator and redesigning it for a specific activity and next from the Root-ai project uses the most advantageous type of manipulator type SCARA.

III. Manipulator

When I designed it, I considered the possibility of the manipulator running over the entire bed and the possibility of the manipulator running in the center aisle. Since the manipulator was designed for a greenhouse, the option of having the manipulator ride over the entire bed was not an option due to the fact, that it would be very difficult to install it in a greenhouse. Therefore, i chosen the option where the manipulator rides in the middle aisle, because the manipulator has better access to the plants and It can do different kind of actions such as picking fruit or removing weeds. Furthermore, i decided to use the SCARA manipulator because it is easier to control than other systems and it can be printed on a 3D printer and then easily assembled. The track on which the manipulator moves through the greenhouse is designed, if it is necessary, it does not have to be completely parallel to compensate for imperfections in the greenhouse.

IV. Digital twin

A digital twin is a virtual representation of a physical system that it makes possibility to test products and systems without the threat of damage or destruction. This representation is the cornerstone of Industry 4.0 as it links the different manufacturing sectors together. The main advantages of the digital twin are:

- speed therefore typing and re-designing the product
- less costly due to the predominantly virtual creation
- it is possible to predict errors and react adequately to them
- reduced waste in digital testing and prototyping
- a more effective and illustrative training tool [3]

In the implementation of this project, firstly i created a digital twin which i will use to test the complete functionality of the system.

A. Siemens NX

I used the Siemens NX program to model the digital twin. This program allows me to create the model using parametric modeling and then control it by using moving constraints. After of setting the individual materials for different parts, it is possible to monitor the stresses on the individual components. Once the model is debugged, the entire system can be connected to the Programmable Logic Controller using communication protocols so the functionality of the program can be tested afterwards. This also allows parallel execution of the program on the real system and in the model simulation at the same time, thus achieving the creation of a digital twin.

B. SCARA manipulator

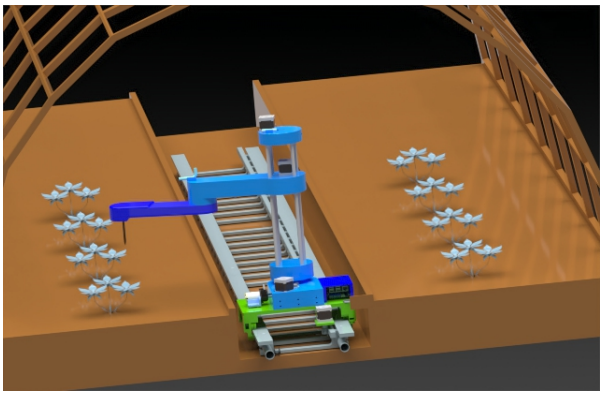


Fig. 1. SCARA manipulator

The base of the manipulator is a model of SCARA manipulator taken from [4]. I has subsequently modified this model to meet the requirements for work in the greenhouse environment. The main modifications consist in redesign of the end manipulator arm and in modeling one level of freedom in the direction of the greenhouse aisle. Most of the parts were modeled to be produced by 3D printing, which i reduced the overall cost and increased the availability of this manipulator.

C. Substitution

The base ensures linear movement of the robot along the centre of the greenhouse, thanks to a toothed comb that serves as a guide rail mounted on the rails. The linear motion is converted to rotary motion by means of a gear wheel attached to the manipulator and rotated by a stepper motor. The linear guide was inspired by HepcoMotion's MHD model. The linear guide consists 4 pieces, where each piece contains 2 bearings to prevent up and down movement and one bearing preventing sideways movement. I chose this principle because its individual parts can be printed using 3D printing method.

D. Arms

To increase the accuracy of the arms, the stepper motors are geared with pulleys. Smooth rotation is ensured by thrust bearings which assist in the rotation of the manipulator. At the end of the arm there is an ultrasonic distance sensor which monitors the distance from the ground. It has also modularity option by the replace of robot head based on the desired action.

V. Electronics

The main driver part of this project is a Raspberry pi 4 connected to a GPIO expander with SPI communication which control DRV8825 modules. These modules are then used to control the actuators. The manipulator has one NEMA 17 stepper motor for each degree of freedom. At the end of the manipulator is a moisture sensor which allows the robot to determine if the plant needs to be watered and how much water it needs. A peristaltic pump is used to pump the water, since this pump can maintain a constant pressure in the pipe and it can draw water even from a clogged pipe. The whole system is powered by 24 volts. The logic part uses 3.3 volt logic.

A. Sensors

The main sensor of the whole manipulator is the soil humidity sensor located in the robot head. There are two main principles of soil humidity measurement. The first is based on the change in conductivity of the humidity in the soil. The disadvantage of this sensor lie in its exposed surface and as a result, corrosion and change in the magnitude of the measurand occurs over time. For our purpose, I have chosen a sensor based on the capacitive principle, where the capacitance changes are based on the change in dielectric. Since the soil has a different permittivity when it's wet or dry, i can determine the difference in capacitance and allow only the necessary amount of water be poured into the seedbed. Another sensor is an ultrasonic sensor which sends out ultrasonic waves that bounce off the soil and determines the distance based on the phase difference of the send and reflected wave. On the board there are two thermistors which change the resistance based on the temperature change in the greenhouse and heating aspects of electronic components which is then measured. The last sensors are microswitches at the ends of all axes which will respond to the robot reaching a specific position.

VI. Software

The software itself was designed and programmed in Codesys which uses an library extension to work with Raspberry pi. This way i created a soft PLC from the Rasberry pi and then it was programing as any other PLCs. The general problem of this solution was that Raspberry pi did not work in real-time but in this application it is not so necessary to work real-time. Particularly the problem was solved by allocating one processor core for real-time operations which is possible only from Raspberry pi 3 model B series. I implemented HMI in the system to simplify the communication between human and manipulator. It is necessary that the Raspberry pi is connected to the Internet which give us the posibility to display HMI interface using a web browser on any device connected to the network. Since i was using a program developed for the PLC it was possible to load this program into our digital twin model at the same time. The disadvantages of this solution were mainly that the Codesys extension for Raspberry pi reserves 12 pins for I/O devices and the remaining pins are used for communications such as SPI, I2C or 1-wire. For this reason I was forced to use an expander with SPI communication which creates an additional 16 I/O pins.

A. Testing programme

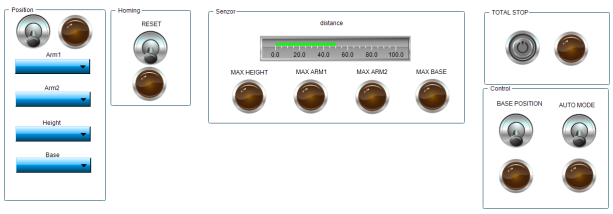


Fig. 2. HMI

To test the digital twin and the real manipulator, i created a PLC program in Codesys using structured text. I also created an HMI for this program. Turning on the homing function puts the robot in a position where all limit switches are closed and sets to zero position. Without the first calibration, the robot can not be operated due to the risk of impact at the ends of the axes. In the Position function, the user sets the specific position of all robot axes. This function can be used, for example, for service tasks or for manual manipulation of the robot. By activating the Base Position function, the robot moves to the position where all axes are least mechanically stressed. When the Auto mode function is switched on, the robot executes a predefined program to water all plants in turn. If a problem occurs, the TOTAL STOP button can be pressed to stop all motors and the robot can not be manipulated until the button is pressed again. This environment also contains indicator lights that show the limit switch is on and a slider that shows the distance of the robot head from the surface.

VII. Realization

The overall implementation can be divided into several steps. The first step was 3D print of all the parts and ordering all the necessary parts. The second step was building the entire robot and creating the wiring. In the last step, I implemented the program and tweaked the bugs.

A. 3D printing

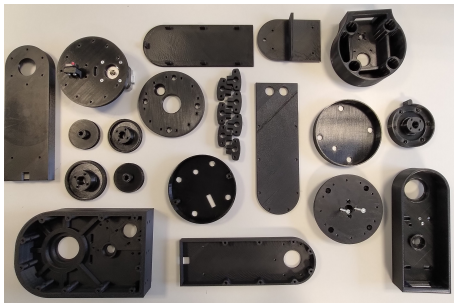


Fig. 3. 3D print

3D print enables fast prototyping of designed parts and the possibility to test them in real conditions. The parts in Figure 3 were created by FDM 3D printing. The material used to print this model was mainly PET-G and ASA material. I chose these materials for their mechanical and heat resistance.

B. Manipulator assembly

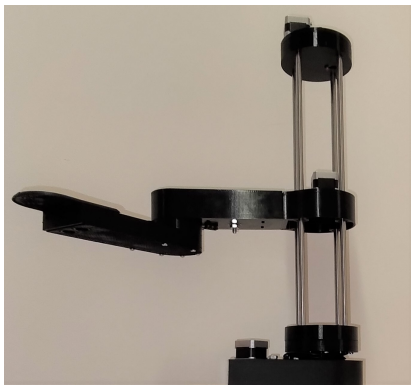


Fig. 4. assembly

From the printed parts, bearings, smoothed rods, belt and bolt with nuts i assembled the resulting SCARA type manipulator. For initial testing i did not connect the manipulator to the base because it would be difficult to verify proper functionality on the manipulator. So i connected 3 stepper motors to the expander and 3 end microswitches to the raspberry and verified the correct functioning of the main part of the manipulator. I then connected the SCARA manipulator to the sliding base and tested the overall functionality of the system. The current model has a arm reach of 500mm and the travel area can be adjusted based on the size of the greenhouse, and curves can be created to some extent if required.It is necessary to have a 24 volt power supply and an 8mm wide flexible hose stretched along the rails to move water to the manipulator head.

FUTURE PLANS

The current version of the manipulator is used only for watering, but already contains preparations for extensions such as an interchangeable head that could plant new crops or to pull weeds. The Codesys extension for the Raspberry pi also allows the connection of a camera, which makes it possible to create a machine vision and to harvest ripened crops.

ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-23-8451 - ”Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence, automation and measurement” financially supported by the Internal science fund of Brno University of Technology.

CONCLUSION

The main goal of this work was to create a robotic manipulator which will irrigate plants with minimal water, based on the humidity in the greenhouse. The result is a SCARA manipulator that drives in the aisle of the greenhouse and decides how much water to add to a given plant based on information from a humidity sensor. The whole manipulator was designed to keep the overall cost as low as possible and thus be affordable to the public. This would result in improved food quality in the greenhouse with minimized pesticide and human labor.

The main advantages over a conventional watering system are that the manipulator can be modularly adapted to pull weeds, remove pests or remove remaining fruit. It is also possible to program the manipulator according to the exact requirements of the plant and thus have different plants side by side in the same greenhouse. All these operations can be monitored virtually without physical present at the greenhouse.

REFERENCES

[1] Molajou, Amir et al. “A new paradigm of water, food, and energy nexus”. Environmental Science and Pollution Research n. pag. 1 Mar. 2023.

[2] Duckett, T.; andBruce Grieve, S. P.; Blackmore, S.: Agricultural Robotics: The Future of Robotic Agriculture. In UK-RAS, UK-RAS: UK-RAS White papers, 2018, ISSN 2398-4414, str. 36. Available from: https://arxiv.org/ftp/arxiv/papers/1806/1806.06762.pd

[3] Singh, M.; Fuenmayor, E.; Hinchy, E.P.; Qiao, Y.; Murray, N.; Devine, D. Digital Twin: Origin to Future. Appl. Syst. Innov. 2021, 4, 36. https://doi.org/10.3390/asi4020036

[4] Dejan: SCARA Robot. 2020. Available from: https://howtomechatronics.com/projects/scara-robot-how-to-build-your-own-arduino-based-robot/

onsemi

jobscz@onsemi.com
www.kariera-onsemi.cz

Jsme úspěšným světovým výrobcem polovodičů s dlouhou tradicí a více než 2200 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 Si a SiC desek a polovodičových čipů**



Nabízíme:
- placené stáže a brigády
- stipendia
- vedení bakalářských /diplomových prací
- exkurze
- prohlídky pro veřejnost
- zajímavé pracovní pozice

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



Design Layout Engineer

„Pro mnoho lidí je práce povinnost, pro mne však práce pro onsemi znamená radost. Považuji za štěstí, že jsem se mohla stát členem skupiny neskutečně vzdělaných lidí a špičkách 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.“

Automatic Robot Path Planning Tool

Tomáš Rotrekl
FEEC Department of Radio Electronics
Brno University of Technology
Brno, Czechia
222862@vutbr.cz

Petr Kadlec
FEEC Department of Radio Electronics
Brno University of Technology
Brno, Czechia
kadlecp@vut.cz

***Abstract*—This paper deals with creation of mazes with polygonal obstacles, expanding those obstacles to find ways through them, describing the environment of the maze, and comparing different ways to solve those mazes using the Dijkstra algorithm, A star algorithm, and the differential evolution with a variable number of dimensions algorithm (VND-DE). A large portion of this work is focused on pathfinding algorithms where we use Dijkstra and A star algorithms to find paths in previously constructed state graphs. We also try to find a set of optimal turning points placed between obstacles of a maze using the VND-DE algorithm. The aforementioned algorithms are then implemented in MATLAB tool to demonstrate and test path-finding algorithms.**

Keywords— Robot Path Planning, path finding algorithms, visibility graph, VND-DE algorithm, Dijkstra algorithm, A star algorithm

I. INTRODUCTION

In our ever-evolving world, the role of robotics is more and more significant. With that comes the need to find solutions to problems with robot movement in an unknown environment. This paper aims to review some of the concepts used for robot path planning. Some of those are: describing the maze, creating a state graph and using different algorithms to find an optimal way through the maze. A MATLAB tool developed to solve all these problems is described in this paper.

II. DESCRIBING THE ENVIRONMENT

First, we need to describe the environment we are trying to find a way through. The environment itself is named the workspace W. It is usually a 2- or 3-Dimensional Euclidean space in which we can describe all obstacles, their position, starting position of our robot, and the location of our target (goal). When we insert all the obstacles, we can refer to the workspace as configuration space C. This space can be divided into two groups: *free space* C_{free} in which the robot can be and *obstacle space* C_{obs} in which it cannot. Configuration space can be reduced by enlarging the obstacles by the size of the robot as shown in Fig. 1. This will mean that we can represent the robot as a point which will simplify further path planning. For successful path planning, the robot and its target location must be both in free space C_{free} [1].

III. PATH PLANNING

In robot path planning, most of the used algorithms consists of two parts: preprocessing in which workspace W is transformed into configuration space C and path planning itself in which we try to find the optimal path in configuration space C.

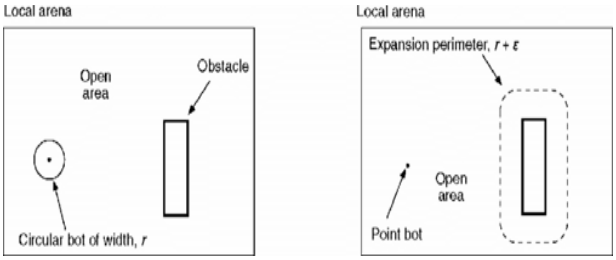


Fig. 1 Expanding obstacle dimensions by largets dimension of the robot [1].

A. Preprocessing

Preprocessing algorithms can be categorized into Cell Decomposition Algorithms, Potential Fields Algorithms, and Roadmap Algorithms. In this paper, we will focus on Roadmap Algorithms specifically on Visibility Graph mainly because of its simplicity. Roadmap algorithms create graphs of pathways on which the robot can move. All Roadmap algorithms are exact – if there is a valid collision-free path they will find it, if there is not, they will confirm that it does not exist. Fig. 2 shows an example of a Visibility Graph where black dots represent our starting point q_{init}, our goal q_{goal} and vertexes of all obstacles (black polygons). Lines represent connections between points with direct visibility. Algorithms used in this paper require all obstacles to be polygons, so any nonpolygonal obstacles must be first transformed into polygons [2].

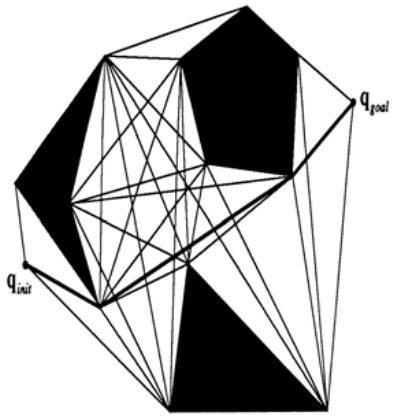


Fig. 2 The representation of the maze as Visibility graph [2].

B. Path Planning

After the creation of a graph structure representing configuration space C, we can apply our path planning algorithm to find the optimal path from our initial position to the

goal position. Before that, we will have to create a state graph which represents all possible states in which our robot can potentially reside. It consists of nodes each representing a position (state) at which our robot can reside and lines between them which represent changes of states. One of the states is initial state S in which our robot is at the beginning of our path planning, another one is the so-called goal state G which coincides with our planned goal. In this paper, we use an undirected, homogeneous, weighted graph which means that: all state changes are possible in both ways, all nodes are treated equally and each line between states has its value which can represent time, cost or some different property. An example of this graph is shown in Fig. 3. How to get from state S to G is up to the algorithm we use for our path planning [2].

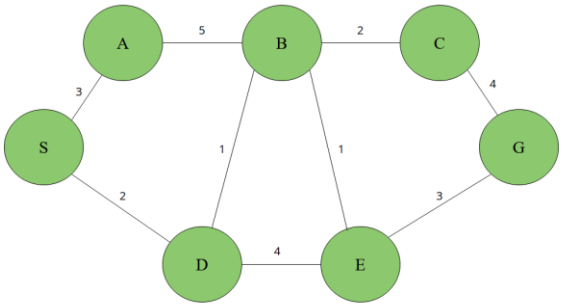


Fig. 3 The representation of the maze as undirected, homogeneous weighted state graph with direct path weights.

IV. Path Planning Algorithms

In this application, we compare three algorithms to find the optimal path in several benchmark mazes.

A. Dijkstra Algorithm

Dijkstra algorithm is said to be the so-called greedy algorithm. Greedy algorithms always proceed to a local minimum at each step. All vertexes have two properties: *visited* (determines if vertex was already visited), and *pathWeight* (calculates minimal cumulative path weight g to this vertex). During the process, we also use direct path weight d between adjacent nodes *n*. The node to which we are calculating weights is designated as current node *c*. Initially, it is node *s*.

The search algorithm process is as follows [4]:

1) Mark initial state node *s* with *pathWeight* of $g(s) = 0$ and as visited.

2) Search all adjacent nodes n, find their direct path $d(n)$ from node *s* and move to the closest one.

3) Search adjected nodes n and add their direct path $d(n)$ to the cumulative weight of current node $g(c)$, if this weight is smaller than the previously calculated cumulative weight $g(n)$ to this node, set new $g(n)$ to this node as $g(n) = d(n) + g(c)$.

4) Mark the current node c as visited and move to the node with the smallest recorded cumulative weight.

5) Repeat steps 3 and 4 until all nodes are visited.

B. A star Algorithm

A star algorithm is an expanded version of the Dijkstra algorithm which achieves better performance on larger problems by using heuristic functions to guide itself towards the goal position. It uses the same process as Dijkstra but calculates not only the smallest cumulative distance of the node from initial node $g(n)$ but adds to it a heuristic value of Euclidian distance h from the node n to goal node g creating function $f(n)=g(n)+h(n)$ which is calculated for all adjacent nodes. Using this heuristic function, the algorithm tends to move towards the goal node [3,5].

C. VND-DE algorithm

The variable Number of Dimensions Differential Evolution algorithm finds the optimal path by maintaining a population of candidate vectors (agents) *x* and creating a new population of agents by removing unfit agents and leaving only feasible ones.

To determine which agent is closer to optimum, we need an objective function. In our case it is a sum of Euclidian distances between each point of the agent including distances between the start and the first turning point and between the last turning point and the goal point.

$$f(x,y)=\sum_{i=1}^{n-1}\sqrt{(x_{i+1}-x_i)^2+(y_{i+1}-y_i)^2} \qquad (1)$$

where *x* and *y* are coordinates of points, N is a number of points in our path including the start and goal [6,7].

Reduced MATLAB code of this algorithm is in Fig.4 and it works it these steps:

1) The initial population is created randomly within a feasible decision space and each agent has its objective function value calculated. Each agent can have a variable length *d* in our case. This has to be an even number because every two elements of the agent vector corresponds to an *x*-*y* pair (coordinates of robot's turning points).

2) With previously known propability propDim we evaluate if agent changes its size in the next generation and change it accordingly .

3) By using three randomly picked agents and our mutation factor *F* we create our mutation vector *v* .

4) Using our crossover factor CR we pick which elements of mutation vector *v* to use in our trial vector *u*.

5) By comparing objective function values of our agent *x* and our created trial vector *u* we determine which vector to keep in our population.

6) steps 2-5 are repeated until we go through all generations with an intention to come closer to the global minimum. For this problem, it is the optimal path guiding the robot safely from the start point to the goal point around the obstacles. However, it is not guaranteed that the final solution will be optimal or even that we find a valid solution.

V. MATLAB Application

We developed a MATLAB tool for solving the robot's path planning problem. The individual steps of the path planning performed in the tool are summarized in the following sub-sections.

input:

```
Gmax, NP, d(2:2:n), CR[0,1], F(0.1+],probDim[0,1], lowlim, highlim
```

Initialize:

```
agmin=1; agmax=fix(NP/size(d,2));
for h=1:size(d,2)                   % for all agent sizes
    for iNP=agmin:agmax             % for selected part of population
        x{iNP}=lowlim+radn(1,d(h)).*(highlim-lowlim); %create agent vector
        fVals(iNP)= objFun(x{iNP}); % objecive function value calculation
    end
agmin=agmax+1; agmax=agmax*(h+1);
end
```

New generations

```
while iG<=Gmax
    for iA = 1:NP
```

Dimension of agent change

```
        r=randperm(NP,3); % generate 3 random different values [1;NP]
        if rand<= probDim               % change size of agent
            newDim= size(x{iA});        % agent size will stay the same
        else
            newdim=size(x{r(randi(3))});% randomly select 1 of 3 agents
        end                             % to change size to (can be same)
        tempX={f_dimcalc(iA,newDim)};   % create target vector
        tempR1=f_dimcalc(r(1),newDim); tempR2=f_dimcalc(r(2),newDim);
        tempR3=f_dimcalc(r(3),newDim);  % create 3 vector for mutations
```

Mutation

```
        v = tempR3 + F*(tempR1 - tempR2); % mutation vector
```

Crossover

```
        dim=1:newDim;
        ind = (rand(1, newDim) <= CR) | (dim == randi(newDim));
            % ind - vector of logical values which determines what parts
            % of trial vector we change
        u(iA, :) = tempX;           % u - trial vector
        u{iA, :}(ind) = v(ind);
        uVals = problem.objFun(u{iA,:});  % calculate objective function
                                          % value
```

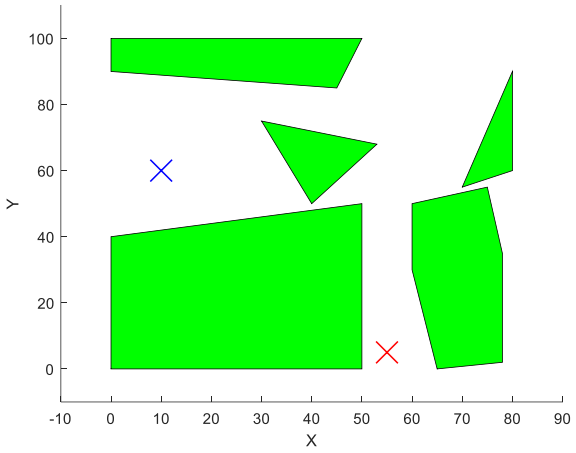
Elimination

```
        if  uVals < fVals(iA,:)             % rewrite agent x with u if
            x(iA,:) = u(iA,:);              % f(u) is better then f(x)
            fVals(iA,:) = uVals;
        end
    end
    iG = iG + 1;
end
```

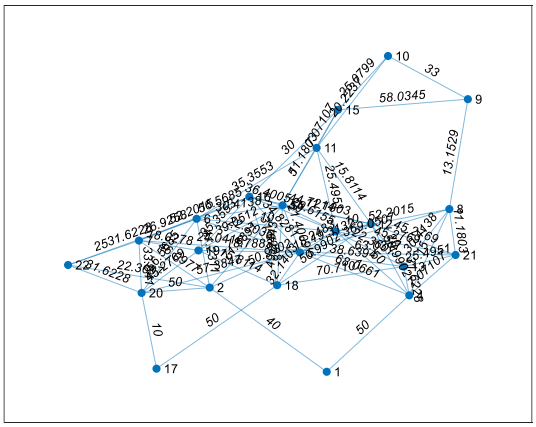
Description:

```
Gmax - number of generations,
NP - number of agents (population),
d - dimensions of agents (from 2 to n, all must be even)
CR - Crossover propability
F - mutation factor
propDim - propability of agents not changing its size
lowlim and highlim - lowest and highest value every agent can have
x - field of agent vectors
fVals - objective function values of agents
```

Fig. 4 Reduced MATLAB code of VND-DE algorithm.



a)



b)

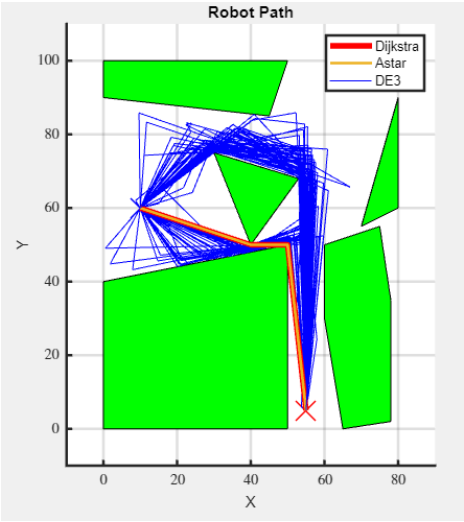
Fig. 5 a) Graphic representation of maze S02 b) State graph of S02.

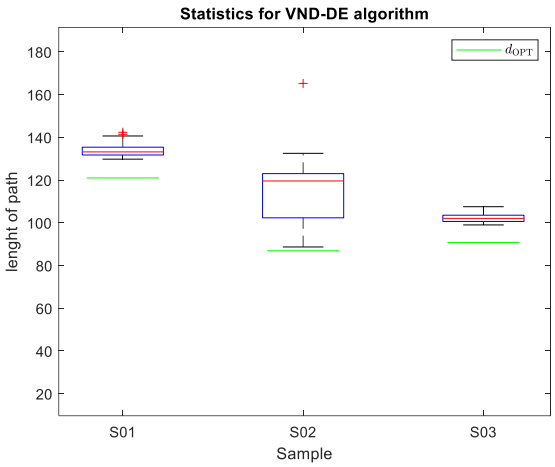
A. *Setup*

To successfully find a valid solution, it is needed to set up the maze with polygon obstacles only. The individual obstacles cannot overlap and the start point and goal points are not inside any of the obstacles. We use reduced configuration space so the robot is represented by a single point (red X) and obstacles are expanded with the largest dimension of the robot. An example can be seen in Fig. 5a). The preprocessing is done by Visibility Graph so all directly visible vertexes are valid in our path planning. This creates our state graph which is undirected, homogeneous and weighted as shown in Fig. 5b). For calculating our path with the VND-DE algorithm, we need to set up few parameters such as the number of generations and agents.

B. *Path planning*

After designing our maze, we can start path-planning algorithms and find optimal solutions. First, we find a path with the Dijkstra algorithm which will find a solution as well as the A star algorithm without a problem because of how our problem is defined. Solving our problem with A star and Dijkstra algorithms is a deterministic process and the solution found by both algorithms is always the same. On the other hand, searching for the optimal path with VND-DE algorithms is a stochastic process. Therefore, the global optimum can be reached only for a robust setting of the algorithm. Statistical data from several VND-DE runs are necessary to evaluate the convergence of this algorithm.





a) b)

Fig. 6 Results of 100 repetitions of the path planning performed by the VND-DE algorithm: a) resulting paths proposed by the algorithm for the sample problem S02; b) boxplots of the path lengths for sample problems S01-S03. Green line d_{OPT} is our optimal path calculated by Dijkstra and Astar algorithms.

RESULTS

Our results show data collected from 100 executions of our program for three different mazes: S01, S02, and S03. In Fig. 6a) you can see 100 consecutive executions of our pathfinding program for S02, in all of them Dijkstra and A star algorithms found the same path. All the path lengths are presented in dimensionless units. You can see statistics for 100 executions of our program on three samples using the VND-DE algorithm in Fig. 6b). The shortest paths were: f_{min} = 129.78 for S01, f_{min} = 88.67 for S02 and f_{min} = 98.99 for S03. For comparison purposes, the shortest paths of A star and Dijkstra were: f_{min} = 120.98 for S01, f_{min} = 86.89 for S02 and f_{min} = 90.79 for S03. The median of agent size was 4 for most results for all the samples which means that in most cases the shortest path calculated by the VND-DE algorithm had two turning points between the start and goal node.

VI. CONCLUSION

This paper presents a tool to build a maze in which a robot can move. Then, it provides us three algorithms for planning a way through the maze and presents solutions found by them. Dijkstra and A star algorithms find the same solution that is shorter than the one found by the VND-DE algorithm. It is important to mention, that even though the VND-DE algorithm was the slowest and found a longer solution, in some cases this solution can be very useful because for certain applications (e.g., the space robots), a longer path with fewer turning points may be preferable to the shortest path.

ACKNOWLEDGMENT

Research described in this paper was supported by the Internal Grant Agency of the Brno University of Technology under project no. FEKT-S-23-8191.

REFERENCES

[1] MAIDA, Anthony S., Suresh GOLCONDA, Pablo MEJIA, Arun LAKHOTIA a Charles CAVANAUGH. Subgoal-based local navigation and obstacle avoidance using a grid-distance field. International Journal of Vehicle Autonomous Systems [online]. 2006, 4(2/3/4), 8 [cit. 2021-12-22]. ISSN 1471-0226. Available at: https://doi.org/10.1504/IJVAS.2006.012203

[2] ABU-DAKKA, Fares. Trajectory planning for industrial robot using genetic algorithms. [online]. Valencia, 2011 [cit. 2021-12-22]. Available at: https://www.researchgate.net/publication/235929101_Trajectory_planning_for_industrial_robot_using_genetic_algorithms. Ph.D. UNIVERSITAT POLITÈCNICA DE VALÈNCIA.

[3] WAYAHDI, Muhammad Rhifky, Subhan Hafiz Nanda GINTING a Dinur SYAHPUTRA. Greedy, A-Star, and Dijkstra's Algorithms in Finding Shortest Path. International Journal of Advances in Data and Information Systems. 2021/02/01, 2(1), 45-52. Available at: https://doi.org/10.25008/ijadis.v2i1.1206.

[4] JAVAID, Muhammad Adeel. Understanding Dijkstra Algorithm. SSRN Electronic Journal [online]. [cit. 2023-03-07]. ISSN 1556-5068. Available at: https://doi.org/10.2139/ssrn.2340905

[5] KOENIG, Sven, Maxim LIKHACHEV a David FURCY. Lifelong Planning A. 155. Artificial Intelligence [online]. 2004, s. 93-146 [cit. 2021-12-31]. ISSN 00043702. Available at: https://doi.org/10.1016/j.artint.2003.12.001

[6] KUKKONEN, S. a J. LAMPINEN. GDE3: The third Evolution Step of Generalized Differential Evolution. In: 2005 IEEE Congress on Evolutionary Computation [online]. IEEE, 2005, s. 443-450 [cit. 2023-03-07]. ISBN 0-7803-9363-5. Available at: https://doi.org/10.1109/CEC.2005.1554717

[7] MAREK, Martin, Petr KADLEC. Another evolution of generalized differential evolution: variable number of dimensions. Engineering Optimization [online]. 2022, 54(1), 61-80 [cit. 2023-03-08]. ISSN 0305-215X. Available at: https://doi.org/10.1080/0305215X.2020.1853714

New home procedure on scara manipulator

1st Michal Marčonek
Deparment of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
230125@vut.cz

2nd Tomáš Benešl
Deparment of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
tomas.benesl@vutbr.cz

***Abstract*—This project deals with a new home procedure system for a Seiko Epson robotic manipulator, which has a new Beckhoff control system. The first attempt was to use the original homing system, but it was found to be unusable with our resources for the first two axes. Subsequently, the possibilities of replacing the original solution and designing a new solution with better parameters are being discussed.**

***Index Terms*—scara, homing, encoders, optical sensors**

I. Introduction

To begin with, it is necessary to introduce the SCARA manipulator we are working with. It is a product of Seiko Epson, which was manufactured in the nineties. The manipulator is equipped with brushless motors by YASKAWA ELECTRIC in power (ordered from the first axis) 200W, 100W, 100W and 50W. The motor of the third axis is also equipped with a brake, which is braked in the idle state. Each motor has a Sony Magnescale RE90C-2048C magnetic encoder on its axis. The original manipulator control system has been replaced by a Beckhoff solution for the following reasons:

- The original control box failed and no spare parts were available due to the obsolescence of the manipulator
- The original software of the control box was primitive for today and did not allow the use of the manipulator in experimental tasks of Industry 4.0
- The only communication bus was a serial line or digital inputs

[8]

The new control box Fig. 1 of the manipulator is equipped with a Beckhoff C6015 industrial computer, two AX5201, AX5203 inverters and multiple input and output cards. The components are interconnected using the EtherCAT protocol which is a special industrial Ethernet extension from Beckhoff. Furthermore, the control box houses the voltage supplies for the control logic and for the inverters.

[9]

II. What's the home procedure?

The home procedure is a set of actions that ensure the automatic guiding of robot parts from a generally unknown position (i.e. any position in the robot's workspace) to a pre-known starting position called the home position. The home procedure is determined by the linear or rotary actuators used. In order to start the procedure at all, it is necessary to know at least the rough position of the robot, i.e. the quadrant

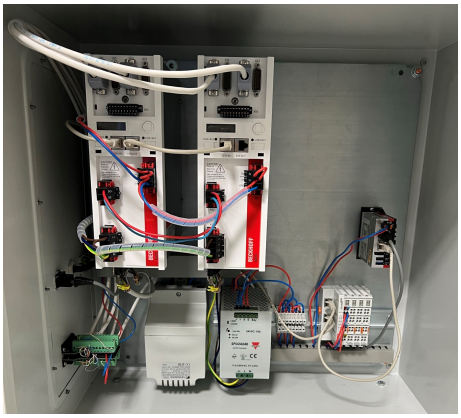


Fig. 1. View of the new control box

for rotary actuators and the interval for linear actuators. Special mechanical or electromechanical elements are used to determine the position (Fig. 2 and Fig. 3) [7]

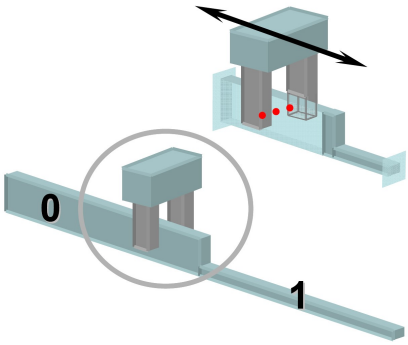


Fig. 2. Homing mark for linear drive [7]

A. Proces home procedury

1) Start the robot and the control system evaluates the values from the home sensors.
2) Based on the evaluated inputs, the robot moves in the corresponding directions.
3) During the movement, the values in the relative sensors are already registered.

4) As soon as one of the home sensors registers a value change, the value from the encoder of the corresponding axis is stored in memory and the axis movement stops.
5) When all axes reach the home position, the robot is homed and stabilized by Hold [1] stopping. [7]

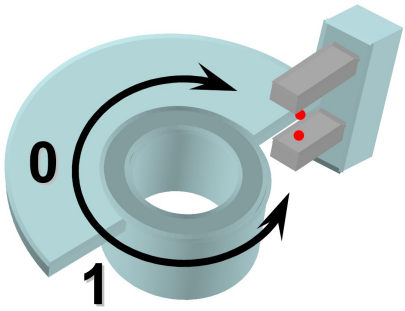


Fig. 3. Homing mark for rotary drive [7]

B. Homing options

1) Homing on limit switch: This is not a simple homing procedure. The robot can traverse its entire range of axes during the procedure. This method is used in 3d printers. During this homing method there must be no obstacle in the working area of the robot otherwise there is a risk of collision. [4]

2) Multipoint homing: This is the method used by our manipulator, for example, and consists in the control system knowing the spacing between the edges of the teeth and the position of the edges Fig. 4. Then only during the homing procedure the control system moves the axis in a relatively small range around $\pm 9°$. This method allows us to already leave the tools in the workspace in the robot's trajectory. However, this method is not perfect and collisions can still occur in special situations. [8]

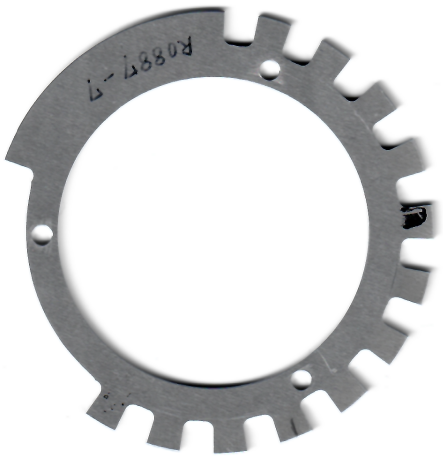


Fig. 4. Original homing wheel of Epson manipulator

3) Homing by memorized previous position: Today's manipulators no longer primarily perform the home procedure according to an absolute sensor, but know their real pose thanks to continuously powered encoders that are equipped with batteries. For example, the current line of Epson LS-B manipulators has a block of lithium batteries in the base of the manipulator. Another example is the Fanuc six-axis manipulator, which is located in the Department of Control and Instrumentation laboratories, the batteries have to be replaced regularly (about once a year). [2], [3]

III. Sensors for homing

1) Sensors for homing with limit switch: For this task, position sensors with a step change in the output variable are used. This type of sensors is very common in 3d printers. Mechanical and optical sensors are among the most used, inductive sensors and Hall effect probes are also used. But of course other types of position sensors with step change resistance can also be used.

2) Sensors for multipoint homing: These are encoders, which can generally be described as mechanical motion sensors that generate a digital signal in response to motion. Thanks to encoders we are able to detect information about position, speed and direction. Encoders can be divided according to the information provided into:

- Incremental - we receive information in the form of directional increments of angles or segments
- Absolute - we get information about the exact angle or segment from the reference point

Both types of sensors are used on the manipulator due to the higher accuracy of incremental encoders and the passive properties of absolute encoders.

IV. Identification of the original multipoint homing principle

Because the manipulator is equipped with a system of multipoint home procedures in the first two axes. When in the joints of the mentioned axes there is Fig. 4 a gear with seemingly identical teeth. The tooth sensing is handled similarly to Fig. 3. Therefore, one revolution of the gear is experimentally measured when the output of the optical sensor is recorded on an oscilloscope. Furthermore, one cannot help but notice the handwritten number in Fig. 4, where this number appears multiple times on the manipulator and the original control box. Therefore, it can be assumed that the original control was tied to our particular manipulator at the time of manufacture.

From the graph it can be seen that the measurement was accurate, but despite all this it would be very difficult and not very reliable to measure the distance between the teeth or the width of the teeth of the homing wheel. This is due to the observation that using this solution, the manipulator can only be brought to a known position on the tenth attempt in unfavourable cases. For this reason, this multipoint homing solution will not be used.

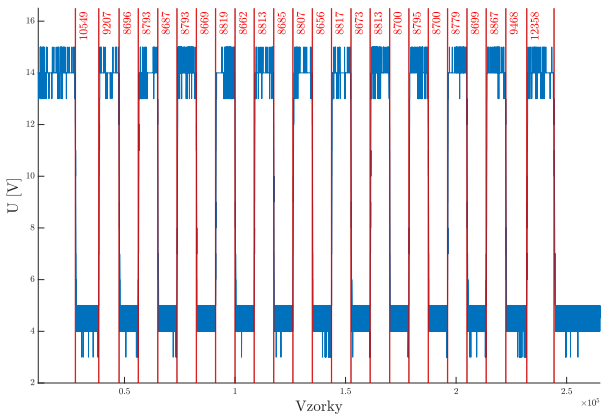


Fig. 5. Processed data in matlab - red vertical always runs along the falling/rising edge at the decision point of the input card, red numbers describe the number of samples between the verticals

V. Choosing the best homing solution

For the sake of preserving the properties, especially the property of minimal movement of the first and second axis during the home procedure. This means that to determine the rotation angle of each axis we need to know their absolute value at any position. This requirement precludes us from using a limit switch for the first and second axes. At the same time, another requirement is to use a solution such that batteries are not needed. This precludes the aforementioned solution of storing the last position in memory and maintaining it using lithium batteries. Now we are left with the use of an absolute encoder. For this encoder, the angle will be encoded in Gray's code because with this encoding it is easy to implement a check of the correctness of the data. However, a problem arises with the need for a small footprint that would allow the sensor to be built directly into the joint in place of the original homing solution. So during internet searches, a special encoder solution [5] was found that uses a single track. The trace is encoded with a specially modified Gray code or pseudocode of [6]. The segment widths of the Fig. 6 pattern are used from the sample solution in [6].

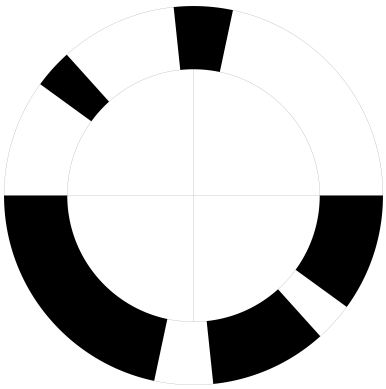


Fig. 6. Code wheel pattern

The advantage of this solution is that the sensors sensing the value from the sample wheel do not need to be all around the perimeter, but it is enough to cover less than half of Fig. 7.

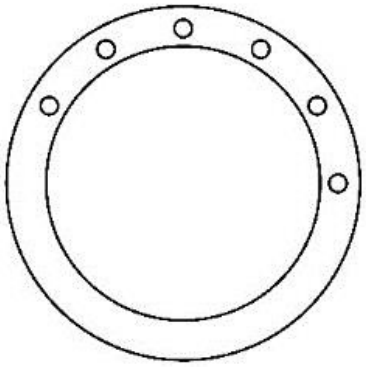


Fig. 7. Sensor locations on the sensor

VI. Design of PCB

Since our solution depends on a uniform distribution of sensors, it is necessary to design a corresponding board, but it is also necessary to take into account the dimensional possibilities on the robot. Therefore, a semicircular board is created, which is equipped with QRE1113 optical sensors. These sensors stand out for their small size and are therefore suitable for this project. Fig. 8.

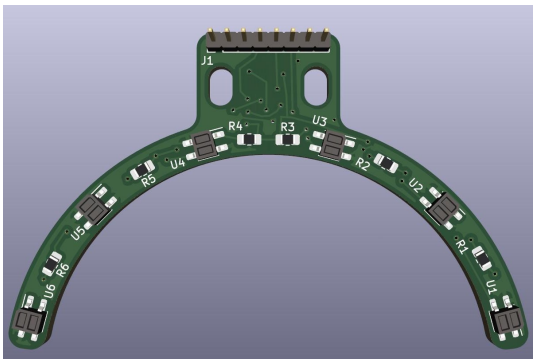


Fig. 8. PCB with sensors

For dimensional reasons, the board could not be designed for industrial 24VDC voltage, but a compromise had to be made. The sensor board is at 5VDC and inside the manipulator body there will be a so-called dongle board changing levels between 24VDC and 5VDC.

VII. Conclusion

In this project, the solution of a new system for the home procedure of the Epson scara manipulator was addressed, mainly from a sensor point of view. We found it possible to design and solve the problem using a single-track absolute encoder with Gray code. The advantage of the solution is higher accuracy and less hardware requirements for the input circuits. For the first two axes the encoder solution is used and for the remaining axes the solution is by Heading II.

ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-23-8451 - "Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence, automation and measurement" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] BURIAN, František. Přednáška - Bezpečnost robotických systémů. *BPC-RBM - stránky kurzu* [online]. Brno: ysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav automatizace a měřicí techniky, 2022 [cit. 2023-01-04]. Dostupné z: https://www.uamt.fekt.vut.cz/ burianf/bpc-rbm/keynotes/bpc-rbm-t12-p-BezpRob.pdf

[2] SCARA ROBOT LS series MANIPULATOR MANUAL. *Google Drive fy Opticontrol* [online]. Brno: Opticontrol, 2020 [cit. 2023-01-04]. Dostupné z: https://drive.google.com/file/d/1Ky2PsHvCWR13LgRLuYswWXoSeKPjuxge/view

[3] FANUC Robot LR Mate 200iD MANUAL. *HAAS CNC* [online]. Oxnard: Haas Automation, 1996 [cit. 2023-01-04]. Dostupné z: https://www.haascnc.com/content/dam/haascnc/en/service/reference/fanuc-manuals/Fanuc%20Robot%20LR%20Mate%20200iD%20Operators%20Manual.pdf

[4] 3D Printer Endstops Types — Best Choice For You?. *TheMechNinja* [online]. Jhansi: TheMechNinja, 2021 [cit. 2023-01-04]. Dostupné z: https://themechninja.com/08/3d-printer-endstops-types-best-choice-for-you/

[5] HOLYBA, Ferenc. Absolute Position Encoder With Single Track Gray Code. *Autodesk Instructables* [online]. San Francisco: Autodesk, 2011, 7. listopadu 2021 [cit. 2023-01-03]. Dostupné z: https://www.instructables.com/Absolute-Position-Encoder-With-Single-Track-Gray-C/

[6] ZHANG, Fan, Hengjun ZHU, Kan BIAN, Pengcheng LIU a Jianhui ZHANG. Absolute Position Coding Method for Angular Sensor-Single-Track Gray Codes. *Sensors* [online]. 2018, **18**(8), 1-6 [cit. 2022-12-31]. ISSN 1424-8220. Dostupné z: doi:10.3390/s18082728

[7] BELDA, Květoslav a Pavel PÍŠA. Homing, Calibration and Model-Based Predictive Control for Planar Parallel Robots. *Institute of Information Theory and Automation* [online]. Praha: Institute of Information Theory and Automation, 2008, s. 1-6 [cit. 2022-12-28]. Dostupné z: http://www.utia.cas.cz/files/Soutez_08/Publikace/Belda.pdf

[8] , SEIKO EPSON CORPORATION. *Scara Roboter Typ BN*. Rev. 5. Düsseldorf, 1998.

[9] BITTNER, Jakub. *Řídicí systém robotu SCARA* [online]. Brno, 2022, 74 s. [cit. 2022-12-26]. Dostupné z: https://www.vut.cz/www_base/zav_prace_soubor_verejne.php?file_id=239265. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav automatizace a měřicí techniky. Vedoucí práce Ing. Tomáš Benešl.

Automated System for Creation of Gigapixel Images

David Mlejnek
Faculty of Electrical Engineering and Communication
Department of Control and Instrumentation
Brno University of Technology, BUT
Brno, Czechia
xmlejn05@vutbr.cz

Jakub Arm
Faculty of Electrical Engineering and Communication
Department of Control and Instrumentation
Brno University of Technology, BUT
Brno, Czechia
Jakub.Arm@vut.cz

***Abstract*— The paper summarizes the design and implementation of a system enabling the creation of gigapixel images. The system adopts an open-source gizmo design enabling pan-tilt motion. The electrical and software part is designed to be scalable and uses ESP32 as a motion controller of stepper motors. The central component embodies Raspberry Pi that communicates with Canon camera via CCAPI interface and sends taken photos to a server. On the server, Hugin software creates a gigapixel image from the photos. The preliminary results suggest the mechanical and electrical part have sufficient parameters to provide the precise and stable movement of the camera.**

Keywords*— *Gizmo mechanism, Pan tilt camera mount, ESP32, Stepper motor, Camera, Gigapixel images

I. INTRODUCTION

This paper focuses on the process of automated capturing of gigapixel images. The topic is split into several part and the partial results obtained are presented. The system will consist of three basic components, namely the camera, pan-tilt mount, the power and control electronics. The assembly should work as follows: In step 1, the user chooses the area to be captured. In step 2, the system switches on that automatically starts capturing the area. In step 3, the captured images are sent to server where they will be assembled by Hugin software. The system output will be a gigapixel image, represented by the captured area. In the future, the system will be improved – for example, it may be mounted on a rail that will serve as the third axis and the system will move on it. The text contains a summary of values of mechanical and optical parameters. The existing results presented deal with the description and control of the pan-tilt mount.

II. GIGAPIXEL IMAGE

A gigapixel image is an image that has more than a billion pixels and the photographer cannot make it use a single photo from a classic commercial camera.

A. Introduction to Gigapixel image

There are only few gigapixel cameras in the world. They have been designed for a specific scientific activity. An example is the camera fitted in the American observatory Pan-STARRS PS1 in Hawaii. With a resolution of almost 1.4 billion pixels, this camera is used to capture the night sky and then to survey the astronomical objects. For classic commercial cameras, the size of the camera chip is a limiting element; for example, Canon EOS 250D has a camera chip with a resolution of 24,1 Megapixels (Mpx). To make a gigapixel image, it is necessary to capture more images and put them together. System with a set of static microcameras

To collect the data (the images) to get the final result, there are two basic procedures, namely using more micro-cameras and capturing all images at one moment or using a single camera while capturing the images gradually and setting the camera position. The system with a set of static micro-cameras (CCD or CMOS sensors) operates so that the field of view (hereinafter the "FoV") of individual sensors scan a certain area and at the same time, they are mutually overlapping. The system AWARE-2 from the article quoted in the References [1] was selected as an example. This system contains 98 micro-cameras, each of them having a resolution of 14 Mpx. All the sensors together have a resolution of 1.372 Gpx; after elimination of overlaps, approximately 0.96 Gpx will remain. The figure (1) from article [1] shows the FoV of the AWARE-2 camera. The same system is applied by the camera fitted in the above-mentioned Pan-STARRS Observatory. Such systems are also used in microscopes to obtain gigapixel images, see article [5].

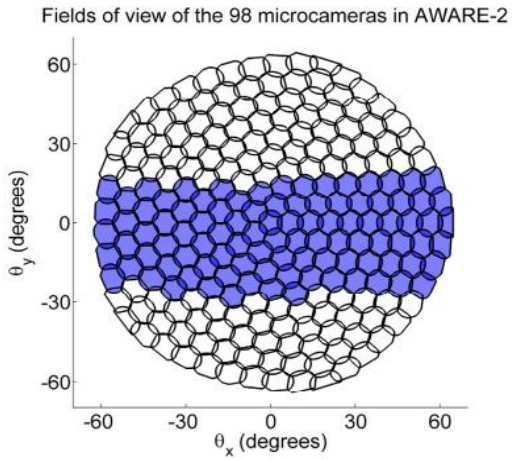


Figure 1 Field of view of special camera AWARE-2 [1]

B. *System with a single camera*

Unlike the preceding system, the system with a single camera is easy to assemble, a commercially available camera and a tripod (stative) are sufficient. As several images must be obtained, the camera must focus on different locations. Only after focusing on the location, we may capture the image. This restricts the application only to the capturing of highly statical sceneries due to the time required to set the camera to the new position. For example, a gigapixel image of a running person cannot be obtained from this system.

C. *Comparing the systems*

In terms of price, the system with a group of the statically set micro-cameras is incomparably more expensive as such systems are produced with specific parameters for exactly specified purposes, especially the scientific ones. The system with a single camera is cheaper, just a camera and a tripod are required. Another important parameter is the time needed to scan all the necessary images. In this regard, the system with more cameras is more advantageous. All images are taken at a single moment. To collect data with the single-camera system, the camera must be moved which takes time. Another parameter is the setting of the system axes. In the solution with more micro-cameras, the axes of individual micro-cameras have been preset and may be calibrated as needed. In the solution with a single camera, the axes must be set in accordance with the requirements for the resulting image. The system is limited by the minimum angle by which the axis may rotate. The system with a single camera was selected as the costs of its assembly are lower.

III. SYSTEM ARCHITECTURE

The Figure 2 shows the designed system architecture.

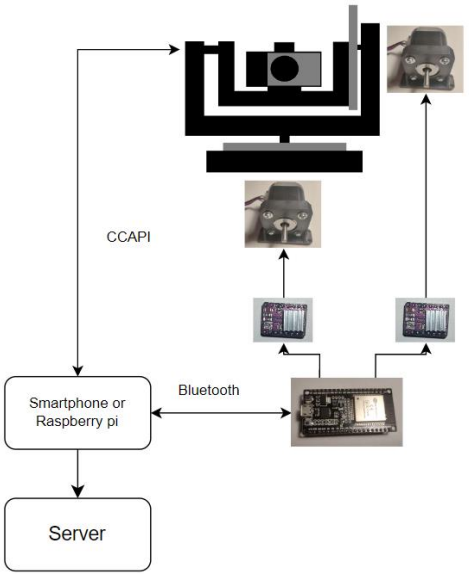


Figure 2 System architecture

Software will be designed for the system control that will be implemented either on a smartphone or Raspberry PI. The software will communicate with the camera via CCAPI and communication with microcontroller will occur via Bluetooth when using the smartphone, and UART when using Raspberry PI. The images will be sent from the phone or minicomputer to a remote server where the computer assembles them into a single image. The images may be assembled using a program for image assembly.

A. *Pan-tilt mount*

The core of the entire system is the pan-tilt mount. It is a mechanical system containing an attachment for the camera and rotating axes allowing the focus of the camera on the required location. For our needs, the freely available pan-tilt mounts. Selection was made out of different pan-tilt mounts on the Thingiverse website, of which the best two models will be described herein. Both pan-tilt mounts operate on the same principles and are suitable for our camera. The only difference consists in the design of the gears and the gear ratio. In terms of the gear ratio which is an important aspect for the parameter of the smallest angle shift, the model Camera Gimbal [6] which has a higher gear ratio on both axes, is better. In terms of the design of gears, the model DSLR Camera Pan Tilt Mount (Stepper Motor Driven) [7] which has arrow-shaped gears, is more suitable. The model Camera Gimbal [6] was selected thanks to its higher gear ratio, but the installation of arrow-shaped gears is planned in the future.

This mechanism allows movements both in horizontal and vertical axis using electric motors. The axes are deposited in

ball-bearings to reach minimum axis friction while moving. The gear wheel of the vertical axis has 160 teeth, while the gear wheel of the horizontal axis has 128 teeth. The drive wheels located on motor axes have 16 teeth. This implies a gear ratio of 1:10 for vertical axis and 1:8 for horizontal axis.

B. Electronic equipment

The camera to be used will be Canon EOS 250D as it is compatible with CCAPI communication with lens EF-S 60mm f/2.8 Macro USM with optical parameters in table 1:

Minimum focus distance (m)	0.2
Filter thread (mm)	52
Maximum aperture – tele (F)	2.8
Aperture blades	7
Maximum aperature (F)	32
Static focal distance (mm)	60
Auto focus	Yes
FoV with crop factor 1,6 and Sensor size 22.5x15mm	13,368deg horizontal 8,934deg vertical

Table 1 optical parameters

As the accurate control of axes is of great importance, a drive allowing an accurate control will be required, and therefore the stepper motors Nema 17 42HD6021-03 with drivers DRV8825 were selected. In the final stage, TMC2209 will replace the stepper driver. It is possible to use servomotors featuring the advantage of having feedback directly incorporated in them, but this option would be rather costly. Therefore, stepper motors were preferred. However, it is possible to create feedback for stepper motors as well. By doing so, they would be equivalent to servomotors. The Figure 4 shows the attached motors.



Figure 3 Pan-tilt mount with stepper motors

To control motor movements, a microcontroller is necessary. In this case, we need a microcontroller that is capable of Bluetooth communication in case that a smartphone is used. For example, the development kit ESP-WROOM-32 complies with these requirements and is already used in the project. The Figure 5 shows the control circuit for stepper motors.

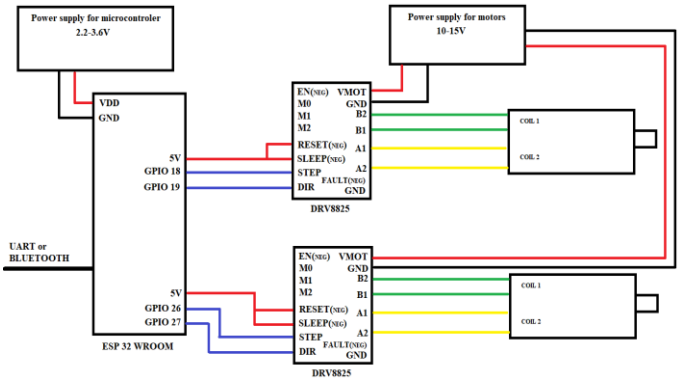


Figure 4 diagram of the control unit

Either a smartphone or a minicomputer may be used as the control element to control the operation of the entire system. This has not been decided yet, but in terms of mobility and simplicity of the entire system, a smartphone will probably be a better solution.

C. Software for image assembly

The software for image assembly must usually be paid for. It is possible to download free versions of such software, but these are restricted either in terms of time, or the resulting image has a visible watermark – such as in case of the PTGui program, therefore the quality of the resulting image is not as good as in the paid versions. Free software also exists, such as Hugin. Also, the free program GIMP is capable of assembling images – it is just sufficient to download addons. Free programs are not equivalent to commercial programs in terms of performance. The Hugin software will be sufficient for basic testing. If a higher-performance interface is required, one of the commercial programs may be used.

D. Summarizing the existing results and further procedure

As a part of the paper, a mechanism was assembled on which a camera was mounted, and the mechanism was put into movement with stepper motors. The next task will be the creation of communication between the ESP-WROOM-32 microcontroller and the phone, and between the phone and the camera and designing software that will control the automated activity of the whole system. Next, the issue of assembling the images into the final gigapixel image must be solved. It would be beneficial to create feedback for accurate control of motor rotation – application of rotation encoders would be suitable for that purpose that would record the rotation position of individual axes. After generation of the feedback, we may consider the work on dynamic control of the system.

E. Basic system parameters

We may calculate the basic system parameters based on the known motor values and driver values. This calculation shows how many steps of motors are required for the rotation of one axis by 360°. It results from the stepper motor datasheet that they make 200 steps per one shaft rotation, and we also know the gear ratios on both axes. With that in mind, the gear ratio is 1:10 for the vertical axis and 1:8 for the horizontal axis. After multiplying the gear ratios and the number of steps of motors, we get the result that the required number of steps per one rotation in vertical axis is 2000 steps and in horizontal axis 1600 steps.

To determine the smallest angle of rotation of the mechanism, we consider 360° and divide it by the number of steps per one rotation. It results that the smallest angle of rotation for the vertical axis is 10′ 48″ and for horizontal axis 13′ 30″. For the maximum micro-stepping of drivers DRV 8825, which is 32, the number of steps per one rotation on the vertical axis will be 64,000 and on the horizontal axis 51,200. This will reduce the smallest angle of rotation on the vertical axis to 19″ and on the horizontal axis to 25″.

Having these values available, it is possible to calculate what change in height will result from the shift by the smallest angle of rotation when capturing an object at a distance of 100 meters. Using the goniometric tan function and the known smallest angle of rotation (angle of the vertical axis), the result is that the change in height on the object being captured will be 31.42 cm. For micro-stepping, the change in height will be 0.92 cm.

Another parameter that we may obtain is the minimum number of images required to assemble a gigapixel image. We know the resolution of the Canon EOS 250D camera which is 24.1 Megapixel (Mpx). To get a gigapixel image, we need a minimum of 1 Gpx. After dividing and rounding up, the result is that we need at least 42 images. This number, however, is difficult to reach and more images will be needed as the assembly must also take image overlaps into account.

IV. Vision

To verify operability of the said concept, we use a smartphone instead of Raspberry PI that will significantly simplify the work in the system in conditions other than the laboratory ones.

It is also intended to test more software solutions to assemble the gigapixel images and evaluate the limits of the mechanism. The mechanism movement in individual axes may be improved by using the arrow-shaped gearing. The mechanism solidity may be improved if we use aluminum structure instead of the plastic one. Implementation of accurate encoders on the rotation axes will guarantee a reliable operation with the possibility of movement regulation.

Future visions for system improvements consist in the depositing of the entire gizmo mechanism on a movable linear axis that will reduce the bias of the final image.

CONCLUSION

The paper verifies the use of the selected mechanism intended to capture gigapixel images. The control assembly was implemented to control the unit movement – for this purpose, the stepper motors were used to ensure accurate positioning and ESP32 for movement control. The smallest step of the mechanism when using micro-stepping is 19″ in the vertical axis and when using micro-stepping is 19″ and 25″ in horizontal axis. The available CCAPI interface will provide for the control itself of the camera while the final image will be created by the Hugin software that will run on the server. The work activities up to now verified the possibility of creating a capturing system using low-end components that guarantee sufficient system parameters.

ACKNOWLEDGEMENTS

The completion of this paper was made possible by the grant No. FEKT-S-23-8451 - "Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence,

automation and measurement" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] GOLISH, D. R., et al. Development of a scalable image formation pipeline for multiscale gigapixel photography. *Optics express*, 2012, 20.20: 22048-22062.

[2] LATO, Matthew J.; BEVAN, George; FERGUSSON, Michael. Gigapixel imaging and photogrammetry: Development of a new long range remote imaging technique. *Remote Sensing*, 2012, 4.10: 3006-3021.

[3] SENTHILNATHAN, R., et al. FPGA Implementation of Closed Loop Control of Pan-Tilt Mechanism.

[4] CABEZOS-BERNAL, Pedro M.; RODRIGUEZ-NAVARRO, Pablo; GIL-PIQUERAS, Teresa. Documenting paintings with gigapixel photography. *Journal of imaging*, 2021, 7.8: 156.

[5] Eric E Thomson, Mark Harfouche, Kanghyun Kim, Pavan C Konda, Catherine W Seitz, Colin Cooke, Shiqi Xu, Whitney S Jacobs, Robin Blazing, Yang Chen, Sunanda Sharma, Timothy W Dunn, Jaehee Park, Roarke W Horstmeyer, Eva A Naumann (2022) Gigapixel imaging with a novel multi-camera array microscope eLife 11:e74988

[6] Thingiverse. Camera Gimbal. fhuable. 2019. Available from: https://www.thingiverse.com/thing:3375167

[7] Thingiverse. DSLR Camera Pan Tilt Mount (Stepper Motor Driven). isaac879.2020. Available from: https://www.thingiverse.com/thing:4316563

Deployment of Linear Programming for Creating Cybersecurity Profiles

1st Michal Stejskal
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
231282@vut.cz

2nd Petr Dzurenda
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
0000-0002-4366-3950

***Abstract*—The cybersecurity job market is growing rapidly, with a great demand for cybersecurity professionals. Currently, there are many cybersecurity curricula, professional training, and certification schemes that have been created recently, however, they are independent of each other and without any synchronization. The recently released European Cybersecurity Skills Framework (ECSF) aims to bring order to this area. This article builds on the work of the European Union Agency for Cybersecurity (ENISA) and extends it with a practical tool that allows professionals to create their desired ENISA cybersecurity profiles. The core of our tool uses a Linear Programming (LP) algorithm to help professionals to find the most suitable combinations of professional training courses to follow in order to achieve their required ENISA profiles. Our experimental results show the efficiency of our implementation, in terms of the speed of finding a solution for different LP solver libraries.**

***Index Terms*—ENISA, ECSF, Cybersecurity Skills Framework, Cybersecurity, Cybersecurity Profile, Linear Programming, Simplex Algorithm**

I. Introduction

Because of the expanding demand for cybersecurity professionals on the European, but also the worldwide market, new cybersecurity terms, roles, and skills are being created and named. However, these names are often quite different. This can bring confusion and misunderstanding about what should be taught in professional courses and what is actually required by the job market. Therefore, ENISA undertook to solve this problem and created the ECSF that unifies the terms used in this specialization [1]. Taking this into account, the Cybersecurity Skills Alliance - A New Vision for Europe (REWIRE) project, created a connection between ECSF skills and knowledge and REWIRE cybersecurity competencies. These REWIRE competencies are then linked back to ENISA profiles. With this methodology, it is possible to map existing university curricula, professional training, and certification courses to ECSF. Furthermore, the REWIRE project designed a web application called Cybersecurity Profiler (CSP) which can serve people in this field, as it includes courses all over Europe and multiple tools, which can help with creating a university curriculum, professional training and certification courses, and understanding requirements for fulfilling job roles. We refer to [2] for more details.

A. Contribution and Paper Structure

This paper focus on our design, implementation, and experimental tests of the computing core of the CSP application allowing identifying which courses, training, and certification courses are recommended for a certain work role.

Section II describes the fundamental knowledge required for understanding the problem. Section III presents the target of the module, which is to assemble courses in such a combination, which fulfills parameters set by the user of the application and match with the chosen profile. Details of implemented libraries and code structure are presented as well. Section IV is dedicated to testing and experiments with different Linear Programming (LP) solver libraries, so that the best library can be selected by setting criteria relevant to this work.

II. Preliminaries

In this section, we describe the ECSF to foreshadow its connection to created CSP application. Furthermore, linear programming and its branch, the simplex algorithm is introduced as the most prominent option for solving the problem of finding combinations from available courses.

A. European Cybersecurity Skills Framework (ECSF)

The ENISA has summarized knowledge, and skills in the field of cybersecurity into twelve profiles [1]. Classification as this gives options not only to individuals seeking proper education but also allows organizations to specify which professions they need and simultaneously training providers can use the same terms to offer their services. Unifying terms in this market will help this area to grow and hopefully provide better conditions for people to become professionals. More experts are required to protect critical infrastructure, privacy and similar vulnerabilities that arise.

Project REWIRE followed up on this work and included skill groups in the created profiles, that reflect which skills need to be mastered to fulfill the profile. Skill groups consist of more detailed skills and knowledge.

B. Integer Linear Programming and Simplex Algorithm

Linear programming is a mathematical discipline that targets finding the best outcome. The objective has to be finding the maximum or minimum of a function. The algorithm aims

to find the maximum or minimum of the objective function while respecting the given constraints, see [3] for more details. The simplex [4] is a standard algorithm often used because it is simple and widespread. Simplex uses bounds to establish feasible region. The boundaries are then traveled along the edges until the optimum of function is reached. This method always presents global optimum if precautions against cycling are applied.

C. Solving packages for ILP

Solvers use different approaches or techniques to compute the solutions of a problem. As each solver usually allows to solve a vast number of different problems, choosing the proper solver can make a relevant difference in performance. It must be clear that solving a problem just by using the simplex method cannot fulfill all types of tasks, and Mixed Integer Programming (MIP) is proof of that [4]. Some problems include values that are required to be binary but the simplex method will produce a best global solution which might not be an integer value. Here the solver has to use its methods to find a solution with integer values. This can be achieved for example by Branch-and-Cut algorithm [5].

The algorithm works as follows, the first step is creating bounds, if the desired outcome is an integer and a continuous number is produced instead, then bounds are created around the value. For example, if the produced number is 4.6 which is not an integer, then the next step is to create a lower bound that would be 4.0 and an upper bound of 5.0. This creates two new constraints which are included for the simplex. Two calculations are performed with bounds appended separately and if solutions consisting of integers are found, then the better of these solutions is considered optimal. Another case is when this does not produce an integer, then branching occurs and similar bounds are set again and calculation is performed on them until a feasible solution is found and considered optimal compared to other integer solutions generated by different branches.

III. CYBERSECURITY PROFILE DESIGNER

The objective for this section is to give a detail description of the module designed for web application. The first subsection describes how the module was designed into detail. Then in next subsection some terms and details about possibilities of external libraries and packages are explained.

A. CSP Search Problem Formulation

First, it is important to outline the problem that needs to be solved. There are 59 courses included in the set at the time. The user can choose a profile that interests him. Profiles and trainings consist of skill groups. Algorithm has to match the trainings to chosen profile so that all skill groups included in the profile are represented by trainings. Because single training cannot fulfill this, multiple trainings have to be put together to achieve possible combinations. The found solution has to consist of the best available courses that match the profile and the selected parameters set up by the user. It is also essential to get an optimal solution in the shortest possible time. The filtering applied by the user includes important numerical values as price, duration, and included skills of the courses. These values together with a number of courses demanded in the solution are subjected to optimization so that the best possible outcome is found. To achieve this, extensive analysis had to be conducted and multiple possibilities were discovered. Between the most promising had ranked LP that on its own offers multiple algorithms and Genetic algorithms (GA) [4].

Regarding GA there are multiple red flags, to begin with. Among the biggest ones ranks the uncertainty that an optimal solution will be found, i.e., the algorithm does not always finish at a globally optimal solution, as in our case, where the final demanded state is not outlined to the algorithm. Another issue worth mentioning is the deviation time from finding each solution.

On the contrary, LP does not have to deal with these fundamental problems, because solving a problem with them is basically following mathematical instructions that lead towards an optimal solution and thus the result is always the same as the same approach is repeated. Unfortunately, there is a problem associated with this. The solution produced by the mathematical approach cannot guarantee that the result will contain whole numbers which are relevant to this work. This problem is solved with the help of solver packages, which will be discussed in more detail in Section III.

B. Application Design

The application currently operates only in python. Data are drawn from JSON local files. After data are extracted from files, they are filtered based on criteria like the language in which are the lessons conducted or the country from which they originate. Another criterion is whether the course ends with certification and also whether it is conducted online, face to face, or both. All mentioned filter settings including mentioned price, duration and mainly chosen profile/s are set up by a user.

The process of optimization then comes into place. Firstly, the target of optimization is acquired, this is statically set for minimization as the objective is to minimize price, duration, and number of included courses. These are values that the solver calculates with. Then variables are created. In this case, variables are courses themselves and they are represented by either one or zero, depending on whether they will be included (1) or excluded (0). That is determined by calculations. Further, the objective function is defined together with constraints and bounds, including the filtering values.

The objective function is presented mathematically in Equation 1, where S_{min} is a minimization of the objective function. Here p_{r_i} stands for price and is added together with d_{u_i} which represents duration. Both of the values are multiplied by b_i which represents a binary variable deciding whether the course will be included or not.

$$S_{min} = \sum_{i=0}^{M}[(p_{r_i}b_i) + (d_{u_i}b_i)] + \sum_{i=0}^{M} b_i \tag{1}$$

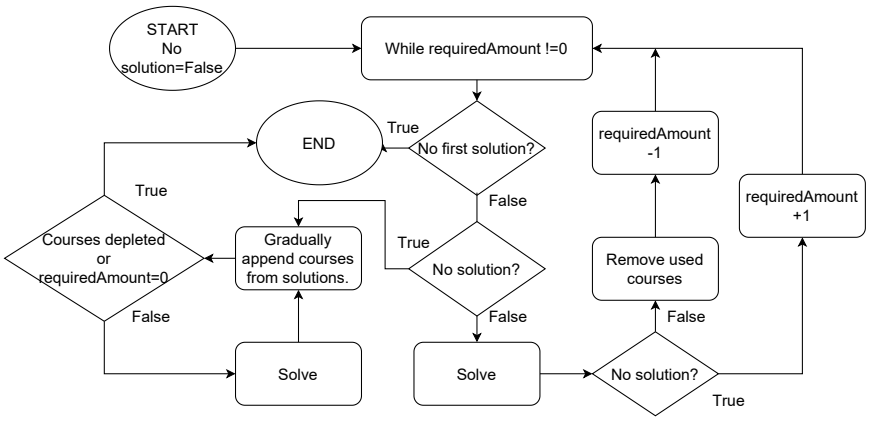


Fig. 1. Decision loop of application

The following equations describe the constraint functions that are applied to create the feasible region and by this fulfill the requirements. The $P_{\min}$ and $P_{\max}$ values defined in equation (2) refer to the price range in which the sum of the courses must move. In the equation (3), the duration limits of the selected courses are determined in the sum.

$$\sum_{i=0}^{M}(p_{r_i}b_i) \geq P_{min} \qquad \sum_{i=0}^{M}(p_{r_i}b_i) \leq P_{max} \tag{2}$$

$$\sum_{i=0}^{M}(d_{u_i}b_i) \geq D_{min} \qquad \sum_{i=0}^{M}(d_{u_i}b_i) \leq D_{max} \tag{3}$$

The constraint given in the equation (4) on the left determines the maximum number of courses included in the solution. On the right side, s indicates the knowledge groups included in the selected courses. The groups in the selected $Prof$ profile must be included in the course group, and s can also contain redundant courses.

$$\sum_{i=0}^{M} b_i \leq C_{max} \qquad \sum_{i=0}^{M}(s\, b_i) \geq Prof \tag{4}$$

The decision loop is the last part of the application. As the user has the option to demand multiple solutions, the loop is required to fulfill settings if possible. The logic of the whole process is displayed in figure 1. In short, decisions are based on multiple factors as whether the first solution was acquired, if the answer is yes, then the cycle repeats. This time if no solution is produced secondary loop is triggered and already used courses, are extracted from the solution list and are appended gradually back to the set used by solving algorithm. This allows the algorithm to find different combinations and terminate itself after all courses from the solution set are depleted. In the opposite case, where the solution was found again, the first cycle repeats itself until it eventually falls into the second loop as well or it runs out of required solutions first.

C. Implementation Details

The application is implemented in Python 3.9.7 programming language.

The choice is due to the wide support of optimization libraries. Regarding different libraries on which whole mathematical functions are based, it offers three most popular packages: Pyomo [6], Scipy [7], and Pulp [8]. Each library uses different syntax to achieve similar functions and they also bring unique functionalities. On their own, these libraries only offer a bridge between the mathematical format and the actual solver itself. What it does, is that it processes given input and relays data in a specific format to the chosen solver.

Scipy [7] is a well-supported module by the community but not by many solving packages. Another disadvantage is its input format acceptability. Scipy cannot work with dictionaries and only calculates with two-dimensional matrix input.

Probably the most widespread is Pyomo [6] library which includes slightly more additional functions like non-linear optimization and can even perform calculations with object-oriented models. Unlike Scipy, Pyomo also allows to presentation of the input data directly from databases, dictionaries, or tables. The repertoire of solver packages is similar to that of Pulp and their syntax is also somewhat similar.

Pulp [8] supports a wide amount of solver packages and also allows binary optimization which is crucial for this article. Besides that mixed integer programming is supported. Pulp has of all libraries easiest problem implementation and with that also comes a very short booting time when the solver is initialized. Pulp is the chosen library for this project and is currently in use together with multiple solvers. The main solver in use is currently CBC (Coin-or branch and cut) [5], but tests were conducted as well with HiGHS [9] and GLPK [10]. We tested all these three solvers in our work. To do so, we use the pulp library as it provides linking and data transfer to individual packages. All discussed solvers are open-source. Multiple commercial solvers would probably outperform all three of those presented, but no tests were conducted on them since we are aiming to implement and release the CSP application under an open source software license.

IV. EXPERIMENTAL RESULTS

In this section, the solver packages, which are potentially candidates for our use in the CSP application, were subjected to testing on the used data sets. The results are summarized in this section and are displayed in Figure 2.

Solver CBC is the default package for Pulp and can be used with Pyomo as well [5]. For MIP problems is applied branch-and-cut algorithm which was discussed at the end of section III. Pulp had previously used GLPK as the default solver, but it was replaced by CBC due to better performance. Another solving package is HiGHS which is still in early development but offers promising results [11]. In addition to MIP, it also allows for solving quadratic programming models and everything can be solved either in series or in parallel.

Figure 2 shows that tests were conducted on two sets represented by 44 and 59 solvers on horizontal axis and time in seconds on vertical axis. The reason why a solution was found faster in a larger set is that fewer infeasible solutions were computed because there were more courses to choose from. Requested amount of combinations was three, but as it was more difficult to achieve this with less courses, actual number of computations was around nine. In contrast with a larger set it was only three, which was demanded. It is important to note that the data on the HiGHS package is not very accurate. There was a relatively frequent failure in the calculation, which was manifested in up to 36.84 % of cases in the first set and around 16.6 % in the second set. If we ignore this, HiGHS does have the best performance, but probably because it is still under development. It has frequent crashes that make it impossible to use this package normally [9]. Despite this, the HIGHS package proved to be the fastest, beating CBC by a hundred milliseconds at the most distinct point. Both solving packages far outperformed GLPK. This, and also its reliability is the reason why CBC was deployed as a solution package. Also GLPK allows only fairly short variable names, and along with HiGHS, have problems with special characters in them.

During tests, average memory usage was around 73,5 MB and GPU usage moved between 3,5 % – 6,7 % on Intel Core i5-6400 processor.

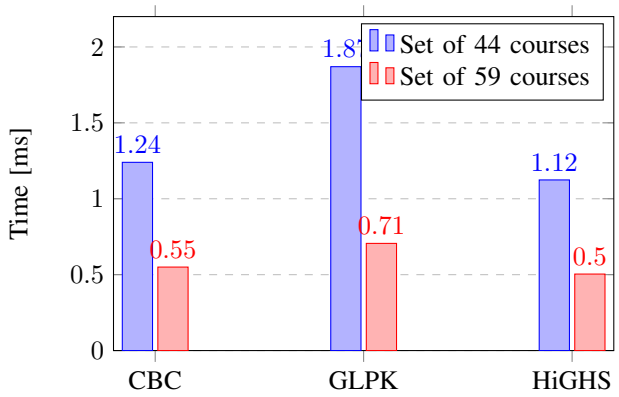


Fig. 2. Performance comparison of different ILP solver libraries

V. CONCLUSION

In this paper, we design, implement, and experimentally test the CSP application computing core. This core allows for identifying which professional training courses are recommended for a certain work role. Methods used for achieving desired results have proven to be efficient and remarkably quick.Linear programming with the simplex algorithm is largely used in different libraries allowing a comparison of their features. The Pulp library was fully compatible with the data format, and broad solver support enabled concurrent testing with promising results supporting the choice to deploy CBC in the final version of the application. The decision cycle seems to be sufficient but could provide more complex decisions in the future. Fundamental requirements for the module were met as it can match courses to profiles based on skill groups. There is room for improvement in the future in terms of more accurate matching. Scales could be implemented to properly display the value of each course and not to consider them all equal in terms of usefulness. More precise searching could be deployed as well because each course contains individual skills and knowledge which are not represented by skill groups in detail.

ACKNOWLEDGMENT

The following funding source is gratefully acknowledged: the ERASMUS+ programme of the European Union (grant 621701-EPP-1-2020-1-LT-EPPKA2-SSA-B 'REWIRE').

REFERENCES

[1] European cybersecurity skills framework [online]. Accessed 2022-12-09. URL: https://www.enisa.europa.eu/topics/education/european-cybersecurity-skills-framework.

[2] Petr Dzurenda and Sara Ricci. Mapping the framework to existing courses and schemes [online]. Accessed 2022-12-01. URL: https://rewireproject.eu/wp-content/uploads/2022/11/REWIRE_R3.4.1_Deliverable-v7-Final.pdf.

[3] Kazuo Murota. Linear programming. In *Computer Vision*, pages 760–766. Springer, 2021.

[4] S. Skiena Steve. In *The Algorithm Design Manual*, volume 2, pages 266–414. London: Springer London, 2008.

[5] John Forrest. Introduction to cbc [online]. Accessed 2022-11-16. URL: https://coin-or.github.io/Cbc/intro.

[6] William Hart and Jean-Paul Watson. Pyomo model documentation [online]. Accessed 2022-11-25. URL: https://pyomo.readthedocs.io/en/stable.

[7] Eric Jones, Pearu Peterson, and Travis Oliphant. Scipy model documentation [online]. Accessed 2022-11-25. URL: https://docs.scipy.org/doc/scipy/dev/index.html#scipy-development.

[8] J.S. Roy, S.A. Mitchell, and F. Peschiera. Pulp model documentation [online]. Accessed 2022-11-13. URL: https://coin-or.github.io/pulp.

[9] Julian Hall, Ivet Galabova, and Michael Feldmeier. Highs package documentation [online]. Accessed 2022-11-24. URL: https://ergo-code.github.io/HiGHS/.

[10] Andrew O. Makhorin. Glpk - the gnu linear programming kit [online]. Accessed 2023-03-08. URL: https://www.gnu.org/software/glpk/glpk.html.

[11] Maximilian Parzen, Julian Hall, Jesse Jenkins, and Tom Brown. Optimization solvers: the missing link for a fully open- source energy system modelling ecosystem [online]. Accessed 2022-11-18. URL: https://zenodo.org/record/6534004.

Network scanner for PowerShell

Dominik Sabota
Brno University of Technology, FEEC, Czechia
E-mail: xsabot00@vut.cz

***Abstract*—This paper describes the proposal and development of a network scanner utilizing Windows PowerShell. The scanner is based on PowerShell's built-in networking modules and functions that supports various scanning options such as target enumeration, basic port scan and host discovery. The scanner's effectiveness is evaluated through various test scenarios, and its performance is compared with existing scanning tools. The developed network scanner is an useful tool for penetration and red-teaming tests realization in order to scan Windows-based company network.**

***Index Terms*—network scanner, cybersecurity, red teaming, ethical hacking**

I. Introduction

Network scanning is an integral part of information technology (IT) security. In today's world, where cyber threats are becoming more sophisticated, it is essential to have a robust and reliable network that can withstand attacks. Network scanning is a critical process that enables businesses and organizations to identify potential vulnerabilities in their network infrastructure, and take appropriate measures to mitigate them. Network scanning involves the process of identifying and extracting information from available devices connected to a network. The primary objective is to recognize available services, detect and identify security and filtering systems, evaluate security, and maintain the functionality of the system of the device. Cybersecurity experts use network scanning as the first step in conducting cyber attacks, penetration testing, or vulnerability mitigation. [1]

II. State of Art

Network scanning plays a critical role in both **red-teaming** and **penetration testing**. Red-teaming involves emulating a cyber attack by an external team of experts to test an organization's ability to detect, respond to, and prevent such attacks. During this process, network scanning is used to identify vulnerabilities and weaknesses that can be exploited during an attack. By mapping out the organization's network infrastructure, red team members can identify potential targets for exploitation and determine the most effective attack vectors. Similarly, penetration testing involves simulating an attack on an organization's network infrastructure to identify vulnerabilities and weaknesses that could be exploited by an attacker. Network scanning is an essential component of the penetration testing process, as it helps identify potential targets for exploitation and provides valuable information about the network's architecture and configuration.By using network scanning tools, red team members and penetration testers can gain a comprehensive understanding of an organization's network infrastructure, identify potential vulnerabilities, and develop effective strategies to mitigate them. Ultimately, this helps organizations strengthen their network security posture and protect themselves from cyber threats. [2]

Nmap (Network Mapper) is a well-known free and open-source tool used for mapping and extracting data about devices connected to a network. It offers a wide range of use cases in the field of network security, enumeration of available devices in the network, or monitoring the behavior of devices and active services. Originally designed for fast and efficient mapping of large networks, Nmap provides full functionality even when scanning smaller networks or individual devices. In addition to the functionalities available via the command-line interface, it is also possible to use the GUI (Graphical User Interface) Zenmap, which is included in the collection of additional utilities of the tool. Nmap is available on all major operating systems, including Linux, Windows, and MAC OS X. [1]

Angry IP Scanner is a memory-efficient, yet comprehensive scanner known primarily for its exceptional speed and ease of use. In addition to basic functionalities such as determining the availability of devices connected to the network using the ping function, obtaining the MAC address of the device, or identifying active ports, the tool offers several options for expanding these functionalities or multiplying the amount of information obtained by using various existing additional plugins and functions, or creating custom functions in the Java programming language. Its unusual speed is due to the implemented process parallelization, which in this case means creating a separate thread for each scanned IP address. Angry IP Scanneris open-source and freely available on every operating system without the need for installation, simply by copying and transferring it to the desired device. [3]

Hping3 is a free and open-source tool that provides advanced functionalities for network testing and analysis. It allows users to create customized packets and send them to network devices, helping in various tasks such as firewall testing, port scanning, and even file transfer between devices with strict firewall rules, by encapsulating the transmitted data into allowed network protocols, and many others. It supports a range of protocols, including TCP, UDP, ICMP, and RAW-IP, and can be used for various purposes such as testing network performance, identifying vulnerabilities in network devices, and detecting the presence of firewalls and IDS (Intrusion Detection System). It can also be used for packet crafting and testing, as well as for creating custom scripts to automate

network testing. Hping3 is freely available on Linux, Mac OS, Windows, and Solaris operating systems, and its command-line interface makes it highly flexible and customizable. [4]

The requirements for creating a custom network scanner were to develop a scanner that can run on devices with older software and older versions of PowerShell, and provide all the features and capabilities of the mentioned scanners that cannot be used on older versions of Windows OS and .NET.

A. Own contribution

Main contribution is own proposal and development of network scanner that is built for Powershell 5.1. Through a thorough comparison and analysis of existing network scanning tools. The scanner was designed to meet the unique requirements of security professionals that realizing the red-teaming testing. Our goal was to ensure that the scanner is highly efficient, scalable, and easily customizable. The resulting tool offers IT professionals a lightweight network scanner that delivers critical security information to help mitigate the risk of cyber threats. The proposed diagram of the structure of the scanner, which illustrates how the program operates and the relationship between certain functions, is shown in (Fig.1).

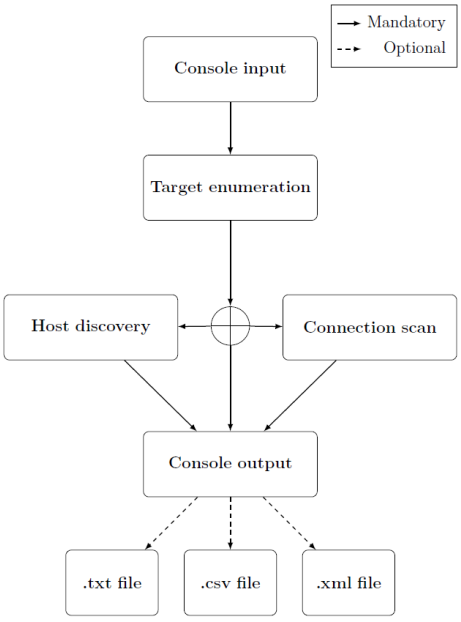


Fig. 1. Scanner structure diagram.

In addition to its current capabilities, this network scanning tool could benefit from future extensions, such as the implementation of a **Stealth scan** method. The Stealth scan aims to avoid detection by minimizing the number of packets sent to the target device. By not completing the *Three-Way Handshake*, the Stealth scan can help prevent the target device from detecting the scan and potentially blocking it. Instead, the **Stealth scan** uses techniques to reduce the number of packets sent to the target, waiting for a response from the device.

Another potential enhancement is the addition of **OS detection** and **service version detection** capabilities. These features could provide more detailed information about the target device, allowing for a more comprehensive evaluation of its security posture. The OS detection capability would identify the operating system being used by the target device, while the service version detection would provide information about the versions of specific services running on the device.

Additionally, implementing **multithreading** could significantly improve the tool's speed and efficiency. This feature would allow for multiple scans to be performed simultaneously, reducing the time required to complete a scan and enabling the processing of a greater number of targets.

Overall, these future extensions could greatly enhance the tool's performance and effectiveness, providing security professionals with a more efficient and thorough means of identifying and assessing potential vulnerabilities within their network infrastructure.

III. Proposal of network scanner tool

A. Target enumeration

Determining the devices to be scanned is an essential and indispensable phase of a network scanning tool. It involves the determination and creation of a list of target IP addresses, processing the received data passed as a parameter to the function.

The function will accept IPv4 address strings separated by commas and written in the following ways:

- A specific standalone IP address.
- Specifying IP addresses by listing the desired values of the last octet, separated by commas.
- Specifying a range of values for the last octet using a hyphen between the desired starting and ending values of the IP address's last octet.
- Defining a network range by specifying the CIDR prefix value after the slash at the end of the desired IP address.

By parsing the received string, the function will evaluate how the IP addresses were entered and write each IP address separately on a new line in an array, thus creating an easily usable list of IP addresses. This function is always executed at the beginning of other functions, and it is also possible to run this function separately to output the desired IP address list.

B. Host discovery

The phase of determining available devices on the network involves verifying the availability of devices by sending ICMP Echo requests (ping). The device is considered available if the source device receives a response from the target device. Conversely, it is considered unavailable if the source device does not receive any response before the timeout period expires.

The function performing host discovery will accept IPv4 address strings passed as a required parameter, which will be

processed and stored in an array variable using the Target enumeration function. Subsequently, it will iterate through the array of IP addresses and verify their availability in the manner described above. The output of this function will be a table and an additional message regarding the number of unavailable devices verified.

C. TCP Connection scan

One of the primary advantages of using a network scanning tool that utilizes the *Three-Way Handshake* method is that it allows for efficient and effective port scanning. During the scanning process, the tool attempts to establish a TCP connection between the source device and the target port, which confirms the availability of the port if successful. This approach provides valuable information to security professionals about the target device and its potential vulnerabilities, allowing for a more thorough evaluation of the target's security posture.

The TCP Connection scan function will accept IPv4 address and port number strings as parameters. The received IPv4 address strings will be processed using the Target enumeration function, just like in the Host Discovery function. It will then proceed to iterate through the array of IP addresses and attempt to establish a connection with the corresponding ports for each IP address. The output of this function will be a table along with an additional message indicating the number of inaccessible scanned ports. [5]

In addition to displaying the results of the functions in the console in the form of a table, it will be possible to save the output in files of various formats by specifying an optional parameter and a path where the file should be created. By selecting the appropriate parameter, a file of the corresponding format will be created, into which the obtained results will be written. The offered file formats are .txt, .csv, and .xml.

IV. Testing of implenetation

Using suitable virtualization applications such as VMware Workstation Player, virtual machines will be set up for the testing. IP addresses of the vulnerable test devices will be determined and tested for availability using the *ping* command.

The Nmap tool on Kali Linux will be utilized to test selected functionalities against vulnerable devices, with information recorded. The same process will be performed using the proposed tool in Windows 11, and the outputs of both tools will be compared. In the (Fig.2) is shown example output of implemented scanner taken from testing environment.

It should be noted that the network scanning tool has been designed to not require any external libraries, making it a more flexible and portable solution. The lack of external dependencies reduces the risk of compatibility issues and security vulnerabilities that could be introduced by external libraries. The tool is being implemented specifically for older versions of the .NET framework to enable its use from a remote or zombie host that may not have the latest version of .NET Core available.

Thorough testing will be conducted on Windows 11 to ensure the scanner's compatibility and portability, utilizing both the older version of PowerShell and .NET framework and the newest tools for analyzing and testing the scanner's functionality. The implementation of this scanner for older versions of the .NET framework, along with thorough testing on Windows 11, will ensure its compatibility and portability, enabling security professionals to utilize it effectively in a range of different environments.

```
Starting TCP Connection Scan at 12/09/2022 23:24:35

31 ports are closed or filtered.


Host            Port Status
----            ---- ------
192.168.111.130   80 Open
192.168.111.130   21 Open
192.168.111.130   22 Open
192.168.111.132   80 Open
192.168.111.132   22 Open


Scanning done: 9 IP addresses scanned.
```

Fig. 2. Scanner output example

V. Conclusion

The purpose of this project was to develop a custom network scanning tool in the form of a usable module for the PowerShell scripting language in version 5.1. The tool offers basic network scanning functionalities and provides a structured report of the results. The project began with an overview of existing network scanning tools, including their functionalities and the operating systems on which they are available. Then, a tool structure was defined using a diagram, outlining the functions to be implemented and the virtual machines used for testing. Finally, the testing process was described to ensure the tool's functionalities were effectively validated. The successful completion of this project demonstrated the feasibility of developing a custom network scanning tool and its potential value in network security.

References

[1] LYON, Gordon Fyodor, Nmap Network Scanning: Official Nmap Project Guide to Network Discovery and Security Scanning". Sunnyvale, CA 94086-6161 United States: Insecure.Com, 2009. ISBN 978-0-9799587-1-7.

[2] Jacob G. Oakley, Professional Red Teaming: Conducting Successful Cybersecurity Engagements (1st. ed.), APress, 2019

[3] KEKS, Anton, Angry IP Scanner [online], https://angryip.org/documentation/.

[4] Linux man page [online], https://linux.die.net/man/8/hping3.

[5] FOROUZAN, Behrouz A., TCP/IP protocol suite. (4th ed.) Boston :McGraw-Hill Higher Education, 2010. ISBN 978-0-07-337604-2.

System for detecting connected devices to the power grid

1st Radim Macho
FEEC, Brno University of Technology
Brno, Czech Republic
221666@vut.cz

2nd Petr Musil
Department of Telecommunications
FEEC, Brno University of Technology
Brno, Czech Republic
musilp@vut.cz

***Abstract*—This paper discusses the possibility of detection devices connected to the power grid. The detection will be performed using data obtained from power-line communication modems connected to the monitored power grid. The data will then be processed using machine learning-based code.**

***Index Terms*—frequency spectrum of the noise, machine learning, noise, power grid, power-line communication**

I. Introduction

On a standard power grid, there is not only standard 50 Hz 230 VAC sine wave signal, but also noise. This noise can be caused by devices connected to the power grid or inductive coupling between electrical network conductors, where originally non-noise signals became noise on other conductors thanks to inductive coupling.

A. High harmonic

Despite the fact that all electrical devices connected to the power grid have some level of influence on the power grid itself, not all electrical devices generate noise on a level that can significantly affect the power grid. Devices that have the most significant impact on the power grid are devices that generate high harmonic [1], such as:

- switched mode power supplies,
- electronic fluorescent lighting electromagnetic control-gear,
- uninterruptible power supplies unit.

Because these devices are among devices that are usually connected to the power grid, there is always some amount of high harmonic noise in the power grid that can cause problems [1]. Problems usually caused by high harmonics are:

- unwanted tripping of circuit breakers - circuit breakers, based on comparing current in phase and neutral, can be negatively affected by high harmonics that unbalance the current in a circuit and can cause unwanted tripping of circuit breakers [1].
- skin effect - because alternating current tends to flow on the surface, the skin effect creates additional loss and heating. Skin effect is more significant on higher frequencies caused by higher harmonics [1].
- overloading of neutrals - in a three-phase electrical system, each phase is shifted by 120° from each other. This makes the combined current on neutral equal to zero when each phase is equally loaded. High harmonics unbalance the load of phases and make the current on neutral, not zero. This unbalance leads to overstressing neutral and generating heat that can lead to fire [1].

B. Power-line communication

Power-line communication, hereinafter referred to as PLC, is a technology that uses the power grid to transfer data. The main principle is based on modulating high-frequency signals representing data to transfer on top of the power grid 50 Hz signal. Frequencies used to modulate data are between 2 MHz and 250 MHz [2].

For the implementation of the power-line communication are required at least two PLC modems, one modem for each point in space that you want to connect. The first PLC modem working as a transmitter is at one side of the PLC connection topology. PLC modem gets data for transmission from a data source. These data are converted into a high-frequency representation and modulated on top of the power grid 50 Hz signal. The second PLC modem connected on the other side of the topology, which works as a receiver and monitors the power grid signal. When a modulated signal from the transmitting modem is found, the second PLC modem demodulates that signal and hands over data to the data destination. In real scenarios, PLC modems can work both as a transmitter and receiver to enable two-way communication [2].

The main advantage of power-line communication compared to Wi–Fi and standard Ethernet cable is that it can connect two speared spaces without the need for additional cabling but still offers fast data transfer rates [2].

The disadvantage is that speed of power-line communication is affected by the power grid to which the modems are connected, such as age, wear, and other devices connected to power grid that generates noise [2].

C. Multiple Input Multiple Output

Multiple Input Multiple Output, MIMO for short, is a commonly used method in PLC technology. This method allows the measuring system to input more input parameters in measured topology and measure more outputs from measured topology [3].

PLC technology uses the MIMO method to create several communication channels between three conductors in a standard wall plug. The standard channel for power-line communication is phase–neutral, but other channels can be used, such as neutral–earth [3].

D. Machine learning

Machine learning is a term for a programming method for developing algorithms and models which allows the program to augment its internal parameters and improve its results without needing to be edited by a real person. This augmentation of internal parameters is based on data that are provided. More data provided leads to better results. Machine learning is currently used, for example, for data prediction, picture content recognition, or for choosing featured content [4].

Machine learning can be divided into three main categories based on learning principles:

- supervised learning - data in the training dataset are labeled with the right results. The algorithm learns to find patterns and assigns them to labeled results. This category is used, for example, in image recognition [4].
- unsupervised learning - data in the training dataset are not labeled. The algorithm search for patterns and relationships but does not assign them to specific results. This category is used, for example, in pattern searching [4].
- reinforcement learning - this method is infinitely learning based on feedback from an external source by giving rewards or penalties. This category is used, for example, in choosing featured content on social media [4].

The field of application of machine learning is constantly growing, and it can be assumed that machine learning will be used even more in the future.

II. Principles of Detection of Connected Devices

A. Measuring the change of frequency spectrum of the noise over a long period of time

Because the frequency spectrum of the noise of the power grid is constantly changing, an experiment for measuring over time was created. Measuring was done instead in the laboratory in the real environment on a flat inside a panel house. Measuring was done on basic topology where one PLC modem was connected to a plug near the flat switchboard. A second PLC modem was connected on an extension cable connected to a plug inside the main room. The topology connection between PLC modems is unknown because it is hidden inside the walls. This topology connection is closest to the real scenario of using PLC modems for data transmission.

Measurements were taken from 7:00 morning to 22:00 at night. Data was measured every hour on a PLC modem connected to a plug inside the main room. The results of the measurements are plotted for better readability into the spectrogram in Fig. 1.

In the Fig. 1 is visible that the frequency spectrum of the noise is relatively stable over time. Despite the fact that no device was connected, disconnected, switched on, or switched

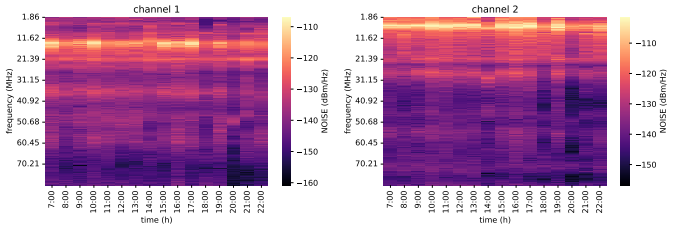


Fig. 1. Spectrogram of measuring the change of frequency spectrum of the noise over the day.

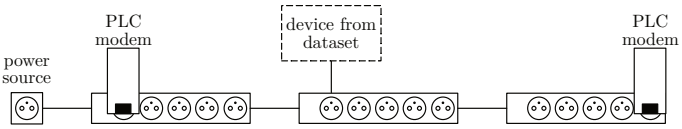


Fig. 2. Topology for measuring the frequency spectrum of the noise of devices in the dataset.

off, except the laptop that was used for collecting measured data and was switched on from beginning to end of measuring, there is a visible change in the frequency spectrum of the noise over the day.

The change in the frequency spectrum of the noise of topology during the day is caused by external influences on the power grid. External influences can be neighboring flats where the electrical devices were used normally. Devices that change their character over time. An example of this type of devices is a fridge which turns the compressor on and off for maintaining the internal temperature.

Based on the experiment results, the change in the frequency spectrum of the noise of the power grid over time is recognizable. Still, the change in the frequency spectrum of the noise of the power grid over time is not that significant, so the change in the frequency spectrum of the noise affects measuring the dataset and following unknown data recognition.

B. Creating dataset of devices

Based on the fact that every device connected to the power grid has some level of influence on the power grid itself, and devices generating high harmonic have a more significant influence on the power grid, a dataset has been created.

The dataset is mainly created from different types of light bulbs, switched mode power supplies, and uninterruptible power supply units. But some interesting other devices were included, such as electricity meters or coffee grinders.

A basic topology has been created for connecting two PLC modems using extension cables, representing a power grid, where one PLC modem is near the power source, standard plug in the wall, and the second PLC modem is located at the end of the extension cables as is shown on Fig. 2. Devices from the dataset are connected in the middle between PLC modems. This topology is used for measuring the frequency spectrum of the noise of devices in the dataset.

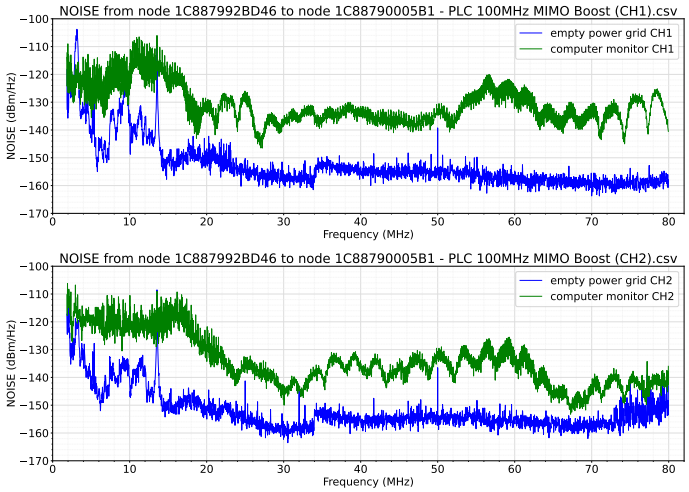


Fig. 3. Frequency spectrum of the noise of empty power grid and with the connected computer monitor.

The PLC modems measure noise on two channels using the MIMO method. Each channel is measured in the frequency range between 1,8 MHz and 80 MHz. For this range are taken 3200 samples for each channel. The representation of measured data is shown in Fig. 3.

As shown in Fig. 3, blue curves represent the power grid with only PLC modems connected and without any other device. Green curves represent the power grid with a connected computer monitor. The computer monitor frequency spectrum of noise significantly differs from the empty power grid. This difference is present because computer monitors belong to devices that strongly influence the frequency spectrum of the noise of the power grid.

For simulating real usage of PLC modems on the power grid, there is an operating program iPerf3 that measures the bandwidth of the network. This measuring represents data traffic between PLC modems.

Three measuring scenarios were created by changing the connection to the power source of the topology in Fig. 2. These scenarios were measured on custom–created topology with extension cables in the laboratory.

1) adding 50 m extension cable, isolation transformer and another 50 m extension cable between main topology and power source. Long extension cables and isolation transformer significantly reduce noise from the power source. That is the reason why this topology provides the clearest data of the frequency spectrum of the noise of a specific device.
2) direct connection topology to the power source. Because this topology is directly connected to the power source the noise from the power grid has a strong influence on the frequency spectrum of the noise of measured topology. This topology has the worst data for detection by the machine learning model but the data resemble the most the real topology.
3) adding 50 m extension cable between main topology and power source. Data measured on this topology have less noise than topology 2 but more than topology 1. This makes data from this topology a great combination of noise from the power grid and the influence of connected devices on the frequency spectrum of the noise.

For each scenario, each device from the dataset was frequency spectrum of the noise measured twenty times to achieve a properly working machine learning model.

When creating a topology in the real world is very important to pay attention to inductive coupling between extension cables, especially with the long 50 m extension cables. For this reason, care was taken to distribute the extension cables over the laboratory floor with even spacing. For the same reason, the crossing of extension cables was prevented.

C. Data processing

Collected data are not used directly for machine learning because each measurement contains 3200 samples for each channel, which makes 6400 samples for each measurement.

As mentioned in chapter II-B, different devices have a different frequency spectrum of the noise shape. These differences can be used for device recognition, but not all devices have such a distinctive shape as a computer monitor. Moreover, measured samples include noise occurring on the power grid. As shown in Fig. 3, there is no curve with a smooth shape.

For this reason, data are evenly split into shorter, equally sized intervals, and then they are statistically defined by the following parameters:

- average value
- minimum value
- maximum value
- most frequently occurring value (mode)
- middle value (median)
- standard deviation

This way, statistically defined data are used for training the machine learning model and the following evaluation of the accuracy of the final machine learning model.

D. Machine learning model

For the entire data processing, and following creation of the machine learning model is used programming language Python as currently the best programming language for creating machine learning based programs.

For creating a machine learning model is used CatBoost which is an open-source library based on gradient boosting on decision trees [5]. CatBoost provides an easy implementation of the machine learning model and a large number of parameter customization for training.

For a properly working machine learning model training dataset is needed, and for evaluation of model precision is needed another dataset that should contain different data than the training dataset. Additionally can be used the third dataset to evaluate model precision during model training. For these reasons, statistically defined data from all three scenarios are mixed together and then separated into three sets, one for the training dataset, one for evaluation of final

model precision, and one for evaluation of model precision during learning training.

Based on the size of the dataset, over seventy output states are defined for the machine learning model. To make the machine learning model accurate twenty samples are used for train dataset, twenty samples are used for evaluation dataset that is used during learning cycles to improve the final accuracy of the machine learning model.

After learning cycles, the cycle with the highest data detection accuracy on the evaluation data is selected as a final machine learning model. This final machine learning model is used for unknown data recognition represented by the last set that has not been used in the learning process, so these data are completely unknown for the final machine learning model.

III. Conclusion

Two final machine learning models were made.

The first one had an edited dataset for learning, evaluating during the learning process, and final data recognition compared to chapter II-D. For machine learning model training was used only data from measuring scenario 1, for evaluating model precision during model training was used only data from measuring scenario 3, and for data detection by the final machine learning model was used only data from measuring scenario 2.

The accuracy of data detection prediction was only 20%, but because data from scenario 2, used as unknown data, contains significantly more noise from the rest of the power grid than from scenarios 1 and 3 because the main topology is directly connected to the plug in the wall. Scenarios 1 and 3 contain less noise because of long extension cables and an isolation transformer. These extensions reduce unwanted noise from the rest of the power grid, and the influence of connected devices is more visible. These differences in data lead to low data detection accuracy.

However, an interesting phenomenon occurred, some devices have such a significant effect on the frequency spectrum of the noise of the power grid that the final model recognized without needing to include data with more noise from scenario 2.

The second model had a dataset as described in chapter II-D. The accuracy of data detection prediction was over 86%, which is an excellent accuracy value. The majority of false detection occurred with devices measured at different levels of switched-on state, where the final model correctly recognized what kind of device from the dataset was connected and that the device was switched on but incorrectly recognized the specific setup of the device.

Suppose reduce desired data accuracy only to detect which type of device is connected, e.g., all light bulbs count as one device and the same for coffee grinders. The accuracy of unknown data detection is 98%.

This leads to the conclusion that it is possible to detect which device was connected to the power grid based on the device frequency spectrum of the noise if the device was measured on the power grid on which it should be detected.

References

[1] CHAPMAN, David: *Application Note – Harmonics: Causes and Effects.* Brussels: European Copper Institute, 2011.

[2] SLÁČIK, Ján a Zdeněk KOLKA: *Síťová komunikace po napájecím vedení.* Brno, 2016.

[3] SCHWAGER, Andreas, Daniel SCHNEIDER, Werner BÄSCHLIN, Altfried DILLY a Joachim SPEIDEL: *MIMO PLC: Theory, measurements and system setup. IEEE International Symposium on Power Line Communications and Its Applications.* 2011, 2011, 48-53. Available from: doi:10.1109/ISPLC.2011.5764447

[4] MAHESH, Batta: *Machine Learning Algorithms. 9th.* Gujarat: International Journal of Science and Research, 2018. ISSN 2319-7064.

[5] *CatBoost Documentation. CatBoost [online]. Yandex, 2023*. Available from: https://catboost.ai/en/docs/

Design of an integrated energy self-sufficient power supply for residential building consumption

Jan Klouda
VUT
FEKT
Brno
214380@vut.cz

***Abstract*— The thesis deals with the design of an integrated self-sufficient system for control and power supply of a residential house. The aim of the thesis is to develop a design for this energy self-sufficient power supply and to evaluate the proposed system from social, economic, and technological points of view. In this thesis, the different technologies for the energy self-sufficient house are evaluated and then implemented in the design to achieve the highest economic and energy efficiency. The resulting system design is then simulated in a predefined residential house in a selected location in the Czech Republic.**

***Keywords*— Self-sufficient house, renewable energy, batteries, photovoltaics electrical installation, power engineering**

I. Introduction (*Heading 1*)

This thesis deals with the design of an energy self-sufficient house for a specific family and the needs of this family. The work begins with the construction of the house, where not only the materials are carefully selected, but also the shape of the house itself. To achieve the best results, the construction of a passive house is chosen. Subsequently, a location was chosen for the possible construction of an energy self-sufficient house. Based on the location and the needs of the occupants of the house, the ideal system for electricity generation and storage is applied. Once the technology is selected, further components such as inverters, charging electronics, and others are selected from domestic stores [1][2][3][4].

Data from the Czech Hydrometeorological Institute helped to size the entire system directly to the site, and the design of the house was thus subject to maximum efficiency and the chosen power generation technology. The result of this work is therefore the design of an electrically self-sufficient house in the given location. An electrically self-sufficient house would therefore be able to meet the electricity consumption of the occupants of the house, using only its own electricity generation, and thus provide comfort every day of the year. The advantage of the proposal will be its flexibility, where at the end of the life of each component of the electrical system, this component will be replaced by a new piece (even from another manufacturer) with the same electrical characteristics. This will give the house a very long theoretical lifetime [3][4][5].

II. Design of the House

The house is designed in open space. The roof is situated on the south side (to achieve the best gains from PV (photovoltaic panels). On the ground floor there is a bathroom, a utility room and a living room connected to the kitchen. In the living room there are three larger windows and a balcony door to achieve heat gains from the sun during the winter period. In the attic there are two bedrooms and a hallway with a washing machine. The entire south roof is covered with PV panels with an output of 18,5 kWp. The design of the home can be seen on Fig. 1.



Fig. 1. Design of the house

A. *Choosing the right materials for a passive house.*

Nowadays, there are a lot of materials for passive house on the market. One of the most common materials are lime-sand blocks from Ytong and baked, sanded, insulated bricks from Heluz. Both of these materials (namely Ytong Lambda YQ, and Heluz Family 2in1) have the same thermal energy storage value and, with a masonry thickness of over 450 mm, meet the conditions for the construction of a passive house. I chose insulated bricks from Heluz (Family 2in1) for their cheap price, availability in the Czech country and excellent thermal insulations. This material also leads in heat. The materials for the base slab are not much different than low-energy and new construction houses. Except for the use of foam glass gravel, which has higher insulating properties than conventional gravel. The materials for the roof here remain high insulation layer with vapour barrier [6][7].

B. *Composition of materials to achieve a passive house*

The materials chosen for the house meet the requirements for a passive house. The reason for this is to reduce heating requirements as much as possible during the winter months. The choice of these materials and their construction details were selected based on the recommendations of the "Centrum Pasivního Domu" (Passive House Centre). In the PHPP software, these materials were assembled into detailed designs. The PHPP software is a tool directly for calculating the energy balance and design of passive houses and thus helped to design also the heat recovery unit for ventilation, calculating the annual

energy gains from electrical equipment and windows. Using this software, the building envelope was designed and the result is the passive house design created [8].

Although the PHPP software can simulate and calculate the energy performance of a building and classify it as a passive house, the energy performance of the building in every hour of the year was needed for detailed calculations that would later be used for the design of the self-sufficient system. For this purpose, the Energie 2023 EDU software was chosen. This software can calculate the hourly energy needs for the whole year from the entered values of the building envelope, occupancy, type of heating, ventilation and other detailed information about the building. This information can be used to provide the electricity consumption for heating, ventilation and hot water for every hour of the year. This information is also very important for calculating the electrical performance of the building. The result of the energy performance of the building is shown in Fig. 2 [9].

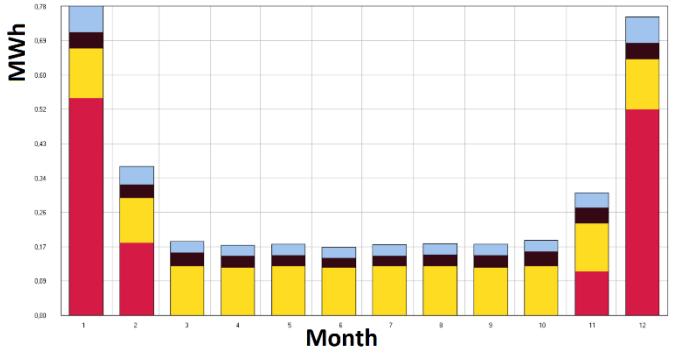


Fig. 2. Simulated monthly energy consumption of the house. Red for heating, yellow for hot water, black for lighting, blue for auxiliary equipment (pumps, etc.)

III. PHOTOVOLTAIC SYSTEM DESIGN

In order to design the PV system, it was necessary to provide enough information that could influence the system. Use the results from previous simulations, and provide a custom profile of the house's electricity use.

A. *Electricity consumption of residents*

To design a PV system for a self-sufficient house, it is necessary to know the electrical consumption of all the users of the house. For this purpose, a smart electric meter was installed in the house of the family under study. This family is the expected future resident of the house. A python program has been developed to read the data from the smart meter and run on a Raspberry Pi microcomputer. This microcomputer is connected via RS485 bus. The created program records the current consumption of the house. This information can be used to determine the peak electrical outputs required to meet the electricity supply and also help create an hourly profile of electricity usage [10].

To make the calculations as accurate as possible, hourly profiles were created for weekdays and weekends separately. An example of an hourly profile of electricity use can be seen in Fig. 3. From these profiles, a profile of hourly electricity values for the whole year was created using correlation. The correlation was made against the predicted changes in consumption over the year available from the PVSOL software. The resulting electricity usage profile is therefore composed of the partial hourly values.

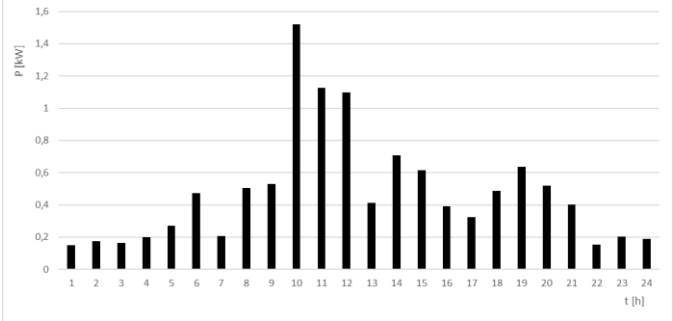


Fig. 3. Figure of hourly electricity consumption of the monitored family during weekend

B. *Combining the results of the house simulations and the electricity consumption of the projected residents*

The electrical performance results of the building were taken from the Energie 2023 EDU software. These simulation results record the hourly electrical energy requirements to provide heating for the building, forced ventilation through heat recovery and hot water. From the measured consumption values of the monitored family (assumed residents), the average hourly electricity consumption values for the whole year were taken.

By combining these two electricity values, a profile of the total electricity consumption of the house was obtained, with all needs provided. Thus, the family would not have to discount their needs and the entire PV system is designed for comfort and reliability.

C. *Selection of photovoltaic system components*

Victron Energy components were chosen for the design of the energy self-sufficient house. This company deals with robust island systems. They are characterised by their reliability, simplicity (easy repairs due to the modularity of the system) but high internal consumption. One Quattro 10 kVA inverter, three SmartSolar MPPT (maximum power point tracker) 250/100 48 V and LFPSmart 25,6/200-a batteries are used for the design. The entire system on the direct current (DC) side is built on 48 V. Due to the modularity, a sufficient number of MPPTs, inverters and batteries can be connected to this 48 V branch [11][12][13].

The photolytic panels were chosen from AEG, type: AS-M1203-H-375. These panels have a power of 375 Wp. 3 panels are connected in series (string). To achieve 18.5 kWp, there are a total of 18 of these strings on the roof. The photovoltaic panels are connected to the inputs of MPPTs [14].

D. *Photovoltaic system simulation*

The entire photovoltaic system was simulated in PVsyst software. This software can simulate islanded PV systems and has a library of Victron Energy products. This library contains all the necessary components with detailed information for the calculation. The electrical profile from previous simulations and measurements, the roof elephant and the azimuth have been entered into the software.

The whole system is based on maximizing the gains from the PV panels, so the house is equipped with up to 18.5 kWp of PV panels. The batteries are undersized relative to the output of the panels to make the resulting system as cheap as possible (in terms of battery and PV panel cost). This puts the entire proposed system into very high levels of unused energy from the PV panels. This was taken into account in the design and the system is thus tuned for the winter months.

IV. Overview of Simulations

In the PHPP, Energie 2023 EDU and PVsyst software, several possible versions of how to achieve self-sufficiency of the house were simulated, along with the different technologies used. To give an example, the energy performance of the house was determined using a ground/water heat pump, air/water heat pump, or using a pallet boiler together with a boiler. In PVsyst, different simulations were created for each of these examples to achieve self-sufficiency. Use a large number of batteries and reduce the number of modules on the roof, achieve full self-sufficiency with zero subordination of the system (so that the residents are not forced to subordinate their lifestyle to the whole system, for example to do laundry and cook only in given situations), or achieve self-sufficiency only partially and use more of a backup source of heat and electricity (see section Backups). The following chapter summarises only one version, which has been selected as the best from an economic, social and environmental point of view.

A. Result of the simulation of a self-sufficient house

The result of all simulations is a building that has no connected external energy sources, is fully self-sufficient, meets the requirements for a passive house, uses mostly ground/water source heat pumps, and the residents do not make any changes to the system (their electricity needs are fully covered and there is no need to change their lifestyle).

The house is powered entirely by photovoltaic panels and energy is distributed by a single 10 kVA inverter. The electrical wiring and equipment in the house is designed for single phase (230 V). Appliances were exclusively selected to be single phase. The inverter is connected to a NIBE S1255 heat pump connected to one phase (1 x 230 V), with the additional power of the electric boiler limited to 2.5 kW (for emergency needs). The lighting is also connected to this inverter and the inverter provides the rest of the electricity needs.Adding the measured peak power of the household and the peak power of the heat pump, the value was 10.32 kW. However, this peak value will not be normal in the system and is just a combination of worst case scenarios (and possible faults). Even at a load of 10.32 kW, the inverter is able to supply this energy in a short time. Due to the replacement of some equipment (new induction cooktop, oven, etc.) this inverter power is sufficient.

A total of 20 LFPSmart 25.6/200-a batteries are used for energy storage (The batteries only use 80% of their capacity, due to battery life). Two are connected in series to achieve the required voltage on the DC side and a total of 10 of these battery strings are used to achieve the required capacity. Three SmartSolar MPPT 250/100 48V connected in parallel on the DC side are used. The PV panels are connected in 18 strings of 3 panels in series. These strings are connected 6-6-6 to the MPPTs. Communication between the devices is provided by the Color Control GX together with the Cerbo GX (from Victron Energy), which also communicates with the VE-BUS BMS (battery management system) unit to ensure gentle charging of the batteries. This device monitors the battery capacity, controls the entire PV system and gives the user an overview of the current status. In Fig. 4 you can see a simplified circuit of the proposed system [15].

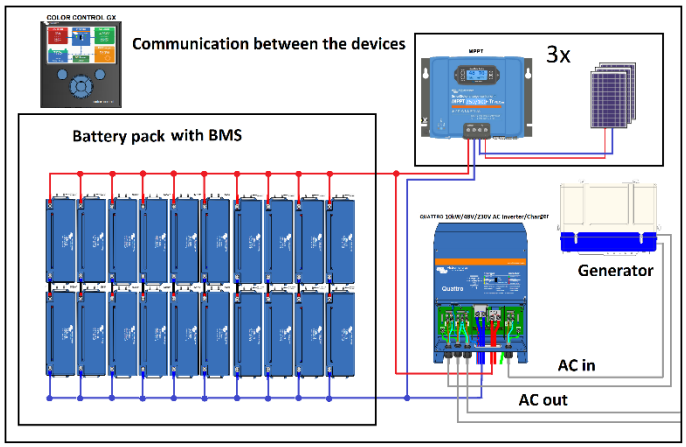


Fig. 4. Simplified photovoltaic system wiring

The proposed system covers the consumption of 3895 kWh of electricity and covers 99.88% of the required electricity (all the losses on the PV system side are included in the design, plus an additional 5% energy reserve). The simulation result of the PV system can be seen in Fig. 5. The missing energy (4.83 kWh) is supplemented by a spare thermal energy source (see section Backups) or by an electrical generator. This loss corresponds to several days (depending on weather conditions) of operation of the backup thermal energy source. It is assumed that this backup source is operated during the winter period when energy is not available. Since the substitute heat source is a wood-burning fireplace, operation is assumed to be during Christmas and other holidays or weekends.

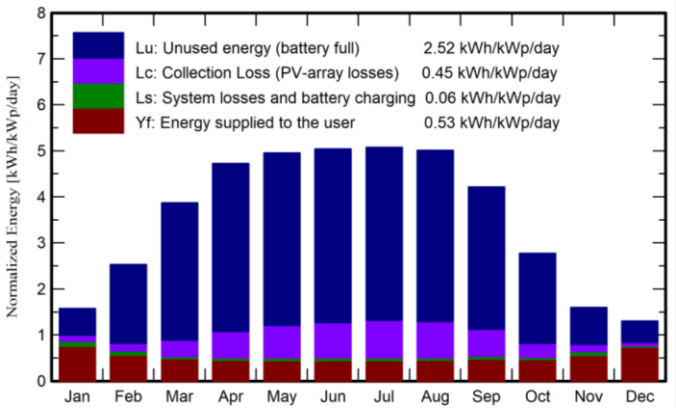


Fig. 5. Simulation result of the photovoltaic system

V. Backups

Because the house is designed for full self-sufficiency, simplicity and reliability are the most important factors. In terms of simplicity, the entire PV system is designed for maximum

modularity. The number of inverters and their output can be varied, as well as the batteries, MPPTs and PV panels. This has the advantage of making the whole system very easy to repair. In terms of reliability, the system is equipped with Quattro 8 kVA and 3 kVA backup inverters. The first inverter is only a backup for the used inverter in case of a failure. The second (less powerful) is a backup to cover the minimum needs of the family and communication in case of failure of all inverters or in case of emergency power shortages.

If a situation arises in which insufficient power is available, the diesel electric central can be manually started and powered, or batteries recharged, from this backup source. All inverters and the entire PV system can also be bypassed (bypassed) and connected directly to the house's consumption.

The thermal energy is backed up by a wood chipped fireplace. The fireplace is located in the living room and is only used during the winter holidays, and some of the weekends, to provide the remaining energy that may be missing from the system during those months. The fireplace is rated for a higher heat output than the heat pump used, so that it can heat the entire house in the event of a heat pump failure. The heat pump has a built-in electric instantaneous heater for hot water or heating. This electric direct heater is only used in the event of a pump failure, or in poor weather conditions (which are not expected in the locality throughout the year). Thus, the heat pump only uses a maximum of 3 kW of power. However, it is not always necessary to achieve this power input. For this reason, the power of the inverter is not used and a power reserve is created [16].

VI. Conclusion

The result of the work is a fully self-sufficient house with a heat pump, heat recovery and a designed photovoltaic system that does not restrict the residents in their daily life. The house is very over-dimensioned in terms of PV panels and in summer it achieves large losses in the form of unused electricity (up to 18616 kWh of total unused energy from PV panels per year). However, the PV system is sized for the winter period where, with the occasional use of a fireplace, it can provide electricity for heating and hot water, while ensuring maximum comfort for the residents. The entire house was simulated in several programs and software, and a custom electrical profile of the residents was created to ensure the highest accuracy of the simulations. The entire electrical system of the house is calculated with a margin and for the worst case scenario that can happen during the lifetime of the house. There are backup generators, inverters and a fireplace in the house. In this way, maximum system reliability and security of power supply is achieved, even with multiple failures. The whole system is designed in a modular way and the number of panels, batteries and the number of inverters can be changed. If the system needs to be expanded in the future.

References

[1] Renewable energy. Africa [online]. [cited 2022-05-02]. Available from: https://afry.com/en/area/renewable-energy?sector=18

[2] Self-sufficient houses. We like to explore [online]. [cited 2022-05-02]. Available from: https://greenliving.lovetoknow.com/Self_Sufficient_Homes

[3] Where to start on a more self-sufficient home. Bioenergyconsult [online]. 2021 [cited 2022-05-03]. Available from: https://www.bioenergyconsult.com/making-your-home-more-self-sufficient/

[4] How you can create a self-sufficient home. Openaccessgovernment [online]. 2019 [cited 2022-05-04]. Available from: https://www.openaccessgovernment.org/self-sufficient-home/59276/

[5] Daily data. Czech Hydrometeorological Institute [online]. [cit. 2023-03-31]. Available from: https://www.chmi.cz/historicka-data/pocasi/denni-data/Denni-data-dle-z.-123-1998-Sb#

[6] HELUZ FAMILY 50 2in1 polished. Izomat [online]. [cit. 2023-03-31]. Available from: https://www.izomat.cz/out/media/58505_05-Technicky-list-CZ.pdf

[7] Thermal insulation blocks Lambda YQ. Stavebniny Janik [online]. [cit. 2023-03-31]. Available from: https://www.stavebniny-janik.cz/files/uploads/Produkty/Ytong/tepelneizolacni-tvarnice-lambda-YQ.pdf

[8] Construction details. Pasivnidomy [online]. [cit. 2023-03-31]. Available from: https://www.pasivnidomy.cz/detaily/

[9] Freely distributable programme Energy 2023 EDU. Department of Civil Engineering, Faculty of Civil Engineering, CTU in Prague [online]. [cit. 2023-03-31]. Available from: https://kps.fsv.cvut.cz/index.php?lmut=cz&part=people&id=52&sub=331

[10] How to communicate with Raspberry Pi 4 B+ using MAX485 module. Industrialshields [online]. [cited 2023-03-31]. Available from: https://www.industrialshields.com/blog/raspberry-pi-for-industry-26/post/how-to-communicate-raspberry-pi-4-b-with-a-max485-module-137

[11] Datasheet Quattro 3kVA 15kVA. Victronenergy [online]. [cit. 2023-03-31]. Available from: https://www.victronenergy.com/upload/documents/Datasheet-Quattro-3kVA-15kVA-EN.pdf

[12] Datasheet SmartSolar charge controller MPPT 250-70 to 250-100 VE.Can. Victronenergy [online]. [cited 2023-03-31]. Available from: https://www.victronenergy.com/upload/documents/Datasheet-SmartSolar-charge-controller-MPPT-250-70-up-to-250-100-VE.Can-EN.pdf

[13] Datasheet 12,8 & 25,6 Volt lithium-iron phosphate batteries Smart. Victronenergy [online]. [cit. 2023-03-31]. Available from: https://www.victronenergy.com/upload/documents/Datasheet-12,8-&-25,6-Volt-lithium-iron-phosphate-batteries-Smart-EN.pdf

[14] High efficiency AS-M1203-H(M6). Aeg industrialsolar [online]. [cit. 2023-03-31]. Available from: https://www.aeg-industrialsolar.de/wp-content/uploads/downloads/solar-modules/datasheets/high-efficiency/half-cut-cells/AS-M1203-H(M6)/High_efficiency_AS-M1203-H(M6)_EN.pdf

[15] Datasheet Cerbo GX-GX Touch. Victronenergy [online]. [cited 2023-03-31]. Available from: https://www.victronenergy.com/upload/documents/Datasheet-Cerbo-GX-GX-Touch-EN.pdf

[16] Fireplace arc 70. Flamen [online]. [cit. 2023-03-31]. Available from: https://www.flamen.cz/produkt/krbova-vlozka-arc-70/

ThermoFisher
SCIENTIFIC

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





Brno, Vlastimila Pecha 12

Acquisition Module for Non-Destructive Testing of Aircraft Composite Structures

Ngoc Nam Pham
Department of Microelectronics
Brno University of Technology
Brno, Czech Republic
243756@vutbr.cz
0000-0002-9760-3339

Jan Leuchter
Department of Microelectronics
Brno University of Technology
Brno, Czech Republic
jan.leuchter@seznam.cz
0000-0003-2845-703X

Jiří Háze
Department of Microelectronics
Brno University of Technology
Brno, Czech Republic
haze@vut.cz
0000-0001-5670-428

***Abstract*— This paper presents an acquisition module for non-destructive testing (NDT) of aircraft composite structures. The module was designed with multiple inputs, which allows data to be acquired in parallel in order to speed up the process of obtaining a large amount of data. The collected data are then sent to a computer for data processing and structural damage diagnosis. The design was described by using VHDL (Very high-speed integrated circuit Hardware Describe Language) and implemented in FPGA (Field Programmable Logic Array) Altera Cyclone II development board. The designed module is expected to monitor composite materials' structure health using the NDT technique of Electrical Impedance/Resistance Tomography (EIT/ERT).**

Keywords—Non-Destructive Testing, Aircraft Composite Materials, Electrical Impedance Tomography, VHDL, FPGA.

I. INTRODUCTION (*HEADING 1*)

In recent decades, the use of composite materials has risen swiftly in many technologies and industries, including aircraft technologies. Thanks to their outstanding mechanical properties, composite materials are increasingly used in aircraft primary structures as alternatives to heavier conventional metal alloys. In modern civil airliners (Boeing B787, Airbus A350XWB) and military fighter (Eurofighter Typhoon), composite materials accounts for up to 50 % of the total aircraft weight. In helicopter (NHIndustries NH90), this value reaches up to 90 %. The main advantages of composite materials, compared with metal alloys, are low weight, higher strength, stiffness, corrosion resistance, better coefficient of Thermal Expansion and ability to produce complex parts, special appearance, and contours [1]–[7].

The key property of composites is their structure which combines two or more materials together to achieve better properties (desirable properties of each element). The complex structure with different phases of materials (fibers and matrix) of composites; however, results in various types of damage in case of accident. Therefore, in order to get safer and more extensive use of composite structures in aerospace applications, it is important to monitor the health of materials which is always conducted in parallel with Non-Destructive Testing (NDT). Continuous monitoring and information acquisition in real-time will significantly increase the operational safety [1], [8], [9].

The state-of-the-art NDT techniques for composite materials were described in [1], [8]. The main ones are visual inspection, optical methods, ultrasonic inspection, laser ultrasonic inspection, acoustic emission, vibration analysis, radiography, thermography, Lamb wave and electrical impedance/resistance tomography (EIT/ERT). Most of these techniques are based on the difference when the test signal propagates through the defect-free structure and through the defect(s) of the structure to identify and locate defect(s). The test signal can be acoustic, ultrasonic, radio wave, or electric current. In some cases, a bulky signal source and transducer can be required for testing. Therefore, this paper focuses mainly on NDT technique of EIT/ERT, where the signal source and transducer are smallest. The example of defect locating using EIT/ERT technique is shown in Fig. 1.

The basic principle of EIR/ERT technique (see Fig. 1) is measuring the voltages on the boundary of investigated object when a small current flows through the object. According to the measurement data, the resistance map is reconstructed using the Finite Element Method (FEM). The accuracy and resolution of this technique increase with the increasing number of both electrode and measurement (16-electrode test requires 16x13=208 independent measurements, 32-electrode test requires 32x29=928 independent measurements). Therefore, it is necessary to speed up the data acquisition, for example by performing measurements in parallel. This paper focuses on design a FPGA-based acquisition module which enables parallel acquired measuring voltage in EIT/ERT testing.

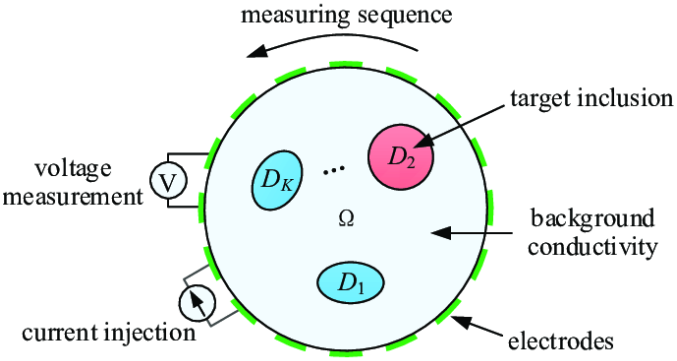


Fig. 1. Principle of EIT/ERT technique [10]

II. Design of the Module

A. Basic concept and used components

Acquisition module is designed for EIT/ERT systems with a parameterizable number of electrodes. The block diagram of EIT/ERT system is shown in Fig. 2. An Analog-to-Digital Converter (ADC) with the differential input is connected between two adjacent electrodes to measure the drop voltage between them. The number of required ADCs is now equal to the number of electrodes. The designed module is operated to parallelly acquire data from these ADCs and send these data to a computer for next processing and computing.

In the design, two types of ADC were selected. The first one is 24-bit, differential, delta-sigma ADC HX711, which has an on-chip Programmable Gain Amplifier (PGA) with gain of 64 or 128. This ADC provides a high accurate measurement of small voltage, but low output data rate (10 or 80 SPS). The second one is 16-bit, differential, delta-sigma ADC ADS1115, with an integrated PGA with maximal gain of 16. Although the gain and resolution of ADS115 are much smaller than HX711, its output data rate is much higher, up to 860 SPS, which makes it more suitable for test signal with higher frequency.

The interface to control and communicate with HX711 is a serial interface via 3 ports - SCK for clock signal, DOUT for data signal, and GND for reference ground. The interface to control and communicate with ADS115 is I2C (Inter-Integrated Circuit) one via 4 ports: SCL for clock signal, SDA for data signal, VDD for power supply, GND for reference ground. In our design, to synchronize measurements, the clock signal is common for all channels (all the ADCs), while the data signal for each channel (each ADC) is independent.

To send data to computer, a simple UART (Universal Asynchronous Receiver/Transmitter) interface was designed in FPGA. A USB-to-UART bridge CP2102 was selected to transmit data via computer's USB ports. For data transmission with higher baud rate, data can be transferred via I2C to ESP-32 microcontroller, which will send these data to computer via Ethernet or wirelessly via Wi-Fi, or ESP-NOW. The details of wireless communication between ESP-32 microcontroller and computer were presented in [11].

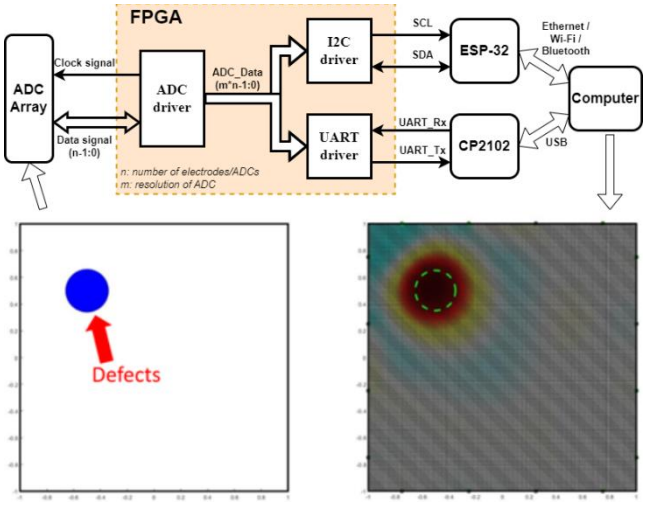


Fig. 2. Block diagram of EIT/ERT system

B. VHDL description

The VHDL description of acquisition module can be separated into 3 drivers, as shown in Fig. 2. The first driver is **ADC driver**, which controls and obtains measured data from the ADC array connected to the investigated object. The second one is **I2C driver**, which sends the acquired data to microcontroller ESP-32 via I2C bus. The last one is **UART driver**, which sends the acquired data to computer via UART bus. All the drivers were described in VHDL as Finite State Machine (FSM). Their state diagrams are shown below.

Fig. 3 shows state diagram of driver controlling HX711 ADC. This FSM is very simple with only 3 states. When no action is required, FSM is in state **Idle**. If ADC data is ready for being read, the SCK bus is pulled down by ADC, FSM changes to state **Read**. After reading 24 bit ADC data, from MSB to LSB, the FSM changes to state **Setup**, the driver sends SCK pulse to configure ADC for next conversion and return to state **Idle**. The number of SCK, which pulses to configure ADC, should not be less than 1 or more than 3.

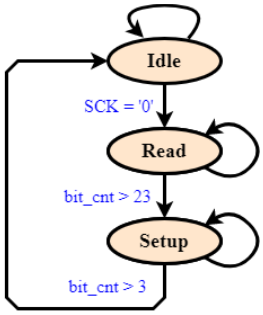


Fig. 3. State diagram of HX711 driver

Fig. 4 shows the state diagram of I2C driver, which is used for both communication between FPGA and microcontroller and controlling ADC ADS1115. When no action is required, the FSM is in state **Idle**, both SCL and SDA lines are at high level. As soon as the start signal is pulled up, FSM changes to the state **Start**, SDA declines while SCL is still in high level. After holding time is defined, FSM changes to state **Addr**, the driver sends 7-bit address of I2C slave with the bit R/$\overline{W}$ to require Read or Write to Slave, then changes to the state **ACK_Addr**. If the address is accurate, the slave sends an acknowledge (ACK) signal by pulling down the SDA line. If it is required to write to slave, the FSM changes to state **Write** and sends data in each set of 8 bit. After receiving the ACK signal from slave, the driver changes to state **Stop**, sets the SCL line to high level, the SDA line in low level; then returns to state **Idle**.

If it is required to read data from slave, the driver will change to the state **Read**. However, in some cases, I2C slave needs more time to manage data and cannot respond to I2C master timely. In this case, the I2C slave can pull down the SCL line, perform an "clock stretching" and force the master into waiting state. In this case, the driver will remain on state **Stretching** until the SCL line is released to high level. After that driver changes to state **Read**, reads data in a set of 8 bits, then changes to state **ACKRead**, and sends an ACK signal to slave by pulling the SDA line low. After sending ACK signal, the driver changes to state ***Stop*** and returns to state **Idle**.

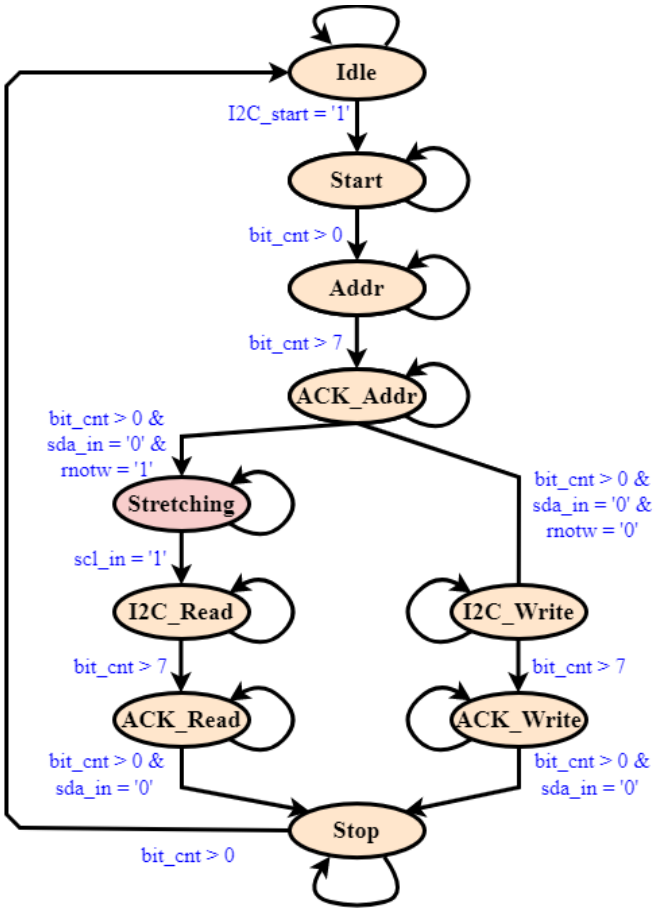


Fig. 4. State diagram of I2C driver

The driver for ADS1115 is designed based on the I2C driver. The first difference is that ADS1115 does not support the clock stretching so the state **Stretching** is skipped. The second difference is that when driver needs to write to ADS1115 (to configure the ADC), it is necessary to send the address of configured register after sending address of ADC with R/$\overline{W}$ bit. In the case of accurate ADC address, ADC will send an ACK signal by pulling up the SDA line.

The state diagram of UART driver is shown in Fig. 5. The driver will remain in state **Idle** if no action is required. Once it is required to send data, the driver changes to state **TX_Start** and pulls down the TX line. The driver then sends the data bit from LSB to MSB in turn, in set of 8 bits. After sending 8 bits, the driver changes to state **Stop**, pulls the TX line higher and then continues sending all the data before returning to state **Idle**. If RX line is pulled down, the driver receives a request to read the data. The driver then reads data from LSB to MSB in set of 8 bits, then returns to state **Idle** and waits for the next request of receiving data.

In this design, the data is sent and received in a set of 8 bits, parity bit after each set of 8 data bits is not used, and Stop bit(s) size is 1 bit. This configuration is called 8N1 UART configuration, which is the most common configuration of UART data transmission. To transmit data successfully, both transmitting and receiving sides must have the same configuration and baud rate.

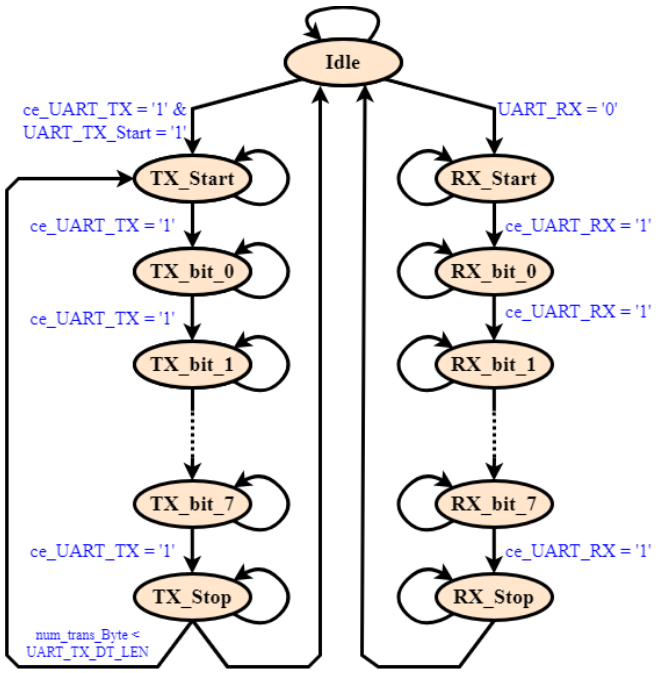


Fig. 5. State diagram of UART driver

III. Implementation and Verification

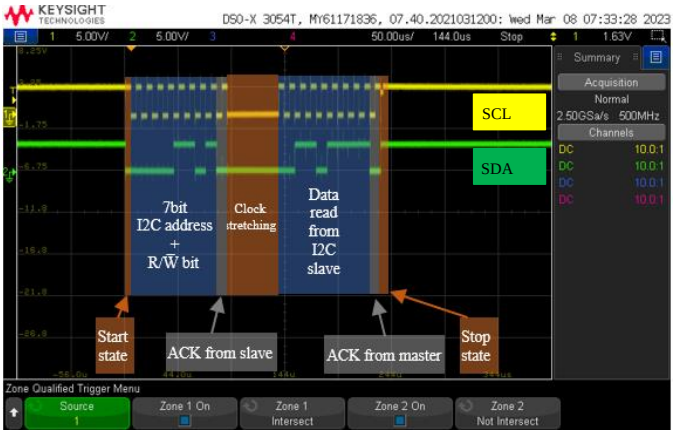


Fig. 6. Waveform of I2C bus – Master read from Slave

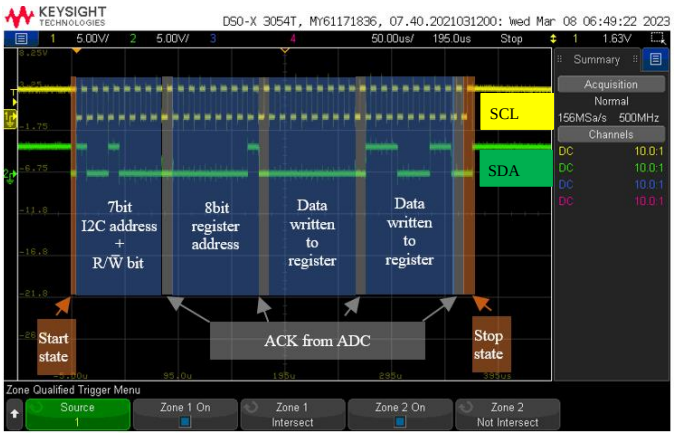


Fig. 7. Waveform of I2C bus – Master write to Slave

VHDL designs were implemented in FPGA development board Altera Cyclone II. The verification of drivers for HX711 and UART communication was presented in [4]. The verification of I2C driver for communication with microcontroller ESP-32 and ADC ADS1115 is shown above.

Fig. 6 shows the waveform of I2C bus in the case FPGA read data from microcontroller, Fig. 7 presents the waveform of I2C bus when FPGA configures the ADC. In Fig. 6, it can be seen that the I2C stretching was performed by microcontroller, SCL was pulled down after ACK. The VHDL design of I2C bus now supports the clock stretching function while communicating with microcontroller.

Acquired data is then sent to computer for further processing. In order to reconstruct resistance/impedance map of investigated object, the MATLAB-based software EIDORS is used. Fig. 8 shows the different results of resistance/impedance map reconstruction with different position of defect. The input data for processing were obtained from simulation in Ansys Maxwell.

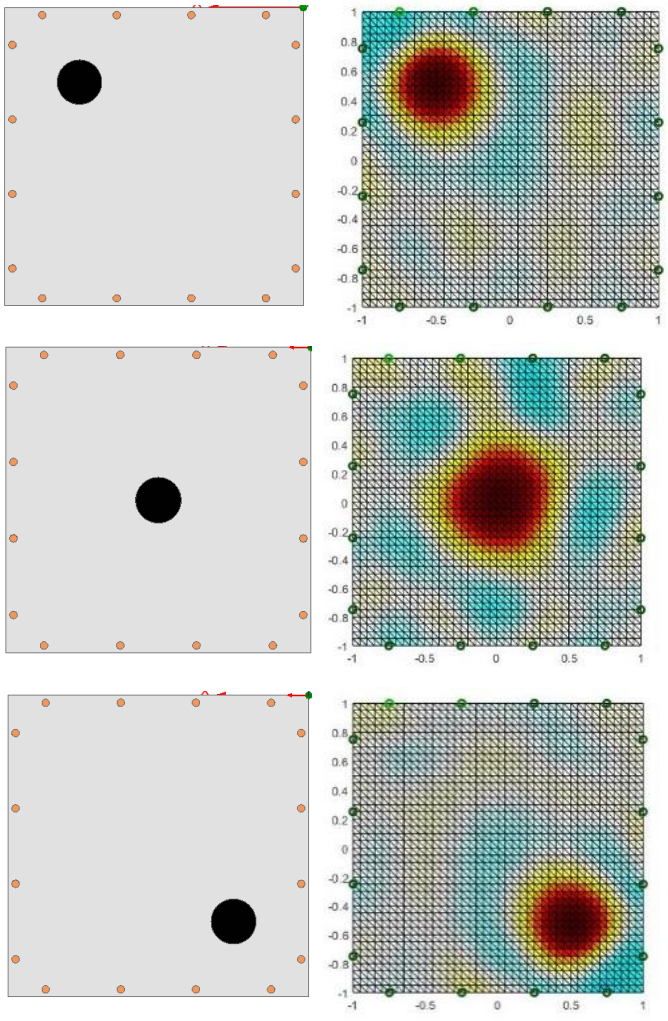


Fig. 8. Resistance map reconstruction using EIDORS

IV. Conclusion

This paper presents a design of FPGA-based acquisition module for NDT technique EIT/ERT. The benefit of using FPGA instead of microcontroller to acquire data is that FPGA enables to operate tasks in parallel while the microcontroller operates in sequence. In this manner, the data acquisition process will be sped up, which allows continuous monitoring and information acquisition in real-time. This requirement is very important for increasing operating safety, especially in aerospace application.

This paper also briefly shows the results of resistance/impedance map reconstruction using MATLAB-based software EIDORS with simulation data from Ansys Maxwell. The results show that EIT/ERT technique is one of the promising techniques for NDT testing when the location of defect can be clearly detected

Acknowledgment

Authors gratefully thank the AERO Vodochody AEROSPACE a.s., Brno University of technology and Le Quy Don Technical University for supporting this work.

References

[1] K. Diamanti and C. Soutis, “Structural health monitoring techniques for aircraft composite structures,” *Progress in Aerospace Sciences*, vol. 46, no. 8, pp. 342–352, Nov. 2010, doi: 10.1016/J.PAEROSCI.2010.05.001.

[2] B. Alemour, O. Badran, and M. R. Hassan, “A review of using conductive composite materials in solving lightening strike and ice accumulation problems in aviation,” *Journal of Aerospace Technology and Management*, vol. 11, 2019, doi: 10.5028/JATM.V11.1022.

[3] C. Soutis, “Fibre reinforced composites in aircraft construction,” *Progress in Aerospace Sciences*, vol. 41, no. 2, pp. 143–151, 2005, doi: 10.1016/J.PAEROSCI.2005.02.004.

[4] N. Nam Pham, J. Leuchter, and Q. Huy Dong, “FPGA-based measurement instrument for the ERT applications of aerospace composite materials,” *Proceedings of the 2022 20th International Conference on Mechatronics - Mechatronika, ME 2022*, 2022, doi: 10.1109/ME54704.2022.9983440.

[5] F. Smith, “The Use of composites in aerospace: Past, present and future challenges,” 2013.

[6] F. C. Campbell, “Structural Composite Materials,” 2010, Accessed: Mar. 06, 2023. [Online]. Available: www.asminternational.org

[7] L. Q. Đại, “COMPOSITES MANUFACTURING Materials, Product, and Process Engineering”, Accessed: Mar. 06, 2023. [Online]. Available: https://www.academia.edu/8105712/COMPOSITES_MANUFACTURING_Materials_Product_and_Process_Engineering

[8] Y. B.-C.-T. e-J. of N. T. & Ultrasonics and undefined 2000, “Emerging NDE technologies and challenges at the beginning of the 3rd millennium—Part I, Part II,” *ndt.net*, Accessed: Mar. 06, 2023. [Online]. Available: https://www.ndt.net/article/v05n01/barcohen/barcohen.htm

[9] “Boller: Encyclopedia of structural health monitoring - Google Scholar.” https://scholar.google.com/scholar_lookup?title=Encyclopedia%20of%20structural%20health%20monitoring&publication_year=2009&author=C.%20Boller&author=F.-K.%20Chang&author=Y.%20Fujino (accessed Mar. 06, 2023).

[10] S. Ren, Y. Wang, G. Liang, and F. Dong, “A Robust Inclusion Boundary Reconstructor for Electrical Impedance Tomography with Geometric Constraints,” IEEE Trans Instrum Meas, vol. 68, no. 3, pp. 762–773, Mar. 2019, doi: 10.1109/TIM.2018.2853358.

[11] N. N. Pham, J. Leuchter, K. L. Pham, and Q. H. Dong, “Battery Management System for Unmanned Electric Vehicles with CAN BUS and Internet of Things,” *Vehicles*, vol. 4, no. 3, pp. 639–662, Sep. 2022, doi: 10.3390/VEHICLES4030037.

Design of a Hardware Accelerator of Demanding Computations with Multiple FPGAs

Petr Zach
Department of Microelectronics
Brno University of Technology
Brno, Czech Republic
xzachp02@vutbr.cz

***Abstract*—This paper compares certain circuits suitable for hardware acceleration of computationally intensive calculations, namely DSP, GPU, and FPGA. The paper presents a list of their advantageous and disadvantageous properties. Based on this information, a proposal for a custom hardware accelerator platform based on several FPGA circuits is presented.**

***Index Terms*—Hardware acceleration, DSP, GPU, FPGA, Xilinx, AXI, design of hardware accelerator**

I. Introduction

In today's world, computationally intensive algorithms are increasingly being used in various fields such as image processing, machine learning, cryptography, etc. Since standard processors often cannot provide sufficient performance for these tasks, specialized hardware platforms are emerging that focus on accelerating computationally intensive algorithms.

These platforms, or hardware accelerators, are specially designed chips or devices that allow for processing a specific task faster than ordinary processors. They are particularly suitable for tasks that are time-consuming, require a large number of iterations of simple computational operations, or allow for parallel processing.

There are several ways to proceed with the hardware design of such a platform. Among the most common components used for accelerating computations are FPGA (Field Programmable Gate Array), GPU (Graphics Processing Unit), DSP (Digital Signal Processor), or ASIC (Application Specific Integrated Circuit). Such a computing system can be composed of one type of computational unit or multiple types. Based on this fact, systems can be classified as homogenous or heterogeneous (a typical example of a heterogeneous system is the combination of CPU (Central Processing Unit) + GPU in PC (Personal Computer) systems). Furthermore, each of the above-mentioned circuits has slightly different properties, making it challenging for hardware designers to make the right decision when choosing the type of acceleration unit for a specific application.

This research deals with comparing critical properties of selected circuits in the field of hardware acceleration, followed by a low-cost hardware accelerator design from a hardware perspective.

II. Comparison of selected circuits

A. Digital Signal Processor

A digital signal processor, also known as DSP, is a type of microprocessor with an architecture designed specifically for performing digital signal processing tasks. The architecture typically consists of three parts: memory units, central processing units, and I/O interfaces.

Advantages of using DSPs:

- The development costs of a portable system based on DSP are usually low. DSP chips are generally cheaper than most other hardware accelerators for portable devices [1].
- Many computer vision and image processing applications involve sequential algorithms, and the architecture of DSP chips is designed for implementing sequential tasks. Multi-core DSPs have added the ability to implement coarse-grained parallelism for low to medium complexity algorithms [1].
- The development time for simple computer vision and image processing algorithms on single-core DSPs is generally relatively short [1].
- DSPs are suitable for handling peripherals, standard ports (such as USB and SATA), and communication protocols (such as TCP/IP) in portable devices [1].

Disadvantages of using DSPs:

- Multi-core DSPs are designed for high-performance computing (HPC) applications with low to medium complexity. They are not suitable for high-bandwidth applications or high-speed applications [1].
- DSPs are better suited for sequential processing. Even though multi-core DSPs are capable of performing coarse-grain parallelism, they are not a suitable choice for increasing processing speed in massively parallel algorithms [1].
- It is usually not efficient to use DSP together with CPU in a PC to increase the processing speed of an algorithm. DSP and CPU have a similar nature of sequential processing, while programming for CPU is simpler and more efficient than for DSP [1].

B. Graphics Processing Unit

Graphics Processing Unit, or GPU, is a type of microprocessor with an architecture designed specifically for image processing tasks. The architecture of GPUs varies depending on the manufacturer and model of the specific GPU.

A graphics processing unit typically consists of multiple texture processing clusters (TPCs). These TPC blocks further include streaming multiprocessors (SMs), texture units, and texture caches (L1). The streaming multiprocessor is composed of scalar processors (SPs), double precision units, two special function units, shared memory, L1 cache, and registers. These scalar processors are the smallest computational units of GPU chips, referred to as CUDA cores in NVIDIA's lineup.

Advantages of using GPUs:

- They are mass-produced and have the best ratio of computing power to cost. Thanks to the entertainment industry, they are relatively inexpensive compared to FPGA circuits [1].
- GPU technology is rapidly evolving and despite technological advancements, the price remains favorable [1].
- High-level programming languages are available for GPU programming. Therefore, development and debugging of code for GPUs is faster and easier than for FPGAs [1].
- Easy use of PCIe interface between GPU and CPU [1].
- They are specifically designed for image and video processing [1].

Disadvantages of using GPUs:

- GPUs are designed to solve problems with massive data parallelism. Therefore, the performance of a GPU is significantly reduced if the data processing is time-consuming or if there is a delay in data flow [1].
- GPU have significantly higher energy consumption compared to FPGA circuits when considering the same performance class [1].
- The main problem when using GPUs in PC-based systems is the data transfer time between the PC and GPU. Therefore, unoptimized GPU code may not help increase processing speed [1].
- The development of low-level functions in GPUs often requires the use of the assembler language. The GPU hardware is pre-structured and has lower flexibility compared to FPGA [1].

C. Field Programmable Gate Array

A field-programmable gate array, also known as an FPGA, is a type of integrated circuit that can be used to implement any digital circuit. Unlike the previously mentioned CPU, DSP, and GPU circuits, FPGA circuits do not have a fixed chip architecture or central processing unit. The programmer designs the FPGA architecture before programming it for a specific application. Despite not having a fixed architecture, FPGAs contain basic blocks that are available to the programmer. FPGA circuits are proprietary, so the structure and nomenclature of FPGAs typically vary by manufacturer.

An FPGA typically contains configurable logic blocks (CLBs), I/O cells, and a programmable interconnect (PI). In order to increase the flexibility of an FPGA, blocks with specific functionality such as block RAM (BRAM), DSP blocks, or microprocessors are integrated into its structure. If a microprocessor is integrated onto the same chip with an FPGA, it is then considered a system on chip (SoC) [1] [2].

Advantages of using FPGAs:

- The processing speed of FPGA is higher than other hardware accelerators. The latest FPGA can process billions of operations per second in parallel using their DSP blocks. No other type of hardware accelerator can achieve such speeds [1].
- Using FPGAs, it is possible to achieve high data throughput, which is why they are advantageous for data acquisition cards [1].
- FPGAs are relatively energy-efficient. Therefore, they are a suitable choice for use in portable devices [1].
- FPGAs enable the implementation of flexible and efficient algorithms through reconfiguration of the FPGA hardware and optimization for a specific algorithm type [1].

Disadvantages of using FPGAs:

- Developing efficient code for FPGA is challenging. Although tools have been developed to simplify and accelerate this process, the programmer still needs sufficient knowledge and technical skills to develop robust code for FPGA [1].
- Long development time for FPGA applications [1].

D. DSP, GPU and FPGA comparison

According to research in literature [1] and examples of DSP circuit usage, they have been shown to not be particularly suitable for high-performance applications. Despite recent advances in multicore DSPs, it is still not possible to utilize them for implementing complex image processing and computer vision algorithms, especially in cases where high data throughput is required. Additionally, DSP circuits are not suitable for PC-based systems, as they do not provide any significant advantages over GPUs other than external interfaces. On the other hand, DSPs are a suitable solution for embedded, mobile, or portable systems where high computational power is not required, but power consumption is critical.

In recent years, there has been a growing trend in the use of graphics processing units (GPUs) for scientific computing. GPUs are relatively inexpensive, the latest generations achieve high processing speeds, and development of complex applications for these circuits is relatively easy. The development time for image processing and computer vision applications is shorter when using GPUs than with FPGAs. GPUs can provide good performance in applications that do not require data acquisition. However, even the latest GPUs sometimes

lack sufficient computational power and/or memory space for complex algorithms. The number of CUDA cores that can be included in a single GPU chip is limited by manufacturing capabilities. One solution is to use multiple GPUs in a single system, but this case also has certain limitations. One of them is data communication and computation management between multiple GPUs. Therefore, the computational speed of multiple GPUs is typically not proportional to the number of GPUs used compared to a single GPU. In some cases, the resulting computational speed of multiple GPUs is even lower than that of a single GPU.

FPGAs are the most flexible hardware accelerator for implementing customized algorithms for digital signal processing, image and computer vision. However, good knowledge of digital logic design, FPGA circuit hardware structure, descriptive languages for digital circuit description (HDL; Hardware Description Language), and programming tools are essential for efficient implementation of complex algorithms in FPGA circuits.

GPU and FPGA are commonly used for implementing image processing and computer vision algorithms compared to DSP circuits. Therefore, it is interesting to compare implementations on GPU and FPGA, however, an accurate performance comparison between them is not practical because each application has different requirements, and the performance of an application depends on the level of programming knowledge and the hardware used. Moreover, in most published works, the comparison between GPU and FPGA is not performed for comparable technological levels. Typically, newer generations of GPU are compared with FPGA. This may be due to the lower cost of acquiring a GPU chip compared to FPGA or their relative programming simplicity [1].

III. System-level Hardware Acceleration Platform Design

For the purposes of this work, a concept of a heterogeneous system with computing components consisting of CPU + multiple FPGAs was chosen. The selection was made with respect to the low cost and flexibility of the resulting platform for hardware acceleration. This combination will allow easy control of communication interfaces and individual peripherals by the central processing unit while also providing computing power of multiple FPGA circuits.

The core of the system was chosen to be an FPGA circuit of the Xilinx XC7Z020-1CLG484I type SoC. This FPGA circuit is integrated on the chip with a dual-core ARM Cortex-A9 processor. The computational FPGA circuits were chosen from the same manufacturer and from the same Artix-7 chip family. Specifically, it is the XC7A100T-2FGG484I circuit. When selecting the computational FPGA, available resources were taken into account, such as the number of DSP blocks, the number of LUT tables, logical cells, as well as high-speed transceivers.

A. Data interfaces

As outlined before, the purpose of a hardware accelerator is to accelerate computational applications by reducing their computation time. This is why it is critical that the communication interface is designed with high data throughput in mind. If this were not the case, delays caused by data transfer could completely erase the time saved by more efficient algorithm computation.

High data throughput is the first criterion when choosing a communication interface. The second criterion is its scalability and usability for connecting multiple computational FPGAs.

The chosen interface for communication between the proposed platform and the surrounding world is an Ethernet interface. Due to the capabilities of low-cost SoC FPGAs, this is the fastest possible communication with support for 1000 Mb/s. Another advantage of this solution is its wide usage across various industry devices. In addition, the platform also offers connectivity via USB 2.0 with a speed of up to 480 Mb/s

With regard to the criterion of high data throughput, it is advantageous to use the available solution from the manufacturer Xilinx for communication between individual FPGA circuits. Xilinx provides support to designers using their FPGA circuits through the design of Intellectual Properties (IP) blocks and their integration into the Vivado Design Suite design environment. Specifically, this includes the IP block called AXI Chip2Chip [4], an adaptable block based on the AXI-4 protocol [3] that provides a communication interface between two chips. It offers the possibility of implementing several types of physical layers of this interface, two of which can be implemented on the selected FPGA, namely:

- LogiCORE IP SelectIO [4].
- LogiCORE IP Aurora 8B/10B [5].

These physical layers differ mainly in the maximum achieved data transfer rate and in the number of occupied I/O pins of the FPGA circuits.

The advantage of the SelectIO physical layer is the possibility of implementing the AXI Chip2Chip IP core even into circuits that do not contain high-speed transceivers (GTP links). This physical layer is therefore implemented using common I/O pins and it is a half-duplex communication. The resulting disadvantage is their relatively high I/O pins consumption. The number of consumed I/O pins is directly proportional to the resulting data throughput. Given that the selected SoC FPGA device does not provide high-speed transceivers, this is the only way to implement the AXI Chip2Chip IP core. For this implementation, a configuration utilizing 46 I/O pins and an expected data throughput of up to 3600 Mb/s was chosen.

The Aurora 8B/10B physical layer is implemented using high-speed transceivers. This physical layer provides full-duplex communication, which, together with the use of GTP links, allows for significant savings in the number of I/O pins used. Given that the computational FPGA circuits contain

8 GTP links, it is advantageous to use this physical layer for communication between them, thus simplifying the PCB (Printed Circuit Board) design.

B. Data Storage

In order to expand the memory capabilities of this board, it contains two DDR3 RAM memory modules providing a memory space of 1 GB.

These RAM memories are connected to the main control SoC FPGA circuit. The complete block diagram of the hardware acceleration platform is shown in Fig. 1.

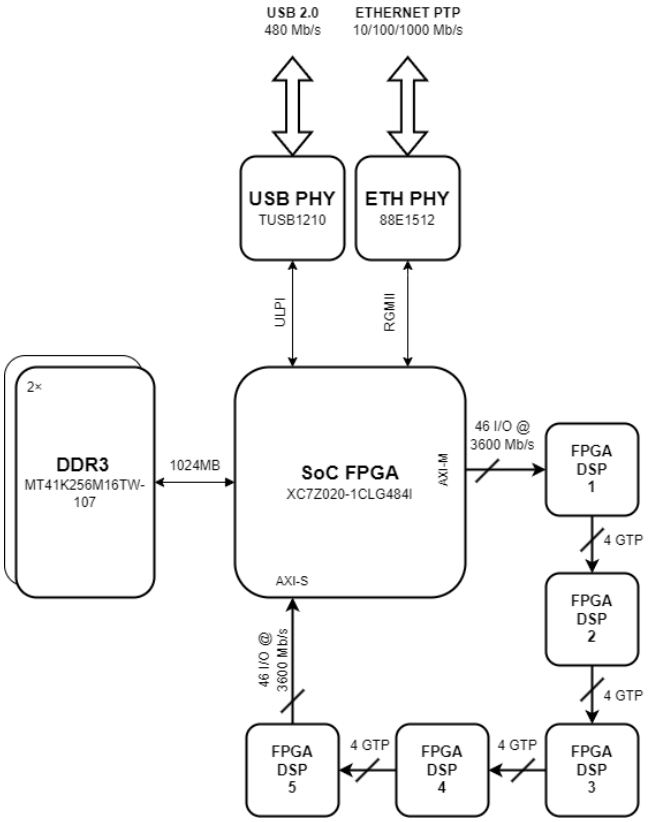


Fig. 1. Hardware acceleration platform block diagram

IV. Conclusion

The result of this work is the selection of circuits suitable for hardware acceleration, followed by their comparison and outlining of advantageous and disadvantageous properties, and thus identifying possible areas for their use. These circuits include DSP, GPU, and FPGA.

Based on the given information, it was possible to design a low-cost hardware accelerator with emphasis on maximum flexibility, compatibility, and performance. For these reasons, a heterogeneous system using both CPU and multiple FPGAs was chosen. The CPU is capable of reliably and efficiently handling the communication interface and implementing communication protocols, while FPGA chips provide high flexibility and performance.

The most critical areas of this hardware acceleration platform were introduced and possible solutions for low-cost chip solutions were explained in detail. The most critical area is the data transfer method. Ethernet interface was chosen for input data to the hardware acceleration platform as it offers the highest data throughput among available options. For data transfer between circuits on the platform, Xilinx's IP core was chosen, again with a focus on high data throughput. Due to the unavailability of high-speed lines in the SoC FPGA, high data throughput is ensured by a large number of I/O ports. In this case, the PCB design will also be critical.

Acknowledgment

I would like to thank Ing. Martin Šťáva, Ph.D for his invaluable advice and guidance throughout the project.

References

[1] HAJIRASSOULIHA, Amir. *Signal Processing: Image Communication: Suitability of recent hardware accelerators (DSPs, FPGAs, and GPUs) for computer vision and image processing algorithms.* Auckland, New Zealand, 2018. The University of Auckland.

[2] ZACH, Petr. *Návrh vývojového kitu s obvodem FPGA [online].* Brno, 2021 [cit. 2023-01-04]. Dostupné z: https://www.vutbr.cz/studenti/zav-prace/detail/134687. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav mikroelektroniky. Vedoucí práce Lukáš Fujcik.

[3] AMBA AXI and ACE Protocol Specification Version E. *Arm Developer [online]. c1995-2022 [cit. 2023-01-04].* Dostupné z: https://developer.arm.com/ documentation/ihi0022/e/

[4] AXI Chip2Chip v5.0: LogiCORE IP Product Guide. *Xilinx [online]. c2022 [cit. 2023-01-04].* Dostupné z: https://www.xilinx.com/content/dam/xilinx/support/documents/ip_documentation/axi_chip2chip/v5_0/pg067-axi-chip2chip.pdf

[5] Aurora 8B/10B v11.1 LogiCORE IP Product Guide *Xilinx [online]. c2022 [cit. 2023-01-04].* Dostupné z: https://docs.xilinx.com/r/en-US/pg046-aurora-8b10b/Aurora-8B/10B-v11.1-LogiCORE-IP-Product-Guide

Testbed for Simulation of MCU Application using RISC-V soft-core

Petr Ohnút
Department of Control and Instrumentation
Brno University of Technology
Brno, Czechia
Email: Petr.Ohnut@vut.cz

Jakub Arm
Department of Control and Instrumentation
Brno University of Technology
Brno, Czechia
Email: Jakub.Arm@vut.cz

***Abstract*—The MCU software testing based on RTOS (Real-Time Operating System) is an extensive area in research and practical domain. The work presents an approach to testing firmware using RISC-V soft-core in the FPGA environment. The testing process embodies a designed test-bench that can operate automatically using a script. The testing procedure rests in the execution of a tested program with a pre-defined input set of signals. The procedure output is stored in a structured file. Such a testing scenario represents the program behavior in a specific situation; using the automated way, more scenario outputs could be obtained. The preliminary results suggest the output files of a program could be analyzed to determine if the tested program comprises an error.**

***Index Terms*—Testing, MCU, RTOS, soft-core, FPGA, RISC-V**

I. Introduction

With time, the demands on the safety of embedded firmware are rapidly increasing in applications where embedded systems control demanding processes whose failure could be catastrophic. [1]

It is financially advantageous to test the designed processor in FPGA or RTL simulations to try to detect unexpected or erroneous behavior of the designed MCU.

This paper deals with the possibility of automatic testing of MCU, mostly FreeRTOS [2], applications in a soft-core RISC-V processor. The processor under test is an open-source RISC-V [3] implementation NeorV32 [4] using Vivado RTL simulator. Other options within simulation tools could include Matlab Simulink, which is used to simulate MCUs by the authors in [6], or JAVA Simulator, which was used by the authors in [7], another interesting simulation options are discussed in [8].

The architecture of the test framework itself is discussed, which has been implemented in Python and provides automatic generation of the Vivado project, execution of individual test scenarios, and subsequent processing of the results.

Furthermore, the test scenarios themselves are described. Most of the tested applications use FreeRTOS, which has a critical bug included in it. The goal is to determine the behavior of the processor during these critical errors. The other test scenarios test peripheral behavior. Several corner case scenarios (e.g., maximum or minimum communication speed) during which a potential error or unexpected behavior could occur were selected for testing.

In the next section, the testbench architecture itself is described together with an example of simulation output data.

Finally, the vision of the project is discussed along with a summary of the current results.

II. Testing framework

The main element of the whole testing process is the testing framework itself, which is implemented in Python (specifically version 3.10.9). The framework is mainly used to run TCL scripts using the Vivado development environment. These scripts then automatically create and prepare a Vivado project or load an existing project for subsequent application testing and run all or only a few selected test scenario. Before each simulation is run, the required files are also checked; these files include, for example, the application image itself or the input data file for the scenario. The simulation and measurement data is processed by testbench, which is described in more detail in section IV. In section IV we will also describe what data are measured and how the data are processed for better readability.

The test framework consists of three classes namely Verifier, Generator and Simulator classes. These classes are described in more detail below. Supported framework options are listed in Fig. I

TABLE I
Supported options

Option	*Description*
`--help (-h)`	help
`--verbosity (-v)`	setting the detailed listing mode
`--generate_project (-gp)`	generate Vivado project
`--project_path (-pp)`	path to an existing project
`--vivado_path (-vp)`	path to the Vivado program
`--simulate (-s)`	simulate single selected scenarios
`--simulate_all (-sa)`	simulate all available scenarios

A. Class: Verifier

The Verifer class serves as the parent class for the Generator and Simulator classes. Its attributes are the logger settings, the Vivado program path, and the Vivado project path. The class

also includes a method that allows you to run commands on the command line.

B. Class: Generator

The Generator class is mainly used to generate a Vivado project. This class is only used if the user chooses the option to generate a new project.

C. Class: Simulator

The Simulator class runs simulations of individual test scenarios, checks the required files, and then processes the simulation output data.

III. Test scenarios

The test scenarios, which are represented by simple C programs, have been designed to either have a bug included (primarily FreeRTOS scenarios) or to be edge cases of a given periphery (e.g., maximum or minimum communication speed). The goal is to observe the processor behavior if the program contains an error or to observe if the processor behaves correctly in the aforementioned edge cases. Several test scenarios have been designed as part of the testing. These scenarios can be divided into two groups namely: testing input/output peripherals, interrupts, and testing FreeRTOS applications.

A. Configuration

Each test scenario can be configured in a JSON configuration file. The user has the option to configure the simulation length for each scenario. Other options for configuring the test scenarios are under discussion for now.

B. Input data

Both input and output files are in format .cvs with the ";" character separating the individual values.

In Fig.III you can see an example of the simulation input file. This example has been modified so it could be shown in the table. The input file contains information about the data for tested peripherals and timestamps at which are data sent to the input.

TABLE II
Example of simulation input data

Time [ns]	rstn_i	gpio_i	uart0_rxd_i
100	1	0	0
80000	1	1	A

C. Interrupts

Two scenarios have been created that use interrupts. The first scenario uses an external interrupt input to the processor, which is controlled by the testbench. The switching frequency can be adjusted within the testbench. The next interrupt test scenario uses a counter that triggers an interrupt when counting up to a selected value. The desired counter value can be set within the C application.

D. Peripherals

UART, GPIO, PWM, and SPI peripherals have been selected for testing. The test scenarios for UART, GPIO, and PWM have already been developed in the thesis [5]. For SPI testing, the scenario that was designed uses either maximum or minimum communication speed. The SPI test scenario is still in development.

E. FreeRTOS

Most of the test scenarios were created for FreeRTOS applications. Each scenario has an intentionally injected bug in it. The injected bugs include deadlock, livelock, race condition, starvation, unbounded priority inversion, and bounded priority inversion.

IV. Simulation

As already mentioned, all scenarios are tested using the simulator included in Vivado. During the simulation of each scenario, data is collected about the states of the outputs of the peripherals. As part of the testing, program counter values, and executed instructions are also recorded. These data are preprocessed after the simulation is completed and are ready for further processing.

The whole simulation process can be seen in Fig.1. The dashed line indicates a process that is optional.

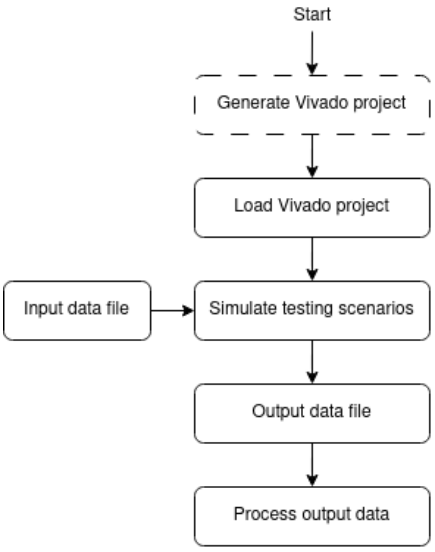


Fig. 1. Simulation process

A. Testbench

During the design, it is necessary to keep in mind that the testbench should be modular and extensible, so a standard structure is given with all the standard verification components, except the scoreboard, that make this possible. This standard testbench structure can be seen in Fig. 2.

A major part of the testbench has already been implemented in [5].

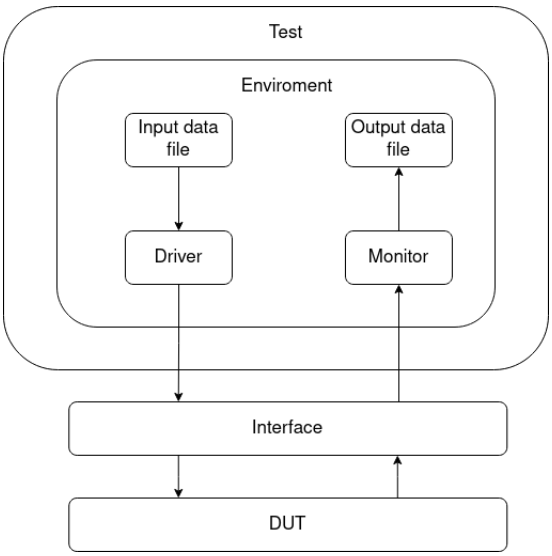


Fig. 2. Testbench structure.

B. Output data

Output data of the simulation are captured by the testbench itself. The testbench captures every change on processor outputs and also the changes of the program counter. The captured data are saved in the output data file which has the same format as the input data file.

An example of captured data can be seen in Fig.III

TABLE III
EXAMPLE OF SIMULATION OUTPUT DATA

Time [ns]	curr_pc	instr	gpio
900	00000034	b0001073	0x02

C. Output data processing

Once the simulation is complete, the output data are processed by a Python script. The processing is done in such a way that for each captured instruction, its assembler equivalent is found, which is then stored in a column next to the captured instruction. This achieves better readability of the captured data. Other methods of processing the output data are discussed.

An example of processed output data can be seen in Fig.IV

TABLE IV
EXAMPLE OF SIMULATION PROCESSED OUTPUT DATA

Time [ns]	curr_pc	instr	asm	gpio
900	00000034	b0001073	csrw mcycle,zero	0x02

In Fig.V we can see the processed data of the deadlock test scenario. We can observe that at approximately 2260 ns, we enter the deadlock state, and the processor under test cycles between two instructions indefinitely.

TABLE V
EXAMPLE OF PROCESSED DATA OF DEADLOCK TEST SCENARIO

Time[ns]	curr_pc	instr	asm
1990	00000080	00000c93	li s9,0
2010	00000084	00000d13	li s10,0
2030	00000088	00000d93	li s11,0
2050	0000008c	00000e13	li t3,0
2070	00000090	00000e93	li t4,0
2090	00000094	00000f13	li t5,0
2110	00000098	00000f93	li t6,0
2130	0000009c	00000417	auipc s0,0x0
2150	000000a0	d6440413	addi s0,s0,-668
2170	000000a4	00000497	auipc s1,0x0
2190	000000a8	f5c48493	addi s1,s1,-164
2210	000000ac	00042023	sw zero,0(s0)
2260	000000b0	00004597	auipc a1,0x4
2280	000000b4	00004597	auipc a1,0x4
2300	000000ac	8a858593	addi a1,a1,-1880
2390	000000b0	00004597	auipc a1,0x4
2410	000000b4	00004597	auipc a1,0x4
2430	000000ac	8a858593	addi a1,a1,-1880
2520	000000b0	00004597	auipc a1,0x4
2540	000000b4	00004597	auipc a1,0x4
2560	000000ac	8a858593	addi a1,a1,-1880

V. FUTURE VISION

The testbed allows an automated execution; therefore, an executor will be developed comprising a suitable scenario definition module. After the execution, a big dataset comprising output files will be stored.

The big dataset allows to examine machine learning and deep learning algorithms to detect errors in an automatic way. Neural networks will be analyzed to classify the detected error. An automated testing system embodying the testbed will be designed to find bugs and errors in an application based on RTOS; the system will be used in the integration testing phase of the V-model of software development.

VI. CONCLUSION

The testbed embodying RISC-V soft-core was redesigned to support an automated way of operation. The output file was redesigned to collect data in more detail enabling an advanced analysis. The testbed structure was modified to follow the standard FPGA test-bench structure; the random generator module was replaced by the fixed scenario list allowing boundary testing. The testbed executes test scenarios that are configurable by a prescription and input file. Some of the FreeRTOS applications were created comprising deadlock, livelock, race condition, and starvation errors. The output files of the applications under the specific scenario show that the errors affect strongly the output file data. In the future, machine learning and deep learning algorithms will be examined to detect errors in the output files.

ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-23-8451 - “Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence,

automation and measurement" financially supported by the Internal science fund of Brno University of Technology.

REFERENCES

[1] J. C. Knight, "Safety critical systems: challenges and directions", Proceedings of the 24th International Conference on Software Engineering. ICSE 2002, pp. 547-550, May 2002

[2] FreeRTOS - Market leading RTOS (Real Time Operating System) for embedded systems with Internet of Things extensions [online]. In: . [cit. 2023-03-11]. Available from: https://www.freertos.org/

[3] RISC-V International – RISC-V: The Free and Open RISC Instruction Set Architecture. RISC-V International – RISC-V: The Free and Open RISC Instruction Set Architecture [online]. Copyright © 2021 RISC [cit. 12.03.2023]. Available from: https://riscv.org/

[4] NOLTING, Ing. Stephan. The NEORV32 RISC-V Processor: Datasheet [on- line]. In: . [cit. 2023-03-11]. Available from: https://stnolting.github.io/neorv32

[5] PRUSÁK, Lukáš. "Automated testbed for SIL/PIL testing of the embedded application using FPGA" [online]. In: . [cit. 2023-03-10]. Brno, 2022. Master's Thesis. Brno University of Technology. The Faculty of Electrical Engineering and Communication. Department of control and instrumentation. Supervised by Jakub Arm. Available from: https://www.vut.cz/studenti/zav-prace/detail/142143

[6] Y. Hong-Jie and M. Hua, "Simulation Implementation of MCU Based on Multisim10 and C Language," 2012 International Conference on Industrial Control and Electronics Engineering, Xi'an, China, 2012, pp. 2022-2024, doi: 10.1109/ICICEE.2012.537.

[7] A. J. Salim, S. I. M. Salim, N. R. Samsudin and Y. Soo, "Customized instruction set simulation for soft-core RISC processor," 2012 IEEE Control and System Graduate Research Colloquium, Shah Alam, Malaysia, 2012, pp. 38-42, doi: 10.1109/ICSGRC.2012.6287132.

[8] R. He, J. Chen, C. Du and Y. Duan, "Research on Embedded System Simulation Technology," 2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 2022, pp. 2597-2601, doi: 10.1109/ITAIC54216.2022.9836882

Spray-dry Synthesis of Perspective Cathode Materials for Li-ion Batteries with Low Cobalt Content

Bc. Hana Hálová
Department of Electrical and Electronic Technology
Brno University of Technology
Brno, Czech Republic
xhalov01@vutbr.cz

Ing. Ladislav Chladil, Ph.D.
Department of Electrical and Electronic Technology
Brno University of Technology
Brno, Czech Republic
chladil@vut.cz

***Abstract*—In this paper the possibilities of synthesis of cathode materials with low cobalt content is defined. The work compares the different kinds of synthesis in terms of time, simplicity and needed temperature. Spray drying synthesis was defined as the most promising one especially regarding using low temperatures and also because of the fast process of synthesis. In the other part of this paperwork, attention was paid to the construction of the apparatus and sample preparation itself. According to the SEM characterization the prepared powder contains particles of size from 300 nm to 1.1 µm while more than 70 % of the particles were smaller than 600 nm. From the results of elemental analysis, the distribution of Mn, Co and Ni was uniform and corresponded with similar spray-drying synthesis on professional devices.**

Keywords—low cobalt content, spray drying, synthesis, Li-ion batteries

I. Introduction

When selecting a cathode material, it is necessary to evaluate several basic parameters of electrode materials, including the electrode potential with respect to lithium, the specific capacity of the material (here it is good to observe both the theoretical value and the currently achieved values, which allow us to identify the current potential of the research work), the total amount of stored energy, the availability of individual elements and recycling options.

Cobalt is a common element used in the cathodes of Li-ion batteries nowadays. It is used because it increases energy density and stability, which in return increases its performance and lifespan. However, the growing tendency of reducing the amount of cobalt is the current trend together with finding alternative materials that can provide similar or even better properties [1][2].

The state-of-art research carried out in the field of cathode materials revealed that in terms of the energy stored in each cathode material, it is useful to observe the value of the electrode voltage with respect to lithium, where the highest values are achieved by $LiCoPO_4$ (4.8 V), $LiNi_{0.5}Mn_{1.5}O_4$ (4.7 V), $LiVPO_4F$ (4.2 V) and $LiMnPO_4$ (4.1 V) [1] [3] [4].

In terms of theoretical specific capacitance, $LiCoO_2$ (274 mAh/g) together with $LiMn_{1/3}Ni_{1/3}Co_{1/3}O_2$ (280 mAh/g) were selected as promising cathode materials. In the case of cobalt-free materials, $LiNiO_2$ (205 mAh/g) and $LiMnPO_4$ (175 mAh/g) were used.

In the case of specific capacitance values already obtained experimentally, the highest specific capacitance is achieved by $LiNi_{0.8}Co_{0.15}Al_{0.05}O_2$ (182 mAh/g) and $LiMn_{1/3}Ni_{1/3}Co_{1/3}O_2$ (160 mAh/g). In terms of further development prospects, the $LiCoO_2$ and $LiMn_{1/3}Ni_{1/3}Co_{1/3}O_2$ materials have the greatest potential, which, despite their high theoretical capacity values, reach only about 52% and 64% of their theoretical capacity, respectively, in the papers published so far.

Consequently, the gravimetric energy density is related to the value of the stress against lithium and to the specific capacity. A comparison shows that the highest commonly achievable energy densities are achieved by $LiCoPO_4$ (800 Wh/kg), $LiNi_{0.1}Co_{0.9}O_2$ (770 Wh/kg), $LiNiO_2$ (710 Wh/kg), and $LiMnPO_4$ (697 Wh/kg) [1] [3].

If we evaluate all the individual criteria and also consider the cobalt content, stability during cycling, and capacity drop, it is interesting for future experiments to focus on the $LiMn_{1/3}Ni_{1/3}Co_{1/3}O_2$ material or variants with reduced cobalt content, as well as $LiMnPO_4$ and $LiNi_{0.5}Mn_{1.5}O_4$ [4] based materials. These materials stand out in particular for their high potential against lithium (more than 4 V), while by appropriate doping with other elements this potential can be further increased (e.g. in the case of the use of Cr in $LiNi_{0.5}Mn_{1.5}O_4$ the voltage against lithium reaches 4.8 V) [4].

II. Synthesis of Cathode Materials with Low Cobalt Content

The perspective cathode materials can be manufactured by several types of syntheses [5]. **Solid-state synthesis** [6] which involves grinding and subsequent drying of precursors is one of the most common types of synthesis of cathode materials thanks to the low cost and simplicity of the process. However, the whole process of solid-state synthesis can last for more than 24 hours. Another possibility to synthesize the cathode materials is **a sol-gel method** [7], which stands out, especially thanks to the homogeneity, and use of low temperatures compared with solid-state synthesis. On the other hand, the whole process is time-consuming and can be relatively expensive. Another synthesis that can be used to prepare cathode materials for Li-ion batteries is **hydrothermal synthesis** [5]. The biggest advantage of the method is the production of highly pure as well as homogenous materials. Nevertheless, the whole process is limited to specific

materials, the use of high pressure and the synthesis time can be several days.

The unique type of synthesis is using the method of **spray-drying**. This method stands out especially thanks to the relatively fast process compared to other possibilities of synthesis and easy scalability for large-scale production [8].

III. SPRAY DRYING SYNTHESIS

The spray drying synthesis can be applied to suspension, solution, or their combination. However, it is needed to ensure the dissolving of precursors of the specific materials in a solvent, such as water, ethanol, or a combination of both. The suspension may contain additives, such as surfactants or dispersants that help to control particle size, prevent agglomeration, and improve the overall homogeneity of the final powder [8] [9] [10] [11].

The process involves atomizing a precursor solution into a spray, which is then dried using hot gas to form a powder. The apparatus for spray drying synthesis consists of: an **atomizer** used for spraying the solvent, a **drying chamber** where the sprayed solution is dried by hot gas and a **collecting chamber**, where the final powder is deposited. These main parts can be seen in Fig. 1.

A. Parameters Affecting the Synthesis

The parameters that affect the whole process can be divided into next 4 parts [9]:

- **Solution properties:** the final powder depends on the concentration of the solution as well as the particle size of precursors. The high viscosity of the solution can affect the size of the droplets and cause slower spray drying.
- **Atomization condition:** spray nozzle type, spray rate and gas flow rate influence droplet size and uniformity of the final powder product.
- **Drying conditions:** the inlet temperature and residence time can cause thermal degradation of the powder product or on the other hand insufficient drying of the solution.
- **Additives:** addition of certain additives e.g., surfactants or dispersants, can affect the droplet size, prevent agglomeration, and facilitate the spray-drying process

Even though the spray-drying synthesis is one of the most promising, especially for mass production, the use in the laboratory is limited by the price of the spray-drying device as well as necessity of a large amounts of precursors. Because of that it is necessary to adapt the aperture for laboratory usage and producing only the limited quantity of powder.

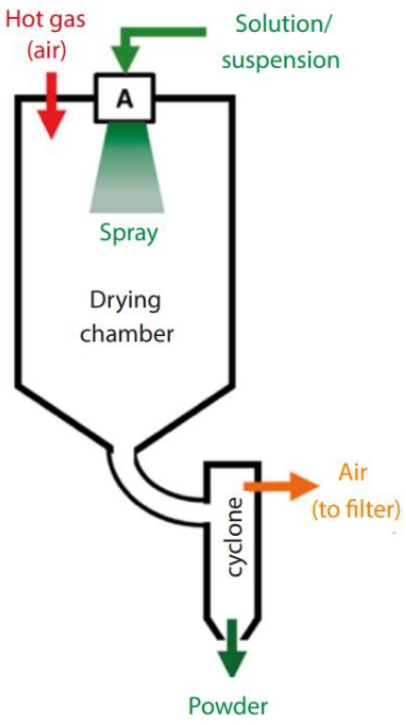


Fig. 1: The main parts of a spray-drying apparatus [8]

IV. MATERIAL SYNTHESIS

The cathode material in the experiment was prepared by spray drying synthesis. The schematic drawing of the laboratory apparatus which was used for the spray drying synthesis can be seen in Fig. 2. As mentioned above, the spray drying consist of an atomizer which sprays the solution to the air, drying chamber, where the droplets are taken by the flow rate of hot air dry, and the collecting chamber, where the final powder is collected.

The stoichiometric mixtures of $Li(OAc) \cdot 2H_2O$, $Ni(OAc)_2 \cdot 4H_2O$, $Mn(OAc)_2 \cdot 4H_2O$ and $Co(OAc)_2 \cdot 4H_2O$ were dissolved in deionized water in order to obtain clear solutions. The chelating agent, citric acid, was added in a molar ratio of 2:1. After that, the solutions were mixed, atomized, and dried in a drying chamber by hot air. During the synthesis, the temperature of the air in the drying chamber was set to 165 °C and the gas flow rate to 3 l/min. The synthesis time was 1 h.

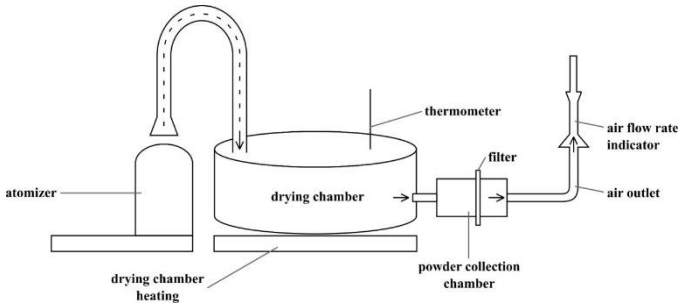


Fig. 2 Aparatus and schematic principle for spray drying synthesis

V. Results and Discussion

The spray drying synthesis can be applied to suspension, solution, or their combination. However, it is needed to ensure the dissolving.

The amount of the light violet powder from spray drying synthesis was approx. 40 mg of NMC. The powder was examined by the Scanning Electron Microscope (SEM). In Fig. 3 it is visible rounded particles of the final powder of NMC. For determining the results and better acknowledgment of the spray-dried powder, diameters of 130 random particles from the SEM picture were measured and size distribution profile were performed. The distribution of particles is separated in intervals based on their size is visible in Fig. 4. The size of the particles varies in the range of 300 nm to 1,1 µm. More than 70 % of the particles were smaller than 600 nm and only a few of them were bigger than 1 µm.

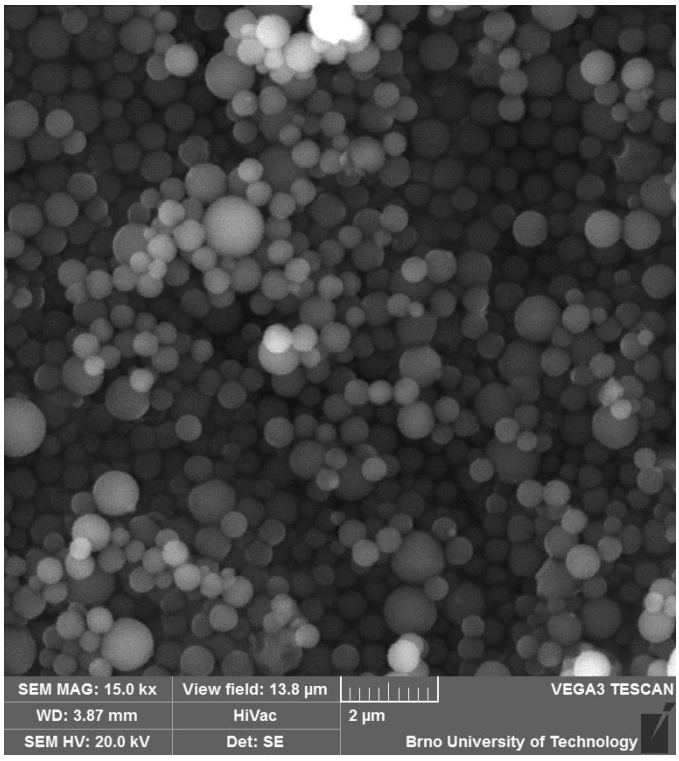


Fig. 3: SEM picture of spray-dried cathode material

According to the elemental analysis the powder consists of manganese, cobalt, nickel, and oxygen with carbon. The spectrum from energy dispersive spectroscopy can be seen in Fig. 5. Element compositions were determined by EDS analysis and the results are summarized in table (Table 1). Results suggest that the ratio between Ni/Mn/Co elements is approximately 1/5/1 which pointed to high manganese content. These results are only indicative because of high organic content which limits the accuracy of measurement (see right column of 3 Sigma error). The element distribution map (Fig.6) indicates that detected elements (Ni, Mn, Co, and C) are uniformly distributed within the volume of samples.

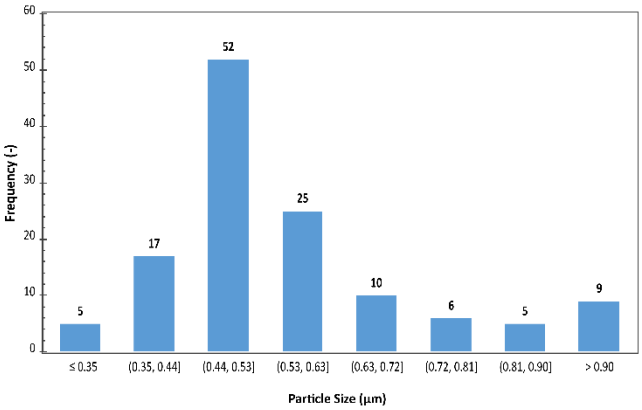


Fig. 4: Particle size distribution of cathode material by spray-drying synthesis (see Fig. 3)

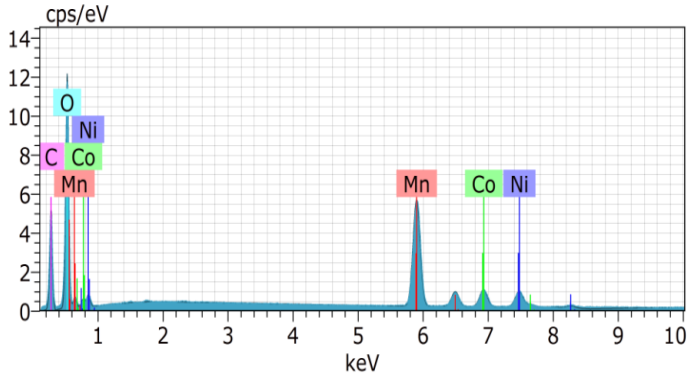


Fig. 5: The EDS of synthetized powder by spray-drying synthesis

TABLE I. The results of EDS element analysis of cathode material by spray-drying synthesis

Spectrum:	**258**				
Element	**AN**	**Series**	**Norm. (wt. %)**	**Atom. C (at. %)**	**Error (3Sigma) (wt.%)**
Carbon	6	K-series	28.62	40.60	5.28
Oxygen	8	K-series	49.52	52.74	8.52
Cobalt	27	K-series	3.22	0.93	0.20
Nickel	28	K-series	2.89	0.84	0.19
Manganese	25	K-series	15.76	4.89	0.70
		Total:	100.00	100.00	

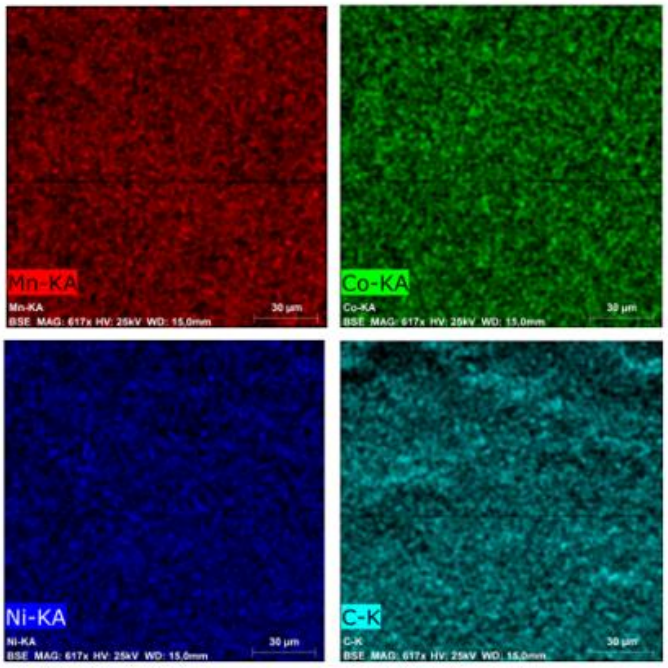


Fig. 6: Distribution of the individual elements (Mn, Co, Ni, C)

VI. Conclusion

The paper discussed several possible syntheses for cathode materials of Li-ion batteries with low cobalt content. For the experiment the spray-drying synthesis was chosen – thanks to the advantages of fast powder formation and simplicity.

The using of common devices instead of the expensive specialized apertures intended directly for the purpose of spray-drying synthesis had no visible effect on the size of particles or the distribution of elements and most of the particles are on the order of hundreds of nanometers and 70 % are even smaller than 600 nm.

Currently, the main disadvantage lies in the low yield of the above synthesis, which is in the order of milligrams. Another problem is the difficulty of air filtration with nanoparticles and their permeation through standard air-permeable filters. Thus, in future work, we will try to optimize the use of filters with different porosity and, together with this, aim to increase the yield of the realized synthesis to the level of hundreds of milligrams, which will allow further practical verification of such synthesized active materials for use in a laboratory-assembled experimental cell.

References

[1] N. Nitta. and F. Wu, "Li-ion battery materials: present and future," Materialstoday [online]. 2015, 18 (5), pp. 252-264, available from: https://doi.org/10.1016/j.mattod.2014.10.040

[2] B. Xu and D. Qian, "Recent progress in cathode materials research for advanced lithium ion batteries," Materials Science and Engineering: R: Reports [online]. 2012, 73(5-6), pp. 51-65, available from: https://www.sciencedirect.com/science/article/pii/S0927796X12000186

[3] T. Kazda, "Modifikace materiálů pro kladné elektrody Lithno-Iontových akumulátorů," Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2015.

[4] M.S. Whittingham, "Lithium Batteries and Cathode Materials," *Chemical Reviews* [online]. 2004, 104, pp. 4271-4301, available from: https://pubs.acs.org/doi/pdf/10.1021/cr020731c

[5] T.A. Wani and G. Suresh, "A comprehnsive review of $LiMnPO_4$ based cathode materials for lithium-ion batteries: current strategies to improve its performance," Jounal of Energy Storage [online]. 2021, 44, part B, available from: https://doi.org/10.1016/j.est.2021.103307

[6] L. Zhang and Q. Qu, "Confined synthesis of hierarchical structured $LiMnPO_4$/C granules by a facile surfactant-assisted solid-state method for high-performance lithiumion batteries," Journal of Materials Chemistry A [online]. 2014, 3, pp. 711-719, available from: https://doi.org/10.1039/C3TA14010E

[7] L. Wu. a S. Zhong, "Synthesis of Cr-doped $LiMnPO_4$/C cathode materials by sol–gel combined ball milling method and its electrochemical properties". Ionics [online]. 2013, 19, pp. 1061-1065, available from: ttps://doi.org/10.1007/s11581-013-0919-9

[8] C. Greenmat, "Spray-Drying of Electrode Materials for Lithium- and SodiumIon Batteries" [online]. Materials, 2018, 11(7), available from: https://doi.org/10.3390/ma11071076

[9] A. Stunda-Zujeva, "Controlling the morphology of ceramic and composite powders obtained via spray drying – A review" [online]. Ceramics International, 2017, (15), pp. 11543-11551, available from: https://doi.org/10.1016/j.ceramint.2017.05.023

[10] Y.J.J Li. and W.CH. Chien, "Electrochemical performance of spherical Li-rich LMNCO cathode materials prepared using a two-step spray-drying method," [online] Ceramics International. 2022, 48(5), pp. 6302-6312, available from: https://doi.org/10.1016/j.ceramint.2021.11.173

[11] T.A. Wani a G. Suresh, "A comprehensive review of $LiMnPO_4$ based cathode materials for lithium-ion batteries: current strategies to improve its performance," Journal of Energy Storage [online]. 2021, 44. ISSN 103307, available from: https://www.sciencedirect.com/science/article/pii/S2352152X21009993

Analysis of UAV sensors data

1st Grégoire ISSASSIS
Faculty of Electrical Engineering and Communication
BURES-SUR-YVETTE, France
254715@vut.cz

2nd Petr MARCOŇ
Department of Theoretical and Experimental Electrical Engineering
BRNO, Czech Republic
marcon@vut.cz

3rd Jiří JANOUŠEK
Department of Theoretical and Experimental Electrical Engineering
BRNO, Czech Republic
jiri.janousek@vut.cz

***Abstract*—The aim of this project is to design and implement software for controlling unmanned aircraft without GPS. This software will be implemented in the control module of the UAV. The basis for the development of the software will be the comparison of information from the GPS sensor and other sensors placed on the UAV.**

Keywords*—*UAV, drone, GPS, Barometer, Compass, Gyroscope, Accelerometer, IMU, Sensor fusion, Pitch, Roll, Yaw.

I. Introduction

Nowadays, drones are used more and more, both in civilian and military operations. Indeed, the most common uses we know of these machines are for leisure or military operations, for example. However, drones are used in many other fields that we do not necessarily think of, such as agriculture, cinema, surveillance, or even first aid, to name but a few. Drones therefore occupy an important place in society that will continue to grow in the coming years.

As with all advanced technology, drones carry sensors for real-time monitoring and control. As drones operate in stressful conditions for the on-board electronics, it is essential to ensure that their sensors can take over from each other in the event of a failure. Thus, drones must be able to continue to function if one of the sensors fails, relying on the other on-board sensors. This problem can be related to instrumentation flight in aviation. Indeed, when a flight data is wrong or insufficient, pilots must know how to use the other data to continue the flight. The same applies to drones. In the event of a GPS failure, the drone must be able to continue its flight to meet its objectives, especially when these are critical objectives.

Also, there are many situations where drones cannot rely on GPS signals to maintain their route. Indeed, in secure areas where GPS signals are jammed, or indoors, where the signal is weak, ensuring flight viability is a major issue for drone specialists.

Finally, GPS is an old technology. Launched in 1995 by the United States, the first experiments with GPS date back to the early 1970s. The missions carried out by drones require more and more precision. However, GPS is reaching its technical limits, which no longer match the accuracy required by drones.

II. Methods

A. Defining a problematic

The first part of the project methodology was to find a problematic. As with any engineering project, we must find solutions to problems. To find a solution that is perfectly adapted to drone users, we had to ask ourselves about their problems. As we have seen before, improving the monitoring and control of a drone at distance is essential. A drone has many sensors, all with different purposes. However, coupling data from several types of sensors can provide similar results to those given by other sensors.

Some sensors, such as GPS, are based on sometimes outdated technology, with an accuracy that may be considered adequate in some use cases, but insufficient in critical operations requiring high accuracy. Furthermore, in the event of a failure of one or more sensors, creating redundancy is one of the most important issues in aeronautics, to be able to continue flying in the event of a failure.

Thus, our team was able to ask questions around various issues: “How UAVs interacts with their environment to compute flight data?”.

B. Choosing a research method

There are several research methods to answer a problem. These methods include the analysis of scientific and academic literature, surveys, interviews, etc.

As we only had a short time to complete our study, we opted for a literature review. There is a large body of literature, both academic and non-academic, on the development, use and maintenance of drones. Closely related to aeronautics and model aviation, the practice of UAVs is not only confined to military uses. There is therefore an impressive database of documents on the practice of UAVs, written by professionals in the sector or by enthusiasts.

All the documentation that enabled us to carry out this project are quoted in this document.

C. Data collection

To carry out this project, we needed to collect a lot of data. Indeed, we needed to compare data from different sensors of a drone, during a complete flight. To do this, we carried out a test flight, during which all the data from all the drone's sensors were recorded. Among the sensors that make up our drone, we find:

- **Airspeed**: sensors that measures the speed of an aircraft relative to the surrounding air, using the differential between the pressure of still air and that of moving air compressed by the craft forward motion;
- **Barometers**: sensors that detect atmospheric pressure;
- **Camera:** optical instrument that captures images of the environment. Coupled with IMU, a camera can achieve visual odometry, which is the process of determining the position and orientation of the drone by analyzing the associated images;
- **Compass**: sensor designed to detect and respond to the magnetic field of the Earth. Items in this family are designed for dead reckoning, digital compass, or a digital magnetometer in the X, Y, Z axis;
- **GPS (Global Positioning System)**: receivers with antennas that use a satellite-based navigation system with a satellite constellation in orbit around the earth to provide position, velocity, and timing information;
- **Gyroscope**: advanced accelerometers that can measure and maintain the orientation and angular velocity of an object. These sensors can measure the tilt and lateral orientation of the object whereas accelerometer can only measure the linear motion;
- **IMU (Inertial Measurement Unit)**: device composed of multiple sensors that can measure and report specific gravity and angular rate of an object to which it is attached. It consists in the addition of gyroscopes and accelerometers sensors;
- **Rangefinders**: also called distance sensors, these sensors work on the principle of reflected light waves. An infrared light is sent from an infrared remote or beacon, then reflected. The reflected light is detected and then an estimate of distance is calculated between sensor and object.

Our drone has many other sensors on board. However, analyzing the data from the other sensors will not be of any interest in our project. Therefore, we will focus on the sensors mentioned above throughout the project. All these sensors are linked to an on-board computer, which operates as the brain of our drone. The drone does not require high computational capabilities. However, the on-board computer needs to be very liable, small, lightweight and energy saving. So, we have opted for a NVIDIA Jetson which is often used on drones and other radio commanded aircrafts. It is very similar to a RASPBERRY Pi. At the end of our test flight, all the data provided by the sensors was recorded on the on-board computer. We were able to retrieve this data using a log file. This is a file extension that can be compared to the logbook of a ship's captain. All the data from all the sensors are listed in a raw form.

D. Data analysis

Once all the data from the flight could be recorded in the log file, we had to start analyzing it, using software. This software allowed us to reorder the raw data by categorizing it, and to display the data in graphical form to find correlations.

Here is the software we used to complete our project:

- **APM Planner 2.0**: is a GCS (Ground Control Station), which is a software application, running on a ground-based computer, that communicates with the UAV via wireless telemetry. It displays real-time data on the UAVs performance and position and can serve as a "virtual cockpit", showing many of the same instruments that you would have if you were flying a real plane. A GCS can also be used to control a UAV in flight, uploading new mission commands and setting parameters. It is often also used to monitor the live video streams from a UAV's cameras;
- **Cartographer**: system that provides real-time simultaneous localization and mapping (SLAM) in 2D and 3D across multiple platforms and sensor configurations;
- **PX4 Autopilot**: open-source autopilot system aimed at low-cost autonomous aircraft. Low cost and availability allow amateur use in small, unmanned aircraft. Mastering PX4 requires a lot of experience, and understand how it works, according to [4] and [5];
- **ROS** (Robot Operating System): it is the OS (Operating System) used by the onboard computer of the drone. We needed to master this OS to operate the drone, but also to calibrate the sensors and retrieve the data from them. To do so, we referred to the ROS documentation [6].

As we have seen before, not all the sensors on board of the drone will be useful for our project. The ones we will analyze are the airspeed, barometers, compass, GPS, IMU accelerometers, IMU gyroscopes, and finally, rangefinders. We will analyze the data from each of these sensors to understand how they work and how do they interact with their environment to collect useful data during the flights.

Before doing these analyses, it is important to understand what a sensor is and what its purpose is. A sensor is a device that detects and responds to some type of input from the physical environment. The input can be light, heat, motion, moisture, pressure, or any number of other environmental phenomena. The output is electrical voltage or intensity. Once we have retrieved the electrical output values, we can analyze them and determine correlations between the environmental variables, i.e., the physical values at the sensor input, and the output variables, i.e., the electrical current or intensity.

i. Airspeed sensor data analysis

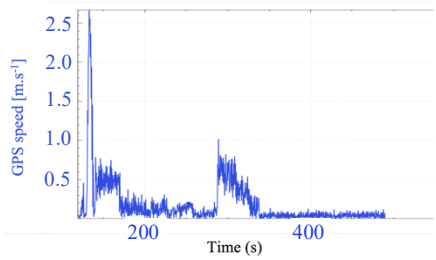


Fig. 1. Airspeed (from GPS) data analysis from APM Planner 2.0

Airspeed sensors are used to calculate the speed of a drone, in addition to the measurements given by the GPS sensor. The sensor is made of a tube placed on the drone, parallel to the ground and in the same direction as the front of the drone. When the drone is in flight, and moving, pressure is applied to it. This pressure enters the tube. The principle is to compare the pressure inside the tube with the pressure on the rest of the drone. Thus, we can estimate the speed of the drone according to the pressure delta. The higher the delta of pressure is, the faster the device goes.

ii. Barometer sensor data analysis

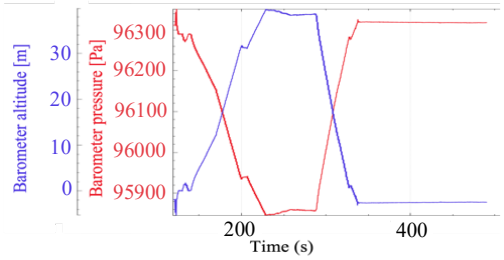


Fig. 2. Barometer data analysis from APM Planner 2.0

From the figure above, we have been able to see a correlation between pressure and altitude. Indeed, as altitude increases, pressure decreases. By calibrating the sensor at the take-off point, the altitude can be estimated as a function of the pressure delta between the take-off point and the current value.

iii. Compass sensor data analysis

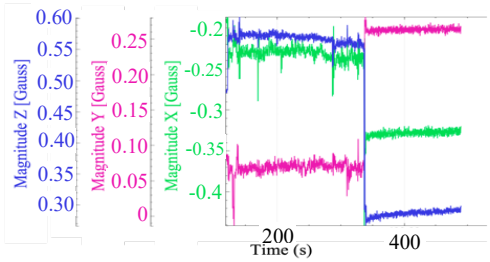


Fig. 3. Compass data analysis from APM Planner 2.0

The compass role is to measure the orientation of the drone in relation with the magnetic north of the Earth. To do this, it is composed of a fixed magnetic bar, bathed in a fluid that prevents any resistance in the medium. Thus, when the drone changes direction, the magnetic bar of the compass reorients itself according to the angle with the magnetic north. Thus, the direction of the drone can be known in real time via three coordinates along the X, Y, and Z axes in a three-dimensional reference frame.

iv. GPS sensor data analysis

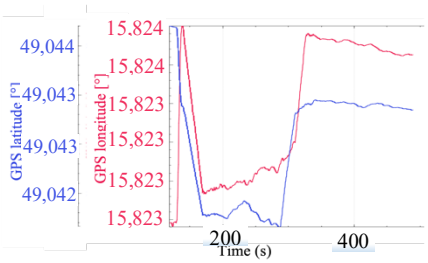


Fig. 4. GPS data analysis from APM Planner 2.0

GPS location requires a satellite constellation. These satellites send the time of transmission by microwave which travels at the speed of light. Knowing this speed and the time difference between the transmission and reception, the distance between the receiver and the satellite is determined. GPS sensors requires a minimum of 3 satellites from the constellation. Thus, we can estimate the distance to each satellite and make a triangulation, knowing that the satellites are geostationary. The coordinates are given according to two variables, latitude, and longitude. Accuracy is between 5 and 10m. This accuracy can be improved with the help of IMU and camera sensors to make visual-inertial odometry, as explained in [11].

v. Gyroscope sensor (IMU) data analysis

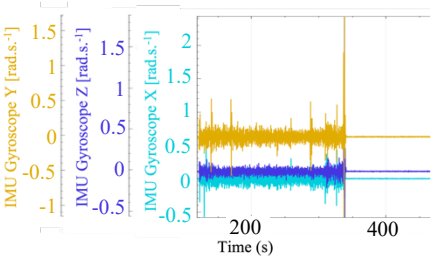


Fig. 5. Gyroscope data analysis from APM Planner 2.0

A gyroscope is a device that measures the orientation and movement of the drone by detecting changes in its orientation. It usually consists of a rotor that rotates at high speed and is mounted on a rotating shaft. When the gyroscope is in motion, the rotational force of the rotor tends to keep its axis of rotation in a constant direction, regardless of the movements of the object on which it is mounted. This property is called "gyroscopic stability" and allows the gyroscope to maintain a fixed orientation in space, even if the object it is mounted on is moving. Thus, the movement of the drone can be known in real time via three coordinates along the X, Y, and Z axes in a three-dimensional reference frame.

vi. Accelerometer sensor (IMU) data analysis

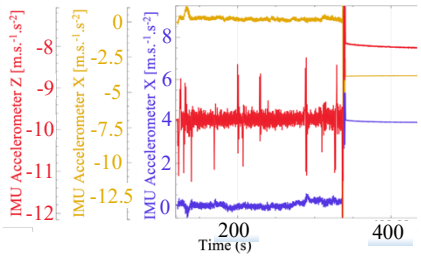


Fig. 6. Accelerometer data analysis from APM Planner 2.0

The accelerometer is a sensor capable of measuring, in three dimensions, the linear accelerations of an object as well as the vibrations. The accelerometer then converts the vibration energy into an electrical signal proportional to the instantaneous acceleration of the object. Thus, the movement of the drone can be known in real time via three coordinates along the x, y, and z axes in a three-dimensional reference frame.

vii. Rangefinders sensor data analysis

Rangefinders are sensors, but also actuators. As we have seen previously, a sensor is a device that detects and responds to some type of input from the physical environment. An actuator does the exact opposite. Indeed, actuators are devices that are responsible for moving and controlling a mechanism or system, converting electrical voltage or intensity, into a physical environment variable. Firstly, the rangefinder emits an infrared wave. This signal will reflect off an obstacle and back to the rangefinder. Thus, the rangefinder acts as a transmitter, as it is an actuator, but also as a receiver, as it is a sensor. The main goal is to calculate the time between emission and reception of the IR wave. The time is divided by two since the wave travels back and forth between the drone and the obstacle.

E. Data correlation

Now that we have been able to analyze data and understand how each of the sensors on our drone works, we will try to put this data together to make these sensors work together. The three most important types of sensors to work together are the barometer, compass and IMU.

i. Barometer and altitude correlation

The barometer gives us the altitude of the drone in real time. Thus, we will give the drone a desired altitude as input and make the drone go up or down, always aiming to reach this value. According to [7], we used the hypsometric equation to compute altitude in real time based on pressure sensor:

$$h = \frac{R \cdot Tv}{g} \ln\left(\frac{p1}{p2}\right) \quad (1)$$

h: altitude from calibration point [m]

R: specific gas constant = 287 J·kg^{-1}·K^{-1}

T_v: temperature [K] (1°C = 273.15K)

g: gravitational acceleration = 9,81 m.s^{-2}

$p1$: pressure from calibration point [Pa]

$p2$: instantaneous pressure [Pa]

ii. Compass and direction correlation

Aircraft convention defines the inclination parameters in terms of three angles: yaw (ψ), pitch (φ), and roll (θ). These angles are referenced to the local horizontal plane, which is the plane perpendicular to the Earth's gravitational vector.

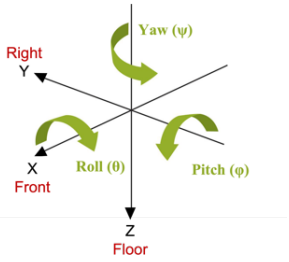


Fig. 7. Aircraft attitude in horizontal plane

The compass provides the yaw value, which is defined as the angle in the local horizontal plane measured clockwise from a true North (Earth's polar axis) direction. This sensor will allow us to orientate the drone towards the desired location using its coordinates. In case of roll and pitch, the Earth magnetic fields will not be perceived the same way as if there was not. So, here are the formulas we used to compute the direction from North, using a compass sensor:

$$X_h = X \cdot \cos(\varphi) + Y \cdot \sin(\theta) \cdot \sin(\varphi) - Z \cdot \cos(\theta) \cdot \sin(\varphi), \quad (2)$$

$$Y_h = Y \cdot \cos(\theta) + Z \cdot \sin(\theta), \quad (3)$$

$$\psi = \arctan\left(\frac{Yh}{Xh}\right). \quad (4)$$

However, as the compass relies on the magnetic fields of the Earth, this sensor is sensitive to magnetic fields other than those emitted by the Earth. Thus, the data from this sensor may be biased depending on the flight environment. Furthermore, the geographical orientation with respect to North, according to a plane parallel to the ground, and given by this type of sensor depends on the inclination of the UAV according to the other planes, in a three-dimensional reference frame. All these specificities about compass sensor are explained in [8] and [9].

i. IMU and inclination correlation

The main goals of the IMU are to reports specific dynamic states such as angular velocity and accelerations. From this data, we can estimate drone attitude angles during flight such as roll and pitch from horizontal plane. By entering an input value, we can move the drone forward by telling it to change its roll or pitch value by activating each motor independently.

To correctly calculate roll and pitch using the IMU, we need to define three acceleration values, aX, aY and aZ. These values are respectively the accelerations of the UAV along the X, Y and Z axes in the horizontal plane. According to [10], here are the equations we used to compute the pitch and roll values:

$$\varphi = \arcsin\left(\frac{aX}{g}\right), \qquad g = 9.81 \text{ m} \cdot \text{s}^{-2}. \quad (5)$$

$$\theta = \arctan\left(\frac{aY}{aZ}\right). \quad (6)$$

F. *Software implementation*

We have not yet been able to deal with this part of the project, but here are the next steps of the project we are considering.

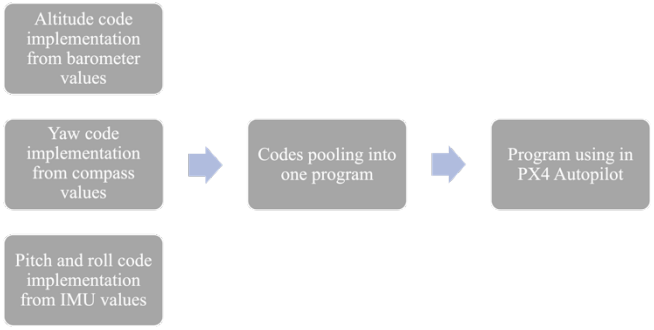


Fig. 8. Bloc diagram of software implementation for our project

As explained in the figure above, we will firstly implement three different codes, one for each of the data correlation formulas we have seen earlier. So, we will have a code that will calculate the altitude of the drone in real time based on the barometer data, using the hypsometric equation (1). We will also implement a code to compute the yaw in real time, using data from the compass within the formulas (2), (3) and (4). Finally, we will implement a code to calculate pitch and roll, using data from the IMU in formulas (5) and (6).

Then, once we will have viable codes for each of these correlation formulas, we can pool them together to retrieve in real time all the values required for flight automation in PX4 Autopilot. As a reminder, this software allows the drone to be piloted remotely, and with automation. Thanks to our software implementation, we will be able to enter the values recovered from the sensors as values to be reached for each of the actions to be carried out by the drone during the automated flight.

III. Discussions and Conclusion

As we have seen previously, UAVs carry many sensors that allow them to have access to a lot of real-time flight information, as flight instruments in aircrafts. Thus, our team was able to ask questions around various issues: “How UAVs interacts with their environment to compute flight data?”.

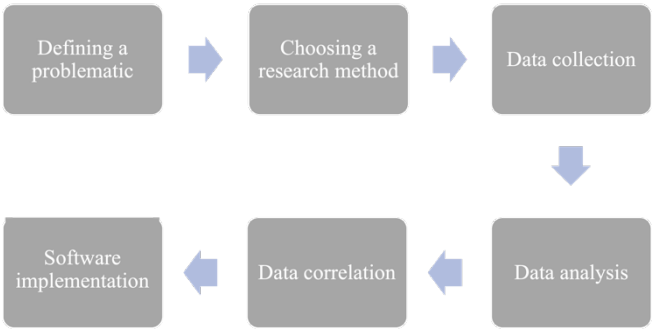


Fig. 9. Bloc diagram of project planification

Here are the methodologies we used to carry out our project. Firstly, we sought to define a problem. Then, we looked for the appropriate research method to implement our project, considering its specificities and the fact that we had limited time. Then we collected all the data that would be necessary for the project. Then came the phase of analyzing this data, so that we could understand how our drone's sensors works. We have then correlated this data with the results of our analysis, so that we could begin to answer our problem. Finally, we will implement the code that will allow us to validate, or not, our hypotheses.

For the moment, our project is not complete. Indeed, even if we have been able to define the perimeter of our project, define a problem, collect the necessary data, analyze, and correlate these data, we still have some work to do. The code implementation will be the last part of our project, that is still remaining.

This project can also be the basis for larger projects about onboard sensors of an UAV. Indeed, the embedded sensors, such as GPS, are under great pressure during flight, cannot be used in all conditions, or are based on outdated technologies. For example, a possible project resulting from this one could be: *Relative positioning of UAV*. The problematics related to this project would be: "Can we control a drone without using a GPS sensor, and how? Is the accuracy obtained by this method more satisfactory than GPS?”. Once we would have figured out how to influence the sensor output variables with commands sent to the drone, we could give limit values and values to reach for each sensor. As explained in [12], sensor fusion is a fundamental aspect of an UAV autonomous flight. Thus, a flight path to be flown would be broken down into a sequence of steps for the drone to perform, seeking to achieve previously established values for each sensor.

References

[1] ArduPilot Development Team, “Non-GPS navigation” , 2023.

[2] ArduPilot Development Team, “ROS with SITL in Gazeboo” , 2023.

[3] Gazebo, “Gazebo Simulation Documentation”, 2023.

[4] PX4 AutoPilot, “PX4 AutoPilot user guide” , 2022.

[5] PX4 AutoPilot, “Flight Log Analysis” , 2022.

[6] ROS, “ROS.org Documentation” , 2023.

[7] American Meteorological Association, “Hypsometric Equation”, 2015.

[8] Honeywell, M.J. Caruso, “Applications of Magnetic Sensors for Low Cost Compass Systems”, 2002.

[9] Honeywell, M.J. Caruso, “Applications of Magnetoresistive Sensors in Navigation Systems”, 1997.

[10] M. Wrona, “Roll and Pitch Angles From Accelerometer Sensors”, 2020.

[11] J. Liao, X. Li, X. Wang, H. Wang, “Enhancing navigation performance through visual-inertial odometry in GNSS-degraded environment”, 2021.

[12] J. Janoušek and J. Semeraád, “Fusion of sensors for autonomous UAV Navigation”, in Proceedings II of the 28th student EEICT, 2022 Selected Papers. 1, Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2022, p. 63-66, ISBN: 978-80-214-6030-1

The Crane Manipulator For The 3D Printer

Bc. Denys Khmil
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
xkhmil00@vutbr.cz

Ing. Jakub Arm Ph.D.
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
jakub.arm@vut.cz

***Abstract*— The work address the assembly and improvement of a machine for automatic servicing of 3D printers. The machine unloads the finished printed object and loads a new printing surface. The planar axes are equipped by closed-loop stepper motor drives to achieve better motion parameters and checking the proper position; the firmware of the drive controllers is upgraded to support CAN communication and S-curve motion. BeagleBone Black is used as the main control unit of the machine; therefore, the extended PCB was designed to connect sensors and unloading manipulators. The control system uses G-code oriented communication protocol. Preliminary results suggest that the machine will operate correctly and fulfill its purpose.**

Keywords—Manipulator, CNC, Stepper motor, Driver, PCB design.

I. INTRODUCTION

The concept of automation has become an integral part of modern industry and business, leading to significant economic benefits and increased efficiency. The field of 3D printing is no exception. The 3D printer itself was a major step forward in the plastic parts production, but it still required human intervention at the beginning and end of the printing process. This article is dedicated to a device designed to automate the unloading of finished products from of 3D printers, with the goal of minimizing human intervention in the printing process. This should bring financial benefits and significantly reduce labor costs in cases where 3D printing is used on a mass scale.

The mentioned above issue was discussed in the bachelor's thesis "The crane manipulator for the 3D Printers"[2] which proposed a device concept for automating the unloading products from a farm of several 3D printers. The thesis was dedicated to the development of a concept, calculations, 3D model creation, and simulation of the device in Siemens NX12 software.

The aim of this thesis is to improve the functionality and reliability of the concept, as well as to build a real prototype of the device with some changes to its design.

II. CONCEPT OF THE MANIPULATOR

The manipulator is a device that can move along the front side of the 3D printers placed next to each other. The function of the manipulator is to pick up the printed part along with the printing surface from the heating bed of the 3D printer after printing is complete and place it in storage. Then, a new, empty, printing surface is taken from storage and placed on the heating bed, allowing the printer to begin the next print.

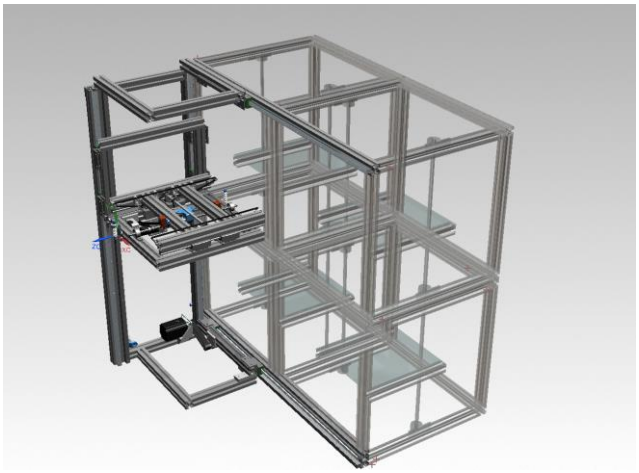


Figure 1 Manipulator 3D model [2]

The manipulator is designed as a 3-axis CNC machine, where X is a horizontal axis, Y - vertical, and Z axis is presented as the mechanism for taking the printing surface from the printer. Figure 1 shows a model of the manipulator that was created in the program Siemens NX12. More details about the mechanical design were described in the article "The crane manipulator for the 3D printers"[1].

During the manipulator prototype assembly, it was decided to redesign some of the manipulator systems to improve overall functionality. First of all, it was decided to decentralize the control of the manipulator.

The concept of the device assumed that Beaglebone Black (mainboard) would be responsible for controlling the motors by sending STEP/DIR commands. In the new concept, the control system is divided into several subsystems that will communicate with each other using CAN.

To implement the new concept, it was decided to completely abandon the use of STEP / DIR drivers, since this solution has low reliability and requires a lot of wires. Instead, it was decided to use a smart motor drivers that can be controlled using CAN communication.

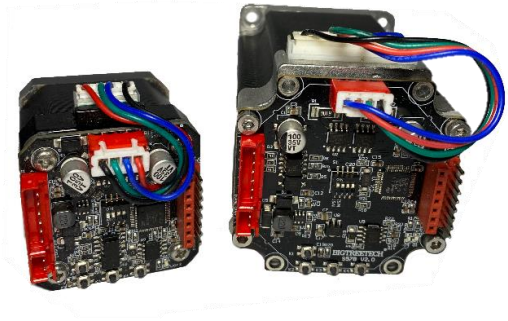


Figure 2 S42 and S57 drivers

The control board for the Z-axis system also required changes. Instead of a simple Arduino-based board, an advanced board was developed. The new board is capable of communicating using CAN and has a motor brake actuator control circuit. Additionally, the board serves as a "hub" for motor drivers, with all three motors located on the Z-axis carriage connected to the board.

III. Motor drivers

In the original manipulator concept, drivers from Makerbase were used. However, after more detailed testing, it became clear that the drivers did not meet the requirements. The following problems were found: the inability to use digital protocols for drivers control and configuration (configuration is only limited to USB), as well as the inability to load custom firmware (lack of an interface for firmware flashing on the board). This, together with the absence of a CAN interface, was the reason for abandoning the use of these drivers.

Instead of these drivers, drivers with a similar conception from the company Bigtreetech were chosen (Fig. 2). These drivers also attach to the back of the motor and have feedback through a magnet placed on the rotor of the motor. Unlike the Makerbase drivers, they have an SWD interface for firmware flashing and UART for configuration. Also, the drivers have a CAN transceiver chip that can be used for CAN communication. However, these drivers also have drawbacks: by default, it is only possible to control the motor using STEP/DIR commands, UART is only used for driver configuration, and the CAN transceiver is not used at all. Taking all the drawbacks into account, the driver firmware needs several modifications to allow controlling the drivers using CAN. Since the drivers are based on the popular STM32F103 microcontroller, this task is achievable.

First of all, the drivers were flashed with the provided binary file to test their functionality, and they appeared to be working properly. As mentioned before, the drivers are controlled with STEP/DIR signals and are able to compensate for skipped steps. After the testing, it was decided to study the firmware from the manufacturer in order to find out its capabilities and potential for improvements. During the study, it was determined that the code was incomplete. Attempts to compile the project ended with a large number of errors. Upon a more detailed examination of the code, it was determined that the code is incomplete. The request for complete code was sent to the manufacturer, but no response was received.

Since modification of the manufacturer code was not the option, there are two possibilities left: to find a suitable project on the internet, or to make a new firmware from the beginning. The first option that was considered was to find a suitable project on the internet. After bit of searching an opensource project dedicated to these drivers was found. In this project a group of volunteers was developing firmware with extended functionality that covered all the capabilities of the driver. However, the project was not completed, and only core functions of the firmware were operational:

- Driver configuration with buttons and OLED display.
- UART and CAN communication using G-Code based protocol. CAN communication was implemented but not used.
- Stepping functions.
- STEP, DIR, ENABLE inputs processing.

Features that have not been implemented or require major improvements:

- PID regulation needed a lot of improvements.
- Control using CAN/UART commands required bugfixes.

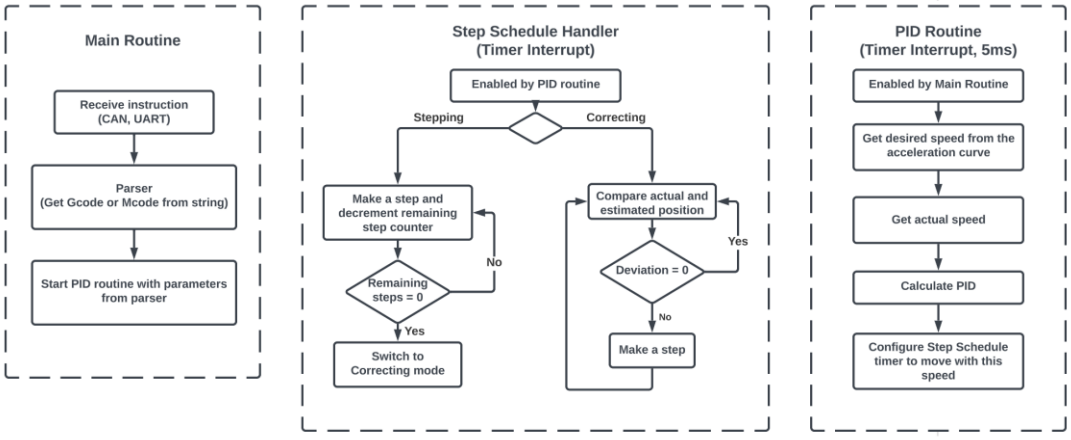


Figure 3 Driver algorithm

- Acceleration control
- Sending notification to the mainboard when done stepping.
- Error notification when destination can't be reached.

After evaluating all the pros and cons, it was decided that this project is a good foundation for custom firmware development for the drivers. After studying the source code of the project, it was decided to make some changes to the main algorithm of the driver's operation. The following algorithm was proposed (Fig. 3): after powering up the driver, the main routine is launched, which is responsible for processing commands received via UART or CAN, as well as for servicing buttons and an OLED display if it is connected to the driver.

The PID routine is responsible for configuring the movement speed. It takes the desired speed from the previously generated acceleration curve and follows it with a PID regulator. It is runs using timer interrupts with a fixed period of 5 ms. The routine starts after the command was received and processed by the main routine. Every 5 ms, the desired velocity is calculated using the acceleration curve and PID controller, which is then used to configure the stepping routine.

The step planner routine, same as PID routine, also works in interrupts, but the difference is that the interrupt period here is dynamic and is adjusted by the PID routine in proportion to the calculated speed. The step planner has two modes: stepping mode and correction mode. In the stepping mode, the driver counts the steps taken. This value is then compared with the ordered steps to be sure that the correct number of steps have been taken. In correction mode, the driver does not count the number of steps taken, but simply corrects the position if necessary based on feedback from the encoder.

To implement smooth movements and unload the main controller, it was decided to implement an acceleration and deceleration algorithm on the motor driver. This algorithm will be implemented in two options. The first option is ramp acceleration. This method is implemented by applying constant acceleration and deceleration at the beginning and end of the movement. On the driver side, this is implemented by increasing or decreasing the speed by an constant amount over time until the desired speed is reached.

The next algorithm is called S-curve acceleration. This is a more complex algorithm that is used in situations where high smoothness of movement and absence of resonances are required. Since the manipulator has large dimensions, smooth movement is desirable. This is achieved through a gradual change of the acceleration parameter. Instead of applying maximum acceleration at the beginning of movement, as in the first case, the acceleration also changes over time: it gradually increases at the beginning, reaches its maximum value, and then gradually decreases. This creates a speed curve that can be calculated using equations:

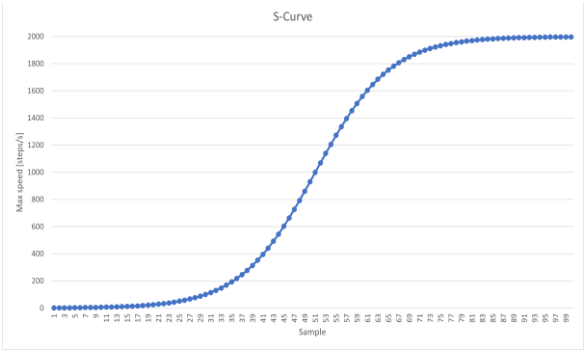


Figure 4 Generated S-curve

$$S_Curve[i] = \frac{MaxSpeed}{1 + \exp\left(-accSlope * (pointOnCurve - slopeMidPoint)\right)} \quad (1)$$

Where:

$$slopeMidPoint = AccelerationTime/2 \quad (2)$$

$$sCurveElementsNum = \text{AccelerationTime/SamplingTime} \quad (3)$$

$$pointOnCurve = SamplingTime * i \quad (4)$$

Figure 4 shows S-curve generated for the following parameters: movement at maximum speed of 2000 steps/s (900 deg/s), the driver should reach maximum speed within 0.5 s.

The Fig. 5 and Fig. 6 shows the results of both algorithms. The graphs depict the calculated speed and the speed measured by the magnetic encoder located on the driver. Despite the fact that the measured value is slightly noisy due to quantization noise, it can be seen that the driver well copes with the task of following the desired speed.

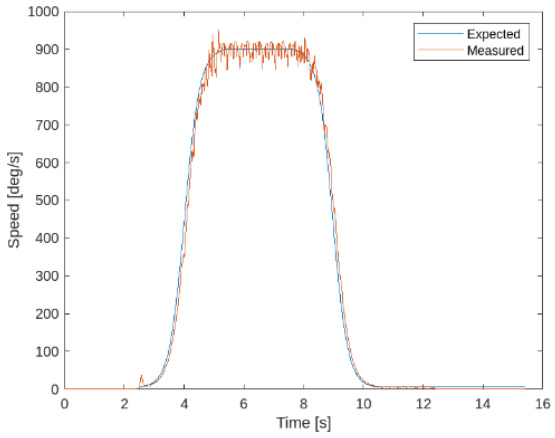


Figure 5 S-curve acceleration

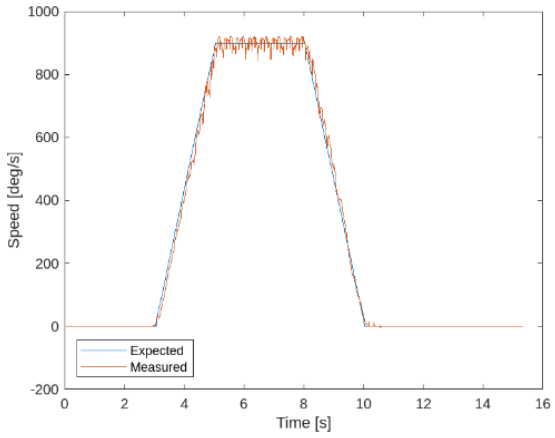


Figure 6 Ramp acceleration

IV. Custom PCB for Z-axis control

The next step is to manufacture a PCB for Z-axis control. Since the concept of the manipulator has changed, the requirements for the PCB will be as follows:

- Support of CAN communication
- Ability to connect up to 3 stepper motors with S42 drivers.
- Ability to connect 4 induction limit sensors.
- Ability to control brakes actuator.
- I2C and RS485 ports for potential extensions.

First of all, it was necessary to choose a microcontroller that would control the board. The STM32F103 microcontroller was chosen, which meets all the requirements since it supports CAN communication, has enough I/O pins, and also has the required interfaces to satisfy the requirements above.

As the board will be powered by 24 V, the next step is to design the power supply circuit. The board requires 3 voltage levels to function properly. MCU requires 3.3 V power, CAN transceiver requires 5 V power, and 12 V is needed for the brake actuator. The maximum current consumption was measured to select the correct voltage converter. The maximum consumption of the board is 3.3 A and it is reached when the brake is released. The AP64501 was selected to convert 24 V to 12 V. This is a 5 A synchronous buck converter with the input voltage up to 40 V. Two AMS1117 are then used to convert 12 V to 3.3 V and 5 V.

The next step was to implement communication interfaces. The board has three types of communication: CAN, RS-485 and I2C. The MCP2551 was selected as a CAN transceiver. The circuit for this chip contains a 120 ohm termination resistor and chip for ESD protection.

MAX485ESA chip was selected for RS-485 communication. It was connected according to the datasheet. Terminal A has to be pulled up, terminal B – pulled down. The circuit includes same ESD protection chip as CAN.

The board contains 5 inputs for limit switches connection. The logic inputs of the microcontroller are isolated using optocouplers. They are synchronizing logic levels of the microcontroller (3.3 V) and the sensor (24 V). Each input also has ESD protection using TVS diode.

And the last circuit on the board is the brake system actuator control circuit. DC actuator was used for unlocking the braking system. It is powered with 12 V and consumes 3 A in active state. Tests showed that actuator doesn't have to be working full power to hold brake in unlocked state. Half power is enough to hold the brake, but full power impulse is still needed while unlocking. Described functionality is achieved using PWM controlled MOSFET transistor. It has 60 V Drain-Source Breakdown voltage and 20 A of continuous current. Transistor can pass 12 A with 3.3 V gate voltage, so it can be controlled directly from the microcontroller.

V. Main control software

The final step is the development of control software for the manipulator. The CODESYS platform will be used, which

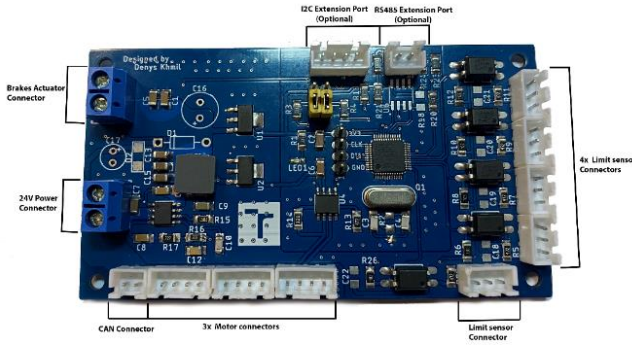


Figure 7 Designed PCB board

can be installed on the Beaglebone and turn it into a PLC. Since CODESYS is oriented towards working with industrial protocols such as CANopen, because of that there is no ready to use library for CAN communication that could be used with the custom G-code oriented protocol. For this purpose, a library will be developed using low-level functions, which will allow for the implementation of a custom communication protocol.

VI. Conclusion

During the work, the mechanical structure of the manipulator was assembled according to the previously developed concept. The electronic and software parts were upgraded by the closed-loop stepper driver which increase the flexibility and reliability of the machine. The available open-source firmware of the driver was modified to be operable, and some functions were added; the drivers now can be controlled using CAN or UART communication and achieve a great accuracy thanks to missed steps compensation. Also, a new Z-axis control board was designed and manufactured to meet the new requirements. Finally, all devices in the control system are connected via CAN bus.

The control system utilizes BeagleBone Black as the main control unit. The main role of the control program will be to manage all connected devices using the G-code oriented communication protocol. The implementation of the software for the main control unit is still ongoing. The software design utilizes CODESYS platform and its ecosystem providing PLC-style application.

References

[1] D. Khmil "The crane manipulator for the 3D printers" Brno, 2020 [Proceedings II of the 26th conference student EEICT 2020]., in press

[2] D. Khmil "The crane manipulator for the 3D printers" Brno, 2020 [Bachelor's thesis. VUT Brno] Available from: https://www.vut.cz/studenti/zav-prace/detail/127025

[3] Welling, Koos, "Automatic Printed Part Removal", Technical Disclosure Commons, (June 24, 2020) https://www.tdcommons.org/dpubs_series/3360

[4] Özsoy, K. , Aksoy, B. & Yücel, M. (2020). Design And Manufacture Of Continuous Automatic 3D Printing Device With Conveyor System By İmage Processing Technology . Erzincan University Journal of Science and Technology. DOI: 10.18185/erzifbed.6

Manufacturing control algorithms for Industry 4.0 demonstrator

1st Petr Klement
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
xkleme12@vut.cz

2nd Ing. Václav Kaczmarczyk Ph.D.
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
kaczmarczyk@vut.cz

***Abstract*—The article deals with the topic of production scheduling and its algorithms for distributed systems in Industry 4.0. In the theoretical part, we will first discuss what Industry 4.0 is, then cyber-physical and multi-agent systems and production planning algorithms in the industry. The practical part deals with the operation of the Industry 4.0 testbed called Barman, which is a school model of an autonomous beverage mixing line, and the implementation of the virtual production mode in the different parts of the cells located in Barman and its full operation and debugging of the hardware part in the production.**

***Index Terms*—Industry 4.0, Cyber-Physical Systems, Barman, TIA Portal**

I. Introduction

This article deals with the topic of distributed production planning algorithms. These algorithms are being used today with Industry 4.0 being the centerpiece of thriving companies. The algorithms deal with better coordination and alignment of the various tasks, processes, and material supply in production. In manufacturing, we can have two types of systems, namely multi-agent or cyber-physical systems, which we will learn about in this article. In the form of a cyber-physical system, a testbed is used industry 4.0 under the name of Barman. This testbed will be using the supplied codes and a virtual production mode will be added. This paper aims to introduce my thesis work with cyber-physical systems. Another goal is to debug the hardware shortcomings of the Industry 4.0 testbed Barman which will then be made operational through Siemens' TIA Portal software and at the same time, a virtual production mode will be added to it, which will be used, for example, for simplified functionality testing of the entire facility and for testing the MES supervisor system.

II. Introduction to Industry 4.0

Before an introduction to cyber-physical systems, we will look at the concept of Industry 4.0 (I4.0). Since cyber-physical systems are part of and based on the characteristics of I4.0. I4.0 is a vast field that cannot be described in a nutshell and explained thoroughly. Therefore below we will introduce the basic principles, and concepts of Industry 4.0 to better relate to cyber-physical systems.

A. General introduction

The concept of I4.0 was first introduced at the Hannover Messe in 2013. The designation 4.0 is meant to be a reference to the fourth industrial revolution, some experts however, consider I4.0 to be a natural evolution of the third industrial revolution. It could say that each industrial revolution has in some way removed the manufacturing process human factor is removed in some way. The basic visions of the so-called fourth industrial revolution emerged in 2011. The idea is that "smart factories" will emerge that use cyber-physical systems. These will take over repetitive and simple activities, that have been carried out by humans. [1]

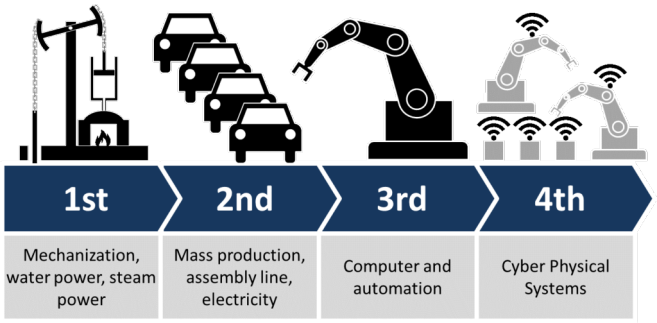


Fig. 1. Depiction of the various industrial revolutions. [8]

Industry 4.0 combines many new technologies and workflows into one. Creating the so-called smart factory. Nowadays, it's completely intelligent factories in the true sense of the word do not exist, but there are projects that are dedicated to it, e.g. Testbed for Industry 4.0 at CEITEC BUT. The idea of a completely smart factory is as follows. A fully autonomous factory that is able to produce a given task autonomously. From supplier to production to export. The main backbone of Industry 4.0 is the cyber-physical systems that implement parts of the production process. On the picture Fig. 1 we can see a historical development of industry.

III. Cyber-physical production systems

Cyber-physical systems (CPS) are a combination of mechanical, electronic, and software components that are capable of responding flexibly to dynamic changes, providing feedback

to control objects and exchange data with information systems and other CPS. They are therefore capable of autonomous behavior and, in future operations these systems will play a key role in optimizing the entire value chain. The basic idea from a system perspective is the integration of autonomous subsystems into complex distributed systems, in particular by enabling efficient communication of each integrated element with the other elements. This integration has led to the decentralization of management and the elimination of the need for a central hierarchical management, which means a transition from embedded systems to systems to cyber-physical systems that link the real world to the virtual world. Here, methods of auto-optimization, auto-configuration, and self-diagnostics will be applied, machine perception and intelligent worker support. However, in the first instance, this cyber-physical system must first be created. [2]

A. Division of cyber-physical systems

Multiagent systems - these are systems of independently behaving agents. These agents do not necessarily have to be physical, they can be just a software system. If the agents are physical, it is already a CPS system. A nice example would be automated conveyors. The advantages of multi-agent systems are great flexibility and fault tolerance. The disadvantages are computational and time-consuming, and the need for some standardization. [3] Holonic systems are multi-agent systems, where the agents form parts of the production. The agents are parts of a reactive control element and a software agent. Holonic agent is thus able to plan, coordinate and actively execute activities. In colonic systems, the agent can be a sensor, an actor, but also a whole machine, a production line, or company. [3]

B. Architecture of CPS

The 5C (Connection, Conversion, Cyber, Cognition, Configuration) architecture in the picture Fig.2 below defines the possibility of creating a CPS on several levels. These range from engaging sensors for data acquisition through data processing and digitization to autonomous optimization and adaptation to changes in external influences. [4] We can see in the figure Fig.2 below where we can see the whole pyramid of the 5C architecture.

IV. TestBed Barman

This part of the thesis aims to introduce the Barman project and its current. Status in terms of software and slightly in terms of physics. One of the practical tasks is to implement virtual mod. The barman can be seen in the picture Fig.3. Barman is divided into eight autonomous cells, two of which are transport and six of which are produced. The cells are located on the workbench in so-called slots. Each slot is equipped with two electrical connectors. The first connector is used to connect the power bus and the second for communication. The communication interface uses the profinet standard. The cells are standardized into two size formats, small cells (w: 330 mm × d: 330 mm × h: 500 mm) and large cells (w: 760 mm × h: 330 mm × h: 1500 mm). There are four slots on the table for cells. There are two slots behind the table for either large or small cells. A small cell is placed behind the table on an expansion bench, which is attached to the side of the rear side of the table. The location of the slots depends on the reach of the handling robot. [5] In my configuration, the bartender has four slots on the table for small cells, nothing more he does not use. [6]

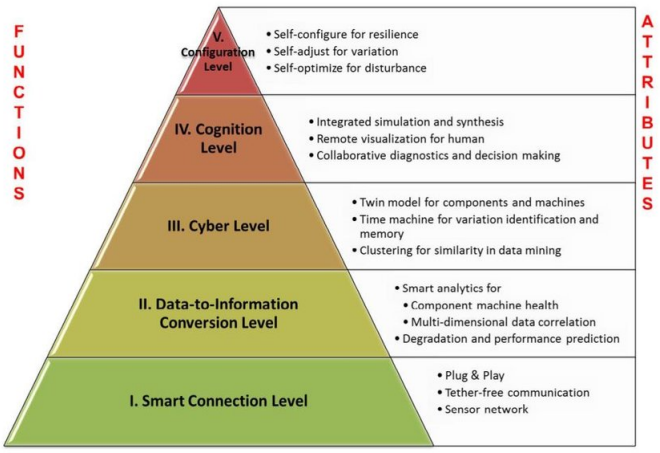


Fig. 2. Architecture 5C. [7]



Fig. 3. TestBed Barman.

A. Transport cells

The conveyor is the member who connects the machine to the public or the operator. His the main task is to transport the prepared drinks outside the robot's operating area. We can see in the picture Fig.4. Thanks to the conveyor it is possible to

ensure the partial safety of the Barman machine. The conveyor is equipped with ten glass trolleys and can detect the state and contents of the glasses. The robot ensures the transport of glasses between the cells and the conveyor. As a conveyor unit is a SCARA robot EPSON H554BN equipped with an effector capable of to grasp the jar. The robot is controlled by an EPSON SRC 310 controller, which uses the interface for the controller is RS232, through which the the robot is controlled by the parent PLC.



Fig. 4. Transport cells - Robot and conveyor. [6]

B. Glass storage cell

The glass storage cell is small in size. Its function is to dispense clean jars and store dirty glasses. For this purpose, a two-axis manipulator and four vertical trays are used, in which the jars are stacked. The manipulator is positioned to cover the two stores on the left side, the two stores on the right side, and the dispensing point located on the axis of the manipulator. [5] [6]

C. Sodobar cell

The cell stores drinking water and create saturated sodium from it using CO2. Soda water is produced in batches, where the size of the batch is based on the size of the company's satiating bottle. SodaStream. After creating one batch of soda, the contents are dispensed into the glasses again using a strain gauge bridge that weighs the glass. [5] [6]

D. Shaker cell

This cell has the task of shaking the contents of the jar. The shaker is powered by a DC motor and operates on the crankshaft principle, which converts rotary motion into linear motion. The cell does not change the contents of the jar, but it is necessary to simulate the movement of the shaker. [5] [6]

E. Soft-drinks storage cell

The cell consists of four soft drink tanks. These tanks are cooled by the table's refrigeration circuit. The cell is designed for dispensing larger quantities of liquids such as colas, tonics, or juices. The dispensing of liquids is mediated by solenoid valves and pulse flow meters. The glass is weighed by a strain gauge bridge which senses the amount dispensed. A cell not yet completed. [5] [6]

V. Barman commissioning

The next goal of the work is to put the testbed into working condition. This was preceded by familiarizing with the entire facility and reading the theses that were involved. Each cell is connected by two connectors. With this came the first problems, especially with the power connector. The bartender was enriched with a safety frame, which has opening windows on the sides, through which we can get in so that we can refill the cells with fluids or remove and replace them. When the plug is plugged into the cell, the windows do not close and the safety stops report an error and the testbed won't start. The solution was to move the cable outlet from the connector to the side where the cable will not interfere with closing the window and the safety contact (circuit) can be connected to the whole device could work. It was necessary to drill a hole in the side of the connector with a diameter of 18mm and then cut an M20x1.5 thread. This had to be done for all cells(4) placed on the testbed table. And mount the grommet on the newly created hole, which had to be unscrewed from its previous location. Into the previous hole, a 3D printer printed a plug to ensure safety, so that it's free of the live wires. After this modification, the entire bartender could be spun up. The next step was to check the program versions in each cell. Individual cells have their own PLC control system, which ensures the operation and performs the tasks of the master system. In checking the individual cells and testing their functionality, I came across the safety PLC that ensures the safety of the entire testbed. In the TIA environment Portal 15V1 environment, I had to install the safety packet to be able to communicate with this PLC and upload or download programs from it. This PLC controls the transport cells as the glass conveyor and the robot. This is where the differences between the program uploaded in the PLC and the supplied program from an earlier thesis that dealt with the safety of the Bartender. This is one safety feature that occurs the in the delivered program, but not in the PLC in the testbed. It does not occur in the other cells differences between the codes. The soft-drink warehouse was programmed in a different version of TIA Portal 17V and this cell was not even added to the overall program, which is a cluster of all cells. It is always possible to upgrade the program to a higher version in this program. For me, this meant installing Tia Portal 17V and converting the entire program. I then added the soft-drink cell to this program and added the PLC and HMI panel to the network configuration in the project. There was one more problem with the safety PLC, it could not communicate with the master from the capacitive sensors at the ifm glass detection conveyor. The fault was a misconfiguration in the Tia Portal where the inputs to the ifm master were not deactivated. After deactivating them everything works. In the picture Fig.5 we can see the correct setup.

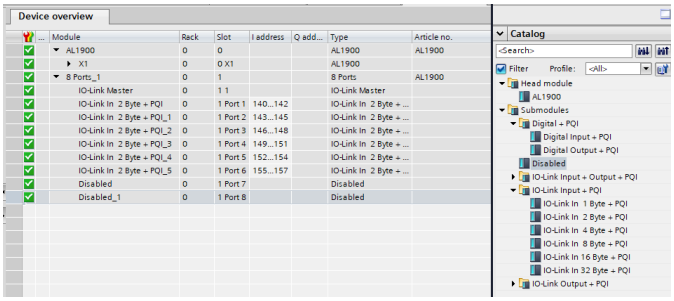


Fig. 5. Modification of ifm master AL1900 for capacitive sensors and communication with PLC.

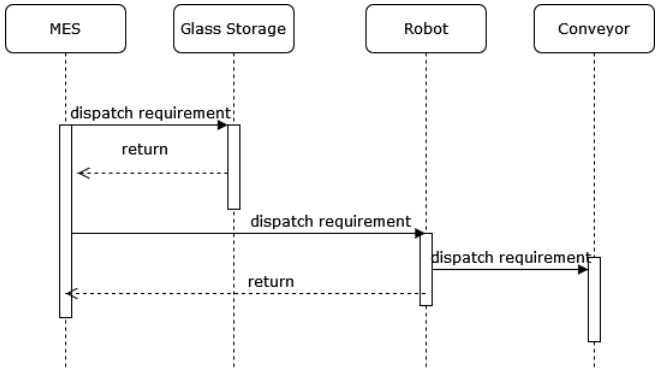


Fig. 6. Sequence diagram of the communication between the MES system and the PLC cells.

VI. Implementation of virtual production mode

Implementation of virtual production mode Individual cells contain 2 modes, namely automatic and manual. In automatic mode the cells receive the commands to be executed, they are sent by the parent system. In manual mode mode, we can control the cell completely by ourselves through the HMI touch screen which has on each cell. Here you can select manual/automatic mode, or you can control the individual actuators of the cell. Simulation mode in the cell, which will only pretend that cell is doing its job. This mode will be called virtual/simulation mode. In virtual mode, the cell will behave as if it were in automatic mode. The parent system sends specific commands to individual cells. A cell will be instructed by by the parent system, jumps to a subroutine where virtual production is executed by a timer that corresponds to the actual time, for example, the picking a glass from the warehouse. When the timer is complete, a message is sent to say that the task has been completed. And that it can proceed further. On the touch HMI screen, buttons are added to start the simulation and display this status. There will also be a display of the time it will take for the task to complete. For example, this mode could be used to test a higher-level MES system.

VII. Describing the production cycle of testbed barman

Entering the product into the production process is done through a web interface where we choose what we want to do. You will enter your name, and email and confirm you are over 18 years of age as they also serve alcoholic beverages. Next, the request is processed through the MES system, which commands the individual cells. For example, if we want an empty glass, we select it in the browser on the web page. The request is handled in the MES system and further, the MES assigns the task to the cell for dispensing the glasses. When the cell is not occupied it takes out the glass and sends the MES that it is done. Next, instruction comes from the MES for the robot to pick up the glass and hand it over to the vacant space on the conveyor. You can then pick it up. We can see this in the picture Fig.6.

VIII. Conclusions and Future work

In conclusion, I would like to say that I managed to run the testbed Barman, removing the mechanical flaws to get the barman running. Checked and consolidated the program of each cell into one project where all programs are together. Set up network communications and resolved minor programming deficiencies. I also added a simulation mode to each production cell that simulates production. These simulation modes are in the testing phase. The MES system needs to be fine-tuned for each item in the procedure for producing a given drink. I will continue to test the simulation and extend the controls on the HMI panel regarding the simulation mode. To make the whole device work both on the simulation side and on the physical side so that the bartender mixes the drink according to the order from the web interface, which we can open on any device where we have an internet connection. And properly document everything and write instructions for operation and commissioning.

References

[1] Průmysl I4.0. Průmysl I4.0 [online]. [cit. 2022-07-11]. Dostupné z: https:// cs.wikipedia.org/wiki/Pr

[2] Kyberneticko-fyzikální systém. Jakub Procházka. Industry 4.0 v oblasti výrobních strojů [online]. [cit. 2022-20-10]. Dostupné z: https://www.vut.cz/ studenti/zav-prace/detail/100622

[3] Kyberfyzikální-systémy. Kyberfyzikální systémy [online]. [cit. 2022- 07-11]. Dostupné z: https://www.iot-portal.cz/2016/08/22/ kyberfyzikalni-systemy/

[4] Digi. transformace. Digitální transformace průmyslu FEKT VUT v Brně. Zpracoval Jan Pásek

[5] KACZMARCZYK, Václav, Ondřej BAŠTÁN, Zdeněk BRADÁČ a Jakub ARM. An Industry 4.0 Testbed (Self-Acting Barman): Principles and Design. IFACPapersOnLine [online]. 2018, 51(6), 263-270 [cit. 2020-03-05]. DOI: 10.1016/j.ifacol.2018.07.164. ISSN 24058963. Dostupné z: https://linkinghub.elsevier.com/retrieve/pii/S2405896318309108 Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[6] Testbed průmyslu 4.0. HUSÁK, Michal. Virtuální dvojče pro testbed Průmyslu 4.0 [online]. Brno, 2020 [cit. 2023-01-01]. Dostupné z: http://hdl.handle. net/11012/189328. Diplomová práce. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav automatizace a měřicí techniky. Vedoucí práce Václav Kaczmarczyk.

[7] 5C architecture. CPS system [online]. [cit. 2022-07-11]. Dostupné z: https://www.iot-portal.cz/2016/08/22/kyberfyzikalni-systemy/

[8] Průmysl I4.0. Průmysl I4.0 [online]. [cit. 2023-03-31]. Dostupné z: https://www.renaix.com/industry-4-0-the-fourth-industrial-revolution/

SUPER HOT CHIPS



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

hr@asicentrum.cz

www.hotchips.cz



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

PŘIJĎ 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

em microelectronic

A COMPANY OF THE SWATCH GROUP

SCARA robot, program functions and use case

1st Martin Štípek
Deparment of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
211185@vutbr.cz

2nd Tomáš Benešl
Deparment of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
tomas.benesl@vut.cz

***Abstract*—This work deals with a new implementation of the control system for the older version of the SCARA robot. It shows a method for a tune-up of the motion controllers using the Ziegler&Nichols method and program introduction into functions for the new use case. This use case is modeled as a large 3D printer using an industrial robot.**

***Index Terms*—SCARA robot, Beckhoff, motion control, Ziegler&Nichols method, 3D printing**

I. Introduction

SCARA robot (Selective Compliance Assembly Robot Arm) is widely used kind of industrial robot. Its specializations are Pick&Place operations. In this article, this robot is configured into the specialized workspace for 3D printing applications. 3D printing was chosen to show a possible different approach to the use case normally implemented with the SCARA robot.

The original control system from EPSON company was destroyed due to an electrical malfunction. During the repairs, it were changed motors and control system and were integrated into the robot. New motors were chosen from the YASAKAWA company (specifically types SGM-02A, SGM-01A, SGM-01A and SGM-A5A) and controllers (types AX5203 and AX5201) and industrial PC (CX6015-0010 model) were purchased from Beckhoff. During previous works on the robotic system, the system was set in motion but it had many problems, like too many errors causing malfunction to motion or not fast enough moving speed.

This article will be describing what was done on the control system, how were tuned the regulators AX500x, how looks visualization of possible future demonstrative robotic workplace and its use case.

II. Motion control of SCARA robot

Every mechanical system needs some kind of selected drive. In motion control of industrial robots, it is best practice to use servo motors, which can provide feedback for the closed control loop.

A. Setting motion to robotic arm

It is possible to control an industrial robot in two modes, automatic or manual. Manual control is configured only axis for servo motors in each axis of SCARA robot. It needs to set references for each axis in the control system and configure power-up sequence (preferably using a function for repeated usage). After that, it is possible to control each axis in small steps, which is called JOG.

In automatic mode, probably using some coordinates in the workspace, it is required to define another virtual axis for the cartesian system (axis X, Y, Z and rotating around axis Z). After starting all eight axis of system (4 axis based on the number of servo motors integrated into a robotic arm and 4 axis defined by the cartesian system) it is possible to use a translator of coordinates. This could be a linear or circular interpolation of coordinates obtained from G-code, for example.

B. Improvments of motion control

The original version of implemented system had problems with electromagnetic interference causing loss of communication and control loop of servo drivers. Despite that the electrical installation had very good shielding of wires, it lack shielded EtherCAT communication. Changing patch cables from Ethernet cat. 5 to Profinet cables (which are used in shielded cat. 6 communications) solved many problems with system errors, especially with invalid communication responses or data N-times in a row. Another problem was causing multiple failures of the control system due to one failure. This was solved by defining multiple error groups. After this, one failure. For example, on driver controller number 2 detects an error code, then it disconnects only the second controller.

Then it needed to measure motor response to control signals from servo drivers. It used an internal measurement project in Scope mode, which simulates an oscilloscope device. By this project, it was measured relations between Actual and Requested position [°], Actual and Requested velocity [°/s], Actual and Requested acceleration [$°/s^2$], and Actual and Requested torque [Nm/A]. It was found that the designed control loop was not suitable for this application so it needed to be modified. Before implementing new constants for controllers, it needed to be changed constants of the servo motors according to their datasheets for better responses on current impulses from drivers. According to the datasheet for setting constants for regulators [1] it needed to be disabled *Position Laq monitoring*. This is due to experimental constants settings based on Torque response and defined cyclic oscillation around the default position.

After this preparation work, it was possible to experimentally test gain of velocity regulators to determine critical gain K_{P_krit}. This critical gain was reached by experimentally increasing gain of the speed regulator up to momentum limit of the motor when it started oscillating around its standstill value (sharp peaks at the edges of the motion requests). Those oscillations were measured using scope project using velo step sequence for rotating selected axis by predefined speed constant during predefined time period of travel. The measured momentum of critical gain is shown in 1 using green color at the end of bode graphs. It was aimed to determine system

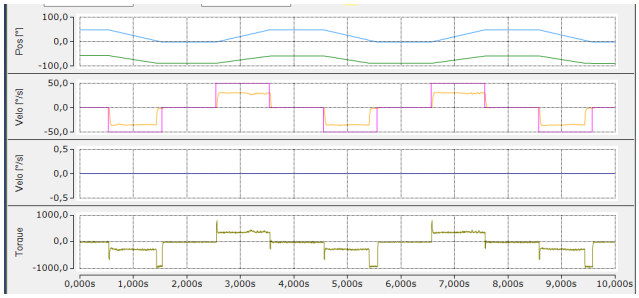


Fig. 1. Critical gain of motor

stability by increasing this value. When it was on current limit (defined by the used motor and regulator) it was determined by torque oscillating around requested position during no movement. This value can be approximately determined by motor buzzing. After finding this value it is possible to find correct gain constants of velocity regulators by applying the formula of the Ziegler-Nichols method (1):

$$K_P = 0.45 \times K_{P_krit} \tag{1}$$

After that, it was taken default value of the integration time constant of 8ms and set new values for position regulators by equation (2):

$$K_v = 1 \times \frac{1000}{min} \tag{2}$$

After this, it was set back *Position Laq monitoring*. With these new regulator constants, it was possible to get better feedback from motors and faster response on current from drivers. It was also possible to increase speeds of axis implemented in the joints of the industrial SCARA robot up to predefined construction limits and reach commutation limits of defined controllers. It was greatly modified gain values (as it is shown in tables II and I). At first look, it is not that big difference, but it provided better motion responses to different current impulses in controlling robotic arm. One of the bode characteristics with newly set parameters is shown in graph 2

C. Future improvements of control system

Furthermore, there will be developed advanced control system using coordinates in the cartesian system. These coordinates will be interpolated by linear or circular trajectories automatically according to preselected curve. It will be possible to follow predefined sequence of coordinates, acquire

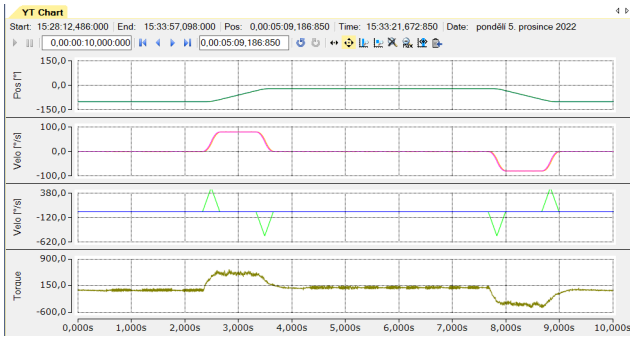


Fig. 2. Bode characteristics of improved motor controll

TABLE I
ORIGINAL SET VALUES OF MOTION REGULATORS

	Original Values			
Axis No.	**1**	**2**	**3**	**4**
K_v [1000/min]	1.6	1	1	1
K_P [Nm/(rad/s)]	0.07	0.02	0.04	0.06
T_n [ms]	8	8	8	8

TABLE II
NEWLY SET VALUES OF MOTION REGULATORS

	New Values			
Axis No.	**1**	**2**	**3**	**4**
K_v [1000/min]	1	1	1	1
K_P [Nm/(rad/s)]	0.05	0.052	0.018	0.04
T_n [ms]	8	8	8	8

them from imported G-code or they could be forced through interface between user and the system.

Simultaneously with that, it will be developed advanced interface of HMI running on a connected PC and a second interface, that will be accessible through a web browser after entering correct IP address on the local network. This web interface could work as remote access to the control system in the future.

III. USE-CASE OF SCARA ROBOT

Every industrial robot needs a tool at the end of its kinematic system. If it is not mounted at the end of the robotics arm, therefor there could not be any use case. In this case, there were not any mounted end effector so this model concept proposes a possible solution for the robot use case.

A. Tool for robotic arm

Most common applications for robotic effectors are: [2]

- Material handling - Gripper - Those are great for moving objects around robotics workspace. Effector grips the object and while holding it, it moves it to a new defined place.
- Move from one place to another on one surface - Pick&Place applications are used for example for transferring items from one conveyor to another at very high speeds - electromagnetic effectors or suction pumps

- Assembly - during this operation parts of final product assembling at very high speed with great accuracy - electromagnetic effectors
- Welding - this uses special tools which must withstand extreme environments (high voltage, dustiness, . . .)

Those are just some of the many possible applications with end effectors. In this article, let's take a look at something different.

Because not many companies do not do large 3D printing, it was decided that the end effector should be used (virtually for its cost) effector for 3D printing. The effector consists of an extruder head, cylindrical body, motor, screw conveyor, and hopper.

- Hopper - Through this conical-shaped tool it is possible to fill in additional material for 3D printing
- Cylindrical body - there is stored almost whole material intended for later use. Through the axis of the cylindrical body passes a screw conveyor. Its purpose is to continuously deliver material into the extruder head for further processing.
- Motor - for purpose of moving the conveyor screw was selected small stepper motor. Its shaft is connected through a conjunction to the conveyor. Around the motor is modeled cage for mounting the effector at the bottom of the third axis of the SCARA robot. Mounting the motor into the cage should stop the momentum forcing body of the motor to rotate and transfer all rotation energy into the conveyor
- Extruder head - It melts down supplied material. It uses electrical energy transferred into heat. The amount of heat required is regulated through the amount of delivered energy and a cooling fan mounted on the effector. The fan prevents overheating the electronics and helps regulate the temperature inside of extruder head

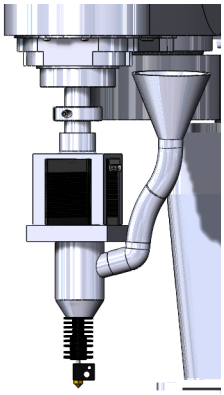


Fig. 3. 3D model of end effector

With this modeled end effector (ref fig. 3) it is possible to print almost everything with most of the common materials. It is possible to use granulated filament in ranges from ABS or PETG up to theoretically aluminum materials. These materials could be purchased from stores already crushed and compressed into pellets. Another way of getting material is crushing scrap from faulty printing or from some plastic waste (PET bottles). This leads to recycling materials at the place of origin. That way it is more economically and environmentally friendly because it saves some of the transportation and at the place of crushing could be sorting bins where the material could be already presorted (etc. separate bins for aluminum cans and PET bottles).

While using input material aluminum alloys, their melting point is starting around 400-500°C. In this case, the whole application must have significant improvements to withstand these extreme temperatures. It would require heat shields to deflect heat from the robot and its electronics, body of the effector made of steel alloys (the melting point of steel alloys is around 1300°C), and a powerful enough heating coil placed inside of the extruder head [5].

B. Workspace and operator safety

The workspace of the used industrial SCARA robot is simulated in figure 4. It consists of a work desk, a similar model to the used robot (specifically newer model EPSON T6 which have 3D model available online [3]), safety pieces of equipment and an automatic feeder already automatically imported in SolidWorks Studio. On the workdesk is mounted robot and feeder. When the robot end effector is almost empty, the robotic arm will move into the default position. There is modeled feeder and automatic refill station for the hopper with actual present material. After that, the robot moves back to the last known coordinate and continues to work. When a 3D print is done, the heated conveyor under the robot moves product automatically to the edge of the workspace, where is removed by operator. Around the workspace were constructed

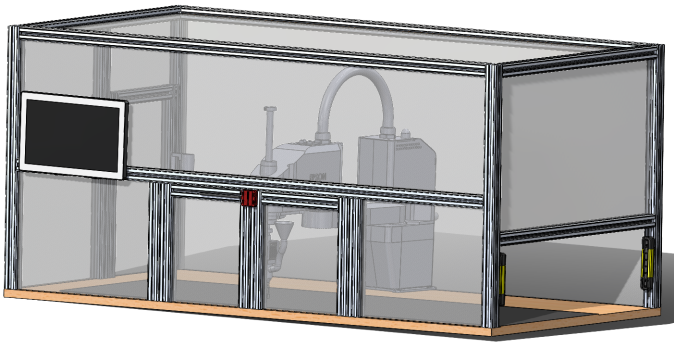


Fig. 4. Modeled workspace

aluminum frames, filled with plexiglass. The robot itself could stop when the force is applied on some link, but this force could be potentially dangerous for user to harm him. This frame prevents operator to be present in the area of the robot when it is in motion. To service devices (add filament, service conveyor, or robot) it is provided with doors and holes in construction protected by safety elements. When the robot is moving, the safety locks will block the doors. Inside each lock is also a sensor for detecting opened door and if it is not closed, the motion of SCARA robot is deactivated. The

output hole is protected by safety light curtain. When there is some object present (like an arm), the robot motion controller is also deactivated. This construction was modeled by using EN ISO 10218-1 and EN ISO 10218-2 standards for robotic environment safety and provides overall safety performance level d of category 3. [4]

IV. Future improvments

At the time of publishing this article, there are many improvements in progress on the SCARA robot. For example, it is still getting better motion control by slightly improving control loop and regulators parameters. It is also getting advanced functions on kinematics control and interpolation interpreter of G-code. It is also in progress teaching robotic arm trajectories and paths by manually pushing the axis. The visualization is being developed for a web browser for remote access to robot, currently possible only from the local network.

V. Conclusion

This article summarizes key points of tunning the used regulators from Beckhoff used in the SCARA robot using the Ziegler-Nichols method. It was developed some program functions, some were greatly improved and some will be added. For example, it will be added selective trajectory planning of the endpoint of the robot passing through coordinates from G-code file.
Next, it was modeled as a possible future workspace integrating large 3D printing applications and functional safety was applied to this model using corresponding standards.
Furthermore, this article outlines future work on the used robotic system and further improvements and futures on the system before the thesis defense.

Acknowledgment

The completion of this paper was made possible by the grant No. FEKT-S-23-8451 - ”Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence, automation and measurement” financially supported by the Internal science fund of Brno University of Technology

References

[1] Tuning Velocity Controller (Ziegler&Nichols). *Beckhoff Information System* [online]. Beckhoff Automation GmbH & Co. KG [cite 2023-03-12]. Available from: <https://infosys.beckhoff.com/english.php?content=../content/1033/ax5000_tuning_guide/3423431051.html&id=>

[2] Robot Grippers and End Effectors: Uses, Benefits, and Cost Analysis. *HowToRobot* [online]. Dirch Passers Allé 27, Denmark: HowToRobot [cite 2023-03-12]. Available from: <https://howtorobot.com/expert-insight/robot-end-effectors>

[3] MH, Oktay. *3D model EPSON T6-B602s* [online]. In: . Azerbajdžán, 2022 [cit. 2022-10-05]. Dostupné z: ¡https://grabcad.com/library/epson-t6-b602s-1¿

[4] ISO 13849-1:2015 *Safety of machinery — Safety-related parts of control systems: Part 1: General principles for design*. European Committee for Standardization, EU, 2015.

[5] What is the Melting Point of Aluminum, and Why is it Significant? *Gabrian International website*. Gabrian International (H.K.) Ltd. [cite 2023-03-29]. Available from: <https://www.gabrian.com/melting-point-of-aluminum/>

Precision landing of multirotor drones using computer vision

Van Phi Nguyen
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
246484@vutbr.cz

Josef Bajer
Department of Aircraft Technology
University of Defence
Brno, Czech Republic
josef.bajer@unob.cz

Václav Křivánek
Department of Military Robotics
University of Defence
Brno, Czech Republic
vaclav.krivanek@unob.cz

***Abstract*— Precision landing is the last and most crucial stage in navigation, and autonomous precision landing Unmanned Aerial Vehicles (UAVs) have many potential applications across a variety of industries. Until recently, most designs relied on the Global Positioning System (GPS) and Inertial Navigation Systems (INS) with sensors such as 3-axis accelerometers, gyroscopes, and magnetometers. However, the accuracy of GPS data is often insufficient, necessitating the use of more precise sensors. In this study, a high-precision landing solution that employs a camera and ArUco markers as the landing platform is proposed. Once the markers are detected through image processing, the relative position between the markers and the drone is calculated. Using the Micro Air Vehicle Communication (MAVLink) protocol, the NVIDIA Jetson Nano board communicates with the flight controller to manage the drone's position, enabling it to fly to the marker's location and land precisely. The algorithm's efficacy was tested and confirmed in both the Robot Operating System (ROS)/Gazebo simulation environment and real-world experiments.**

Keywords—ArUco markers, drone, MAVLink protocol, NVIDIA Jetson Nano, OpenCV, precision landing, ROS/Gazebo, UAVs

I. Introduction

The interest in UAVs has grown rapidly in recent years due to their wide range of applications, making them one of the hottest technologies today.

Landing a UAV is acknowledged as the last and most critical stage of navigation [1] especially when 80 % of the hazard UAVs' accidents are related to the landing process [2]. Thus, accurately estimating the position and orientation of UAVs is crucial for safe self-landing. The traditional methods of using GPS and inertial sensors have shown significant inaccuracies, resulting in a deviation of up to 10m from the intended landing position.

Recently, different types of UAV landing approaches have been studied. In particular, a number of computer vision-based approaches have presented. Oualid Araar and his colleagues [3] succeeded in autonomous tracking and landing on a moving platform. This indoor experiment showed that the UAV can successfully land on the target landing platform (that contains multiple fiducial markers). Chen et al. [4] developed a vision approach and the use of LiDAR sensors for tracking and localization systems. The landing target platform is a sample with a red rectangle and a green circle inside. The drone can track the landing target's position (in horizontal plane *x*,*y*- coordinates) and the altitude parameter is provided by the LiDAR sensor. J. Wubben and his colleagues [5] succeeded in landing a drone using the same ArUco markers landing target. However, their algorithm is built on the ArduSim library, which is a novel real-time flight simulator, oriented to the development of flight coordination protocols for multicopters, performing planned missions or forming a swarm.

This work aims to design, implement, and verify a hardware and software system solution to enable a drone to autonomously land at a specific location. The primary goal is to develop a vision-based landing system that uses the drone's built-in image processing system to detect the landing target. One difference in this work is that the UAV can detect the landing area when it is high above the ground (height > 8 m). The proposed landing algorithm employs the ArduPilot firmware while using inexpensive sensors (only a Jetson camera is required) and achieving very low landing accuracy (less than 20 cm).

II. System Architecture

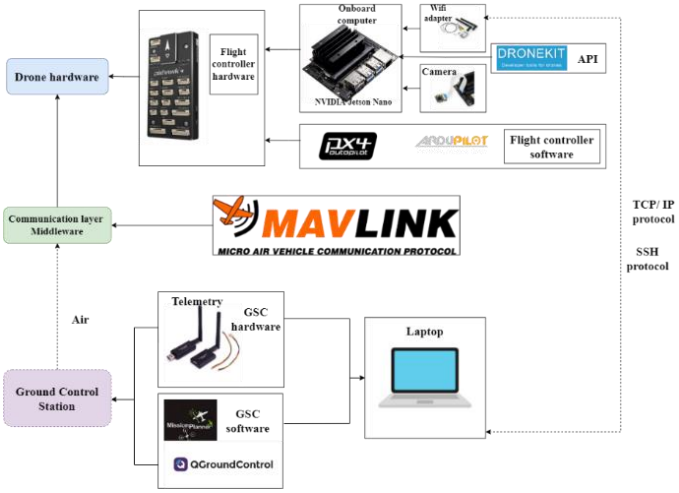


Fig. 1. Overview of hardware and software structure for autonomous landing.

A. Main hardware

Fig. 1 illustrates the system architecture, which includes essential parts of the UAV onboard system:

- The flight controller – Pixhawk 4 which receives the information from different sensors such as GPS, barometer, and LiDAR, etc.
- A wide-angle fish-eye camera connects to the NVIDIA Jetson Nano by a Camera Serial Interface (CSI). The camera

needs to be calibrated to avoid distortion and improve the performance of the precision landing algorithm.

- The onboard computer NVIDIA is in charge of image processing, sending controlling signal to the flight controller and communicating with the ground control station.

- And other general components as Electronic Speed Controller (ESC), brushless DC motors, Li-Po battery, telemetry, etc.

Two types of connections exist between the drone and the ground control station: telemetry-based connections using the MAVLink protocol and router-based networks using the TCP/IP protocol.

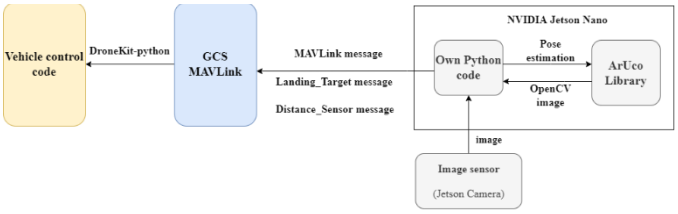


Fig. 2. Software architecture.

B. *Landing platform – ArUco markers*

ArUco markers are square fiducial markers that can be used for camera pose estimation and have great potential for identification and localization in 3D space [6]. They are similar to QR-codes but contain only their identifiers. ArUco markers come in various configurations of $N \times N$ (N = 4; 5; 6; 7), but they all have a black border that enables fast detection in the image and identifies the code within it. For this project, landing pads use markers from the ArUco marker library's DICT_ARUCO_ORIGINAL, which has size 7 x 7 markers. The border consists of black columns and rows, with data encoded in a 5 x 5 matrix inside (see Fig. 3). A white square corresponds to a value of 1, and a black square corresponds to a value of 0. The first, third, and fifth columns serve as parity, while the second and fourth columns store the marker identifier in natural binary code.

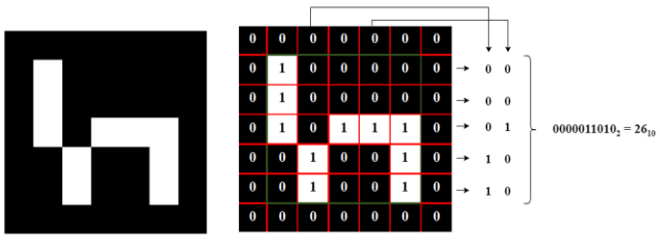


Fig. 3. Example of ArUco markers and decoding it.

Open-Source Computer Vision (OpenCV) is employed to process the video input from the camera to detect the ArUco marker. Additionally, it visualizes the parameters and bounding box of the detected markers, estimates their poses and attitudes relative to the camera frame. After image processing, the final output obtained is either the position of the ArUco marker within the image frame (x_{pixel}, y_{pixel}) in pixels or the marker's relative position $P(x, y, z)$ in meters with respect to the camera.

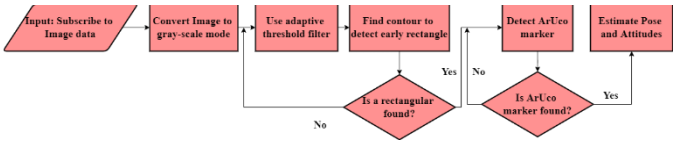


Fig. 4. The main image-process algorithm by OpenCV.

III. MAVLink Protocol for Landing Procedure

The MAVLink protocol fully supports communication with the UAV. This protocol is a very lightweight messaging protocol for communication between vehicles (particularly drones) and ground stations [7], and also between the drone's flight controller and onboard computer NVIDIA Jetson Nano.

The main MAVLink message utilized in this study is Landing_Target. Two types of parameters can be sent in this message via MAVLink version 2 to direct the drone towards the landing platform. The first parameter is the relative position $P(x, y, z)$ of the marker, as determined by the image processing output. The second is the set of parameters ($angle_x$, $angle_y$, z), representing x, y-axis angular offset off the target from the center of the image, and the parameter z in $P(x, y, z)$ representing the distance from the drone to the marker.

$$angle_x = \frac{\left(x_{pixel} - \frac{horizontal_{resolution}}{2}\right) \cdot horizontal_{fov}}{horizontal_{resolution}} \quad (1)$$

$$angle_y = \frac{\left(y_{pixel} - \frac{vertical_{resolution}}{2}\right) \cdot vertical_{fov}}{vertical_{resolution}} \quad (2)$$

- *horizontal_fov*, *vertical_fov* [rad] – the horizontal, vertical camera Field of view (FOV)angles.
- *horizontal_resolution*, *vertical_resolution* [pixel] – the horizontal, vertical resolution of the camera.
- x_{pixel}, y_{pixel} [pixel] – the position of the marker center in the image frame, on the x-axis, y-axis.

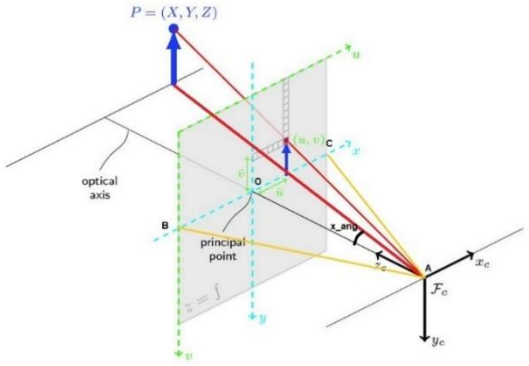


Fig. 5. The target's relative position on the captured image, adapted from [8].

Assuming the drone is equipped with a rangefinder, the Landing_Target message operates correctly. However, if a rangefinder is not available, there are two solutions: install a suitable rangefinder or create a virtual "MAVLink rangefinder". The proposed algorithm addresses this issue and utilizes the depth parameter of the ArUco marker by implementing a virtual distance sensor through a different message format of the MAVLink protocol, known as Distance_Sensor.

The necessary broadcast rate for transmitting MAVLink messages varies based on the landing speed and desired accuracy. MAVLink documentation indicates that Landing_Target and Distance_Sensor messages necessitate broadcast rates of 10-50 Hz and 5-20 Hz, respectively. After conducting research and flight tests, a suitable camera resolution for this algorithm is 640×480 at 30 FPS. This resolution satisfies two requirements: detecting a 30×30 cm landing pad at a height of 15 m and transmitting MAVLink messages at the required broadcast rates. It is important to note that modifying the image resolution necessitates camera recalibration.

IV. Landing Procedure

There are two conditions to detect a marker: (1) the marker must be completely inside the picture and (2) each square in this marker must be uniquely identified and analyzed (black or white).

Precision landing algorithm using MAVLink protocol
start timer 30 s
alt ← drone.location.global_relative_frame.altitude
while *alt* > z_3 **do**
if *alt* > z_2 **then**
id_to_find ← *id_1; marker_size* ← *id_1_size; marker_height* ← z_1
elseif *alt* < z_2 **then**
id_to_find ← *id_2; marker_size* ← *id_2_size; marker_height* ← z_2
end if
detect marker *id_to_find*
if not detected **then**
Change to *Loiter_mode*
if timer exceeded then
abort landing by GPS
end if
else
Change to *Land_mode*
get P(*x*, *y*, *z*)$_{id\ to\ find}$
get(*angle_x, angle_y*)
Send MAVLink *Distance_Sensor* message (integer (z_m))
Send MAVLink *Landing_Target* message (*angle_x, angle_y*)
end if
end while
descend and disarm UAV

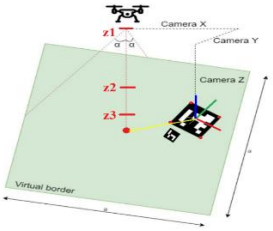


Fig. 6. The illustration of the drone landing procedure.

During drone landing, its altitude plays a crucial role. If the drone is too close to the landing pads (i.e., $z_3 = 0.5$ m), the marker will exceed the camera's field of view, making it too big to be fully captured in the image frame. Conversely, if the drone is flying too high (e.g., 15 m) and the marker size is only 30×30 cm, the camera will fail to detect the target. However, increasing the marker's size on the ground can enhance the camera's ability to detect it, thereby improving the maximum allowable altitude at which the camera can spot the marker. As long as the camera can detect and identify the marker, the precision landing mode will be activated.

However, using a large landing pad to increase the detection altitude can result in a larger error, making landing on the center of the marker challenging. To address this, multiple markers of varying sizes are combined to improve detection. Specifically, this work employs two ArUco markers, *id_1* and *id_2*, of different sizes. The camera will search for markers *id_1* or *id_2* at altitudes below z_1 (8 m in real-world experiments and 30 m in the simulation environment) or below z_2 (2 m in real-world experiments and 6 m in the simulation environment) respectively, depending on the drone's current altitude. If no marker is detected, the drone will switch to mode Loiter to maintain its position, heading, and altitude, and if the searching exceeds 30 seconds, the mission is aborted, and the drone will land using GPS. When a marker is detected, the relative position *P(x, y, z)* of the marker with respect to the camera is obtained by the NVIDIA Jetson Nano's pose estimation of the ArUco markers. The drone then switches to mode Land, and the onboard computer calculates the parameters *angle_x, angle_y*, sending two MAVLink messages to the flight controller. The drone will continue to descend and disarm until its altitude is below z_3.

V. Results

The landing platforms are strategically placed in different locations. The markers are positioned 4-5 m away from the drone to ensure they are always within the landing area where the UAV is searching for markers. A successful landing is defined as the UAV landing within 20 cm of the intended marker's center. All test flights were performed in the simulation to verify the system's functionality, and the results were 100% successful in the ROS/Gazebo simulation. The position error is calculated by comparing the UAV's position to the marker's center.

The average deviation of the UAV from the marker center was 4.97 ± 2.98 cm on the *x*-axis and 5.14 ± 3.26 cm on the *y*-axis in the simulation environment (from the height of 30 m). In real-world flight tests, the UAV had an error of only 3.30 ± 1.46 cm on the *x*-axis and 2.33 ± 1.36 cm on the *y*-axis (from the height of 8 m).

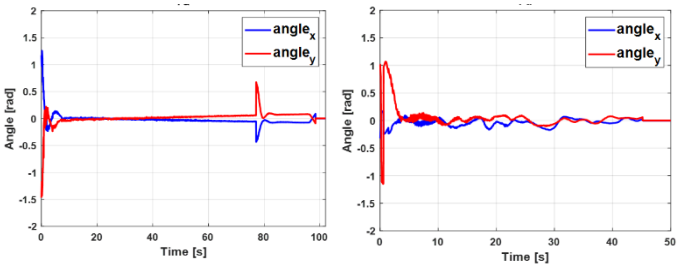


Fig. 7. angle_x, angle_y angular offset of the target from the center of the image in the simulation ROS/Gazebo (left) vs. in the real-world flight tests (right)

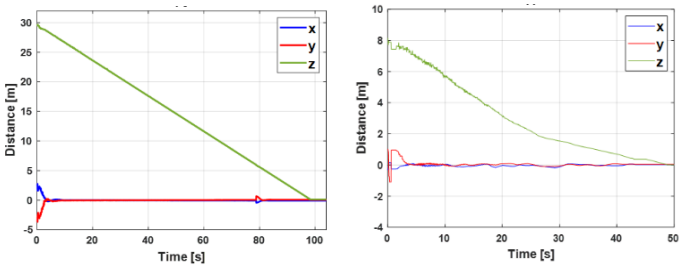


Fig. 8. Velocity of the drone in the horizontal plane in the simulation ROS/Gazebo (left) vs. in the real-world flight tests (right)

During the first 10 s of the simulation environment test, the drone attempted to move towards the center of the landing pad while gradually descending. Due to inertia, *angle_x* and *angle_y* oscillated around zero. The drone changed its course to land on a smaller marker at a height of 6 m in the 78th-79th s (Fig.7). After a few seconds, the drone achieved equilibrium on the smaller marker, and *angle_x* and *angle_y* were recomputed.

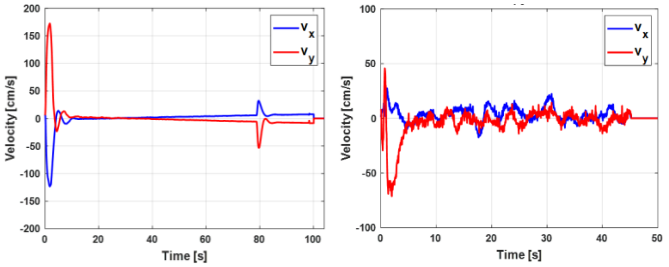


Fig. 9. Position of the drone with respect to the markers on the ground in the simulation ROS/Gazebo (left) vs. in the real-world flight tests (right)

In real-world flight tests, the inertia had a more significant effect, and *angle_x* and *angle_y* oscillated around zero. During the precision landing process, the Pixhawk determined the appropriate velocity for the drone in the horizontal plane based on the target's angular offset from the image center and the relative position of the markers and the drone. When the deviation angle was not too wide, and the distance between the drone and the markers on the *x* and *y* axes was not too great, the drone barely moved. However, in real-world flight tests, due to inertia, the UAV overshot its target and began braking behind it, then flew in the opposite direction, overshooting again, and repeating this process until the target's angular offset from the image center was small enough. Fig. 9 depicted the UAV's smooth landing trajectory towards the landing pads.

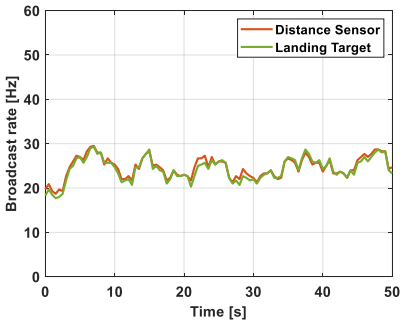


Fig. 10. The broadcast rate of two MAVLink messages in real-world flight tests.

The NVIDIA Jetson Nano sent the MAVLink messages Landing_Target and Distance_Sensor to the Pixhawk 4. The broadcast rate is exactly what was expected and desired.

TABLE I. COMPARISON OF RESULTS

Source	Accuracy [m]	Max. altitude [m]	Landing speed [m/s]	Out-doors	Mov-ing target
Our algorithm	0.05 – 0.11	8 – 50*	0.2 – 0.3	Yes	No
GPS – based	1 – 3	∞	0.3 – 2.0	Yes	No
ArduPilot PL with IR-Lock sensor [9]	< 0.3	15	n/a	Yes	Yes**
Araar et al. [3]	0.13	0.8	0.06	No	Yes
J. Wubben et al. [5]	0.11	20 – 30	0.1 – 0.3	Yes	No
Chen et al. [4]	n/a	2.5	0.23	No	Yes

** Depending on the size of the ArUco markers and the resolution of the camera.*
*** Apply for both the moving target and the stationary one.*

Our approach applies to the common Vertical Take-Off and Landing (VTOL) UAVs, whose flight controllers support ArduPilot firmware and MAVLink protocol for landing procedure with an inexpensive camera required only. Under comparable circumstances, ArduPilot community trials on landing by GPS or with an IR-Lock sensor [9] were also carried out. Experiments by Araar et al. [3] performed with the AR.Drone 2.0 quadrotor from Parrot. Chen et al. [4] worked with a T-Lion board, which installs the flight controller, powerful processors and several essential sensors together. These two proposals are carried out in-house without considering environmental factors. J. Wubben et al. [5] utilized their own ArduSim library to implement the algorithm, building on the similar idea about ArUco markers.

VI. CONCLUSIONS AND FUTURE WORK

The ROS/Gazebo simulator was utilized to evaluate the precision landing approach. The method's accuracy is remarkable, potentially achieving less than 20 cm. The ArduPilot firmware allows for the adjustment of landing time via the vertical landing speed parameter. This landing approach employs a single camera sensor that serves as both a target detector and a rangefinder, providing depth parameters from a ground marker. Consequently, it is a cost-effective solution that eliminates the necessity for a depth sensor, such as LiDAR or Sonar. While the algorithms in this landing approach meet the requirements for precision landing, they are not the most optimal. Future work aims to improve the algorithm by integrating the UAV's state, including acceleration and velocity, into the control loop to enhance flight behavior. Additionally, the plan is to use cascaded ArUco markers instead of one or two ArUco markers as the landing platforms, which will improve both landing accuracy and speed. Lastly, the algorithm should be developed further to accommodate scenarios involving night vision and less favorable weather conditions.

REFERENCES

[1] X. Pan, D.Q. Ma, L.L. Jin, Z.S. Jiang. *Vision-based approach angle and height estimation for UAV landing*. In Proceedings of the 1st International Congress on Image and Signal Processing, CISP 2008, Sanya, China, 27–30 May 2008; Volume 3, pp. 801–805.

[2] D. Tang, F. Li, N. Shen and S. Guo, *"UAV attitude and position estimation for vision-based landing,"* Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology, 2011, pp. 4446-4450, doi: 10.1109/EMEIT.2011.6023131.

[3] O. Araar, N. Aouf, & I. Vitanov. *Vision Based Autonomous Landing of Multirotor UAV on Moving Platform*. J Intell Robot Syst 85, 369–384 (2017). https://doi.org/10.1007/s10846-016-0399-z

[4] X. Chen, S. K. Phang, M. Shan and B. M. Chen, "*System integration of a vision-guided UAV for autonomous landing on moving platform*," 2016 12th IEEE International Conference on Control and Automation (ICCA), 2016, pp. 761-766, doi: 10.1109/ICCA.2016.7505370

[5] J. Wubben, F. Fabra, C.T. Calafate, T. Krzeszowski, J.M. Marquez-Barja, J.-C. Cano, P. Manzoni, "*Accurate Landing of Unmanned Aerial Vehicles Using Ground Pattern Recognition"*. Electronics 2019, 8, 1532. https://doi.org/10.3390/electronics8121532

[6] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

[7] Dronecode Project. *Introduction - MAVLink Developer Guide*. MAVLink Developer Guide [online]. [cit. 2023]. Available online: https://mavlink.io/en/ (accessed on 10.03.2023)

[8] Dronecode Project. *Landing Target Protocol*. MAVLink Developer Guide [online]. [cit. 2023]. Available online: https://mavlink.io/en/services/landing_target.html (accessed on 10.03.2023)

[9] ArduPilot Dev Team. *Precision Landing and Loiter with IR-LOCK*. ArduPilot - Versatile, Trusted, Open [online]. 27.02.2023 [cit. 2023]. Available online: https://ardupilot.org/copter/docs/precision-landing-with-irlock.html (accessed on 27.02.2023)

Multiagent systems for distributed production

Martin Žiačik
FEEC - Department of Cybernetics, Control and Measurements
Brno University of Technology
Brno, Czech Republic
xziaci00@vut.cz

Václav Kaczmarczyk
FEEC - Department of Cybernetics, Control and Measurements
Brno University of Technology
Brno, Czech Republic
kaczmarczyk@vut.cz

***Abstract*— Manufacturing companies today face a dynamic environment in which economic, technological, and customer trends are changing rapidly. It is requiring increased demands for flexibility and agility in production management and planning. To handle the growing complexity of systems, we need new methods where decentralization is playing an important role. In this context, the challenge is to develop an innovative, agile, and reconfigurable architecture for distributed production management using new paradigms in response to market demands. The use of multi-agent systems that reflect these demands in their design can be a suitable alternative to traditional centralized production management. In this document, we present a theoretical knowledge base of intelligent manufacturing systems and briefly introduce the implementation of such a production management model.**

***Keywords*— scheduling, planning, C# algorithm, multi-agent approach, optimization, simulation, Industry 4.0**

I. Introduction

The field of manufacturing systems is becoming more and more intricate, with a plethora of unexpected and unforeseen events constantly emerging. These events may include rush orders, resource malfunctions, late change requests, and order cancellations/modifications. The conventional approach for resolving these issues relies on a centralized structure where decision-making functions are concentrated in a single node. Such a centralized point is Manufacturing Execution System (MES). This system serves among others for job planning and management, and all relevant data must be transferred to them. As a result, these centralized approaches are unable to adapt to the fluid and dynamic nature of modern industrial systems where unexpected events occur frequently. [1] To overcome the challenges of complexity and interoperability in the industry, it is crucial to implement new production system architectures. These architectures would serve as a fundamental driver for reducing complexity and promoting interoperability between various manufacturing systems. Also, thanks to the initiative of Industry 4.0 traditional production systems are being transformed to create modern and intelligent factories. This transformation is made possible through the integration of Multi-Agent System (MAS) architectures and agent-based paradigms which comprise a group of agents that are independent and autonomous. Concrete implementation of the MAS in the field of manufacturing production is called Cyber-Physical Production Systems (CPPS). [2] CPPS is a manufacturing control system concept that combines software (cyber) and hardware (physical) applications. It represents the integration of physical production processes with digital technologies, enabling the communication and coordination of machines, systems, and humans. CPPSs are also known as one of the fundamental cores of Industry 4.0. New technologies such as the Internet of Things (IoT), Artificial Intelligence (AI), cloud computing, or Service Oriented Architecture (SOA) are needed in connection with the above-mentioned CPPS. [3]

II. Scheduling algorithms in production

The discussed scope of the production includes a range of tools from operational analysis, statistical discovery and management. The issue is typified by the conflicting demands that are placed on production. On the one hand, there is the desire to maximize the utilization of production capacity with the shortest possible downtime and minimum inventories. On the other is the reality of limited resources, whether material, human, or time. This is why the efforts of many companies are mainly focused on the creation of a detailed schedule of production operations.

Three primary categories of methods exist for distributed manufacturing systems classification: [4]

- Operations research methods - work with exact mathematical models and algorithms. These methods include for instance linear programming, the simplex method, the branch and bound algorithm, or the Lagrangian relaxation method. The result is always the most optimal solution. However, become impractical when dealing with larger problem scales and constraint quantities, as the computation time increases exponentially.
- Heuristic methods - are used in cases where there is no known general algorithm for solving the problem, or where due to the complexity of the analyzed system it is not possible to construct a mathematical model. The principle is checking the search space and keeping the searched paths which contribute to forming the final solution. The result of stochastic methods are sub-optimal but in practice are often acceptable. One of them is the Hill Climbing algorithm.
- Artificial intelligence methods are using artificial intelligence tools such as neural networks, machine learning theory, fuzzy logic, or genetic algorithms.

III. HETERARCHICAL ARCHITECTURE

In terms of interrelationships and layout, we distinguish four types of control architectures for manufacturing. Each architecture represents a unique approach to production control system design. [5; 6] For purpose of this document we will focus on heterarchical architecture. These eliminate the concept of division into superior and subordinate entities, and a cooperative approach and the pursuit of complete local autonomy of entities come to the fore. This form of architecture is completely decentralized, made up of independent entities that have their own control algorithm and follow certain rules. Cooperation between the entities is ensured through a negotiation procedure. The entities negotiate with each other to arrange adaptive planning on the face of external stimuli, changes in production, or disturbances. Collaborative heterarchy, founded on negotiation should conform to the following regulations: [7]

- All parties have the same entitlements to resources,
- All parties operate autonomously,
- Entities have equal entry and availability to one another,
- All entities must strictly comply with the regulations of the system as a whole.

The main advantages of such an arrangement include high fault tolerance, flexibility, and a higher degree of eventual redundancy. The heterarchical approach also reduces the level of software complexity due to increased modularity. The shortcomings of heterarchical architecture include the risk of the non-deterministic and unpredictable behavior of the system as a whole since the absence of a parent component does not allow global optimization. Prediction is virtually impossible because future states are always the result of interactions between entities which can result in undesirable or even dangerous behavior of the system - technically called emergence. Some of the problems of this approach are also related to the heterogeneity of the used components and their incompatibility with each other. The lack of definition of methods of procedures and standards to coordinate the development of distributed control concepts also remains a problem. [7]

IV. INTELLIGENT MANUFACTURING SYSTEMS BASED ON AN AGENT-BASED APPROACH

We will consider an intelligent manufacturing system (IMS) as a system composed of multiple interdependent elements in an environment that possesses distributed intelligence and is in a cooperative heterarchy relationship. Distributing a complex problem among several parts of an intelligent system brings new innovative uses. Multi-agent systems and heterarchical systems based on the holonic principle could be the suitable implementation of such IMS which reflect the all mentioned requirements. [8]

Production systems based on the multiagent principle consist of virtual entities – agents. For the successful accomplishment of goals, agents use the principles of cooperation, mutual coordination, and negotiation mechanisms. Each entity's crucial role within the system is to act in a manner that maximizes the overall performance, ensuring its efficient operation and endurance, especially during critical situations. [9] It is worth mentioning one of the IMS paradigms for the manufacturing enterprise, which supports the necessity of implementing hierarchical structures to guarantee conflict resolution between autonomous entities. The goal of IMS is to combine the stability of centralized architecture while providing the dynamic flexibility of heterarchical systems. The result will be a hybrid structure, inspired, for example, by the functioning of living organisms in nature or the mechanisms of human society.

V. CPPS REFERENCE ARCHITECTURE

Architectures of intelligent manufacturing systems based on a multi-agent or holonic approach require formalizing. MAS concepts in the field of automation are represented by the name cyber-physical production systems (CPPS). CPPS is an innovative and promising paradigm for the future of manufacturing control. The smart manufacturing architecture of CPPS is defined by the combination of an embedded system i.e. a physical part equipped with process automation elements, and a cybernetic part representing the agent software.

Generally, there are two types of agents in CPPS that represent entities in the manufacturing domain:

- Resource agent - represents an intelligent software superstructure over the physical production equipment. The production equipment is supposed to be encapsulated as an agent. It can represent all resources across the production process from conveyors, cranes, and AGV systems to human resources. It transparently represents and maintains an up-to-date list of services it or the physical part can provide.
- Product agent - represents an intelligent software superstructure over the manufactured product. In particular, such an intelligent product should have the means to proactively allocate production requirements using a communication interface oriented towards resource agents. It should be also able to monitor its status, as well as the status of production equipment or other product agents. Based on this information, it can fully manage and optimally plan its production, taking into account also external factors entering the production process.

The proposed design pattern stipulates that the mentioned agents should be regarded as mandatory for an agent-based CPPS architecture, as they each perform essential functions. This pattern provides a MAS architecture that can handle the complexity of production and adjust as adaptively as required by CPPS. [10]

VI. ADAPTIVE ORGANIZATION IN CPPS

Using appropriate distributed control algorithms, individual machines (resource agents) and products (product agents) can make their own production management decisions regarding resource allocation and coordination using an automated form of negotiation. The intelligent product thus seeks the services of intelligent resources and thus manages its own production. The main advantage of such an approach is the above-mentioned feature of automated re-configuration if the planned production is changed in some way. An automated negotiation process takes place for each operation. It is based on supply and demand. [3]

The interface between the physical world and the control algorithm model is provided by the Asset Administration Shell (AAS) module. Each autonomous agent of decentralized production, for example, a product, a machine, a robot, a conveyor, a line, or a rack in a warehouse, is equipped with this module, which covers a multitude of functions. It provides directed identification and communication, status or lifecycle monitoring, or configuration of agent properties. In an SOA architecture, the module is also responsible for advertising its services, as well as leveraging the services of other agents in the network. AAS module interface over each agent also brings high scalability to the system, especially when incorporating a new agent - plug & produce (P&P). Web and cloud-based solutions are used to administer the specifications and manage the negotiation of such services. [11] Communication between physical devices may have a different implementation, in literature is often chosen the standard protocol OPC UA. Implementation is approaching another Industry 4.0 principle - IoT (Internet of Things). It is also necessary to mention the digital twin as another Industry 4.0 concept, which is linked to the integration of CPPS. A digital twin can create detailed virtual models of physical objects in virtual space to simulate their behavior in the real world and provide feedback during the system development phase, or serve as a reference for detecting errors during operation. The aim is thus to create intelligent factories that effectively reflect the requirements of multifaceted production management. [3]

VII. PROTOTYPE IMPLEMENTATION

The simulated manufacturing technology model was coded in C#. We took advantage of object-oriented programming principles and the .NET framework. This platform makes our solution possible to use cross-platform.

The algorithm aims to simulate dynamic manufacturing scheduling and rescheduling in the context of the above-mentioned multiagent approach. A simulated manufacturing environment consists of several resource agents, in a parallel arrangement with relationships among themselves. Resources may become unavailable and new resources may be added, while the start time and processing duration of a task can also vary. The complex combinatorial nature and its dynamic characteristics make it an excellent example for utilizing multi-agent systems.

To recognize whether a resource agent can meet the requirements of order was used a standardized model of segments. Each product agent has a defined list created by job decomposing into possible manufacturing actions. Each of these actions is represented by one Production Segment which contains the order of action in sequence, required resources, and other feature data. On the other hand, on the resource agent side, its capabilities are represented by a Process Segment structure. The segment structures used in the agents meet the basic concept of the ISA-95 data model. A segment in a recipe or production route is created by using the Personnel, Equipment, and Materials models in context to describe a unit of work in terms of resources.

The availability of the production path for each individual order (request) coming from a product agent is tested sequentially. After adding a new product agent (order), our proposed scheduling algorithm first tests the resource's own scheduler for available production capacity. Sequentially is tested the capacities of all required process nodes from the following operations i.e. testing the availability of the entire production path of the relevant product agent. The entire selection of the most suitable production path through the resource agent structure is in the control of the product agent. Also, the product agent requests entry into the reservation queues of the selected resource agents. From scheduling methods, we use an objective function of the total weighted job completion time to optimize the scheduling of the model situation. It is often interpreted as a measure of production work in progress. The deterministic rule simply selects the technology node that offers the earliest execution time from the provided options.

As our goal is to create an agent-based manufacturing scheduling system for a practical manufacturing enterprise, the system must be capable of dynamically responding to unexpected production orders and situations. To address possible emergent scenarios it is necessary to implement a mechanism for:

- Adjusting the schedule following a task's delay,
- Reorganizing the schedule after a machine malfunction or failure,
- Modifying the schedule after the cancellation of orders,
- Revising the schedule to accommodate new unexpected or emergent orders.

In our implementation communication is asynchronous and can be of point-to-point (between two agents), broadcast (one agent to all agents) type. For negotiation has been used Contract Net Protocol (CNP), is shown in Figure 2. The use of a bidding protocol like Contract Net is frequently referred to as a crucial element for an open and dynamic agent's environment. The product agent sends an ask for bids, evaluates them according to its selection criteria and awards the contract to the best offer(s). Between agents are sent defined dynamic objects which encapsulate the information being transmitted. These objects can be divided into two categories: requests and assertions.

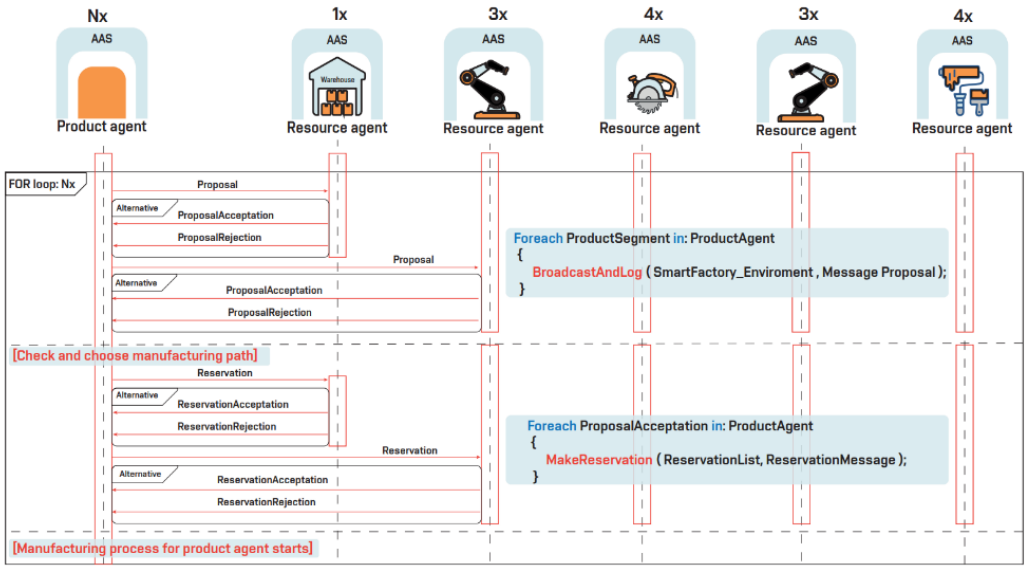


Figure 2: Direct communication between agents (CNP).

Future improvement could be using a framework that guarantees syntactic interoperability between different MAS. The FIPA ACL or KQML standards are manipulative languages based on communication acts where message content could be represented in XML language, which is an established standard for inter-computer communication.

VIII. Conclusion

In this paper, we discuss the current state-of-the-art in manufacturing control systems, especially innovative techniques and efficient mechanisms for dynamic scheduling mechanisms. Our attention is focused on the description of the multi-agent structure with distributed intelligence as a modern and effective tool in the field of management and organization of the production sphere.

In some degree of abstraction, our system simulates distributed intelligence across the manufacturing enterprise. The scientific problem that arises from the above is to identify and link sets of technologies in manufacturing to result in the realization of an intelligent heterarchical system. Rather than developing complex high-level mechanisms for multi-agent cooperation and negotiation, we focus on simple and reliable mechanisms. The proposed mechanism is based on negotiation supported by the bidding protocol Contract Net. The result is a straightforward yet effective solution for dynamic manufacturing scheduling. A multi-agent approach in manufacturing highlights autonomy, modularity, and openness of the manufacturing process. In today's highly competitive environment, such innovative production management can also lead to a significant competitive advantage.

Acknowledgment

The completion of this paper was made possible by grant No. FEKT-S-23-8451 "Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence, automation and measurement" financially supported by the Internal science fund of Brno University of Technology.

References

[1] BARBOSA, José, Leitão PAULO a Adam EMMANUEL. Structural Self-organized Holonic Multi-Agent Manufacturing Systems [online]. 0-13 [cit. 2023-03-09]. Dostupné z: https://uphf.hal.science/hal-03467679/file/HoloMAS'13-JB-PL-EA-DT.pdf

[2] SAKURADA, Lucas a Paulo LEITãO. Multi-Agent Systems to Implement Industry 4.0 Components: IEEE Conference on Industrial Cyberphysical Systems (ICPS) [online]. [cit. 2023-03-10]. Dostupné z: doi:10.1109/ICPS48405.2020.9274745

[3] SEITZ, Matthias, Felix GEHLHOF, Luis CRUZ SALAZAR, Alexander FAY a Birgit VOGEL-HEUSER. Automation platform independent multi-agent system for robust networks of production resources in industry 4.0. Journal of Intelligent Manufacturing [online]. [cit. 2022-12-27]. Dostupné z: https://link.springer.com/article/10.1007/s10845-021-01759-2

[4] PINEDO, Michael L. Scheduling Theory, Algoritms and Systems [online]. Fifth Edition. NYU Stern School of Business New York, NY, USA [cit. 2022-11-13]. ISBN 978-3-319-26580-3.

[5] DUFFIE, Neil A. An Approach to the Design of Distributed Machinery Control Systems [online]. IEEE, 435-442 [cit. 2022-11-07]. Dostupné z: doi:10.1109/TIA.1982.4504105

[6] DILTS, D.M., N.P. BOYD a H.H. WHORMS. The evolution of control architectures for automated manufacturing systems. Journal of Manufacturing Systems [online]. 1991, 79-93 [cit. 2022-11-07]. Dostupné z: doi:0278-6125

[7] DUFFIE, Neil A. Synthesis of Heterarchical manufacturing systems: Computers in Industry [online]. 167-174 [cit. 2022-11-07]. Dostupné z: doi:0166-3615

[8] UPADHYAY, D. B. A review paper on Multi agent base intelligent manufacturing system [online]. [cit. 2023-03-10]. Dostupné z: https://ijaers.com/Paper-Aug%202014/IJAERS-AUG-2014-011.pdf

[9] KUBÍK, Aleš. Inteligentní agenty: Tvorba aplikačního software na bázi multiagentových systémů. 1rd edition. Brno: Computer Press, Computer Press. ISBN 8025103234.

[10] CRUZ SALAZAR, Luis, Daria RYASHENTSEVA, Arndt LÜDER a Birgit VOGEL-HEUSER. Cyber-physical production systems architecture based on multi-agent's design pattern: Comparison of selected *approaches mapping four agent patterns* [online]. [cit. 2023-03-10]. Dostupné z: https://doi.org/10.1007/s00170-019-03800-4

[11] ARM, Jakub, Tomas BENESL, Petr MARCON, Zdenek BRADAC a Frantisek ZEZULKA. Automated Design and Integration of Asset Administration Shells in Components of *Industry 4.0 [online]*. [cit. 2023-03-01]. Dostupné z: https://doi.org/10.3390/s21062004

Emotion recognition system

Jan Wójcik
Department of Control and Instrumentation
Brno University of Technology, Faculty of Electrical Engineering and Communication
Brno, Czech Republic
xwojci00@vutbr.cz

***Abstract*—This article deals with the design of an emotion recognition system, which should be used as a communication tool for people with an autism spectrum disorder. Camera data will be used for real-time emotion recognition, with appropriate output for the user. The paper deals with areas such as face detection, extraction of relevant features, finding a suitable dataset, or designing a classifier.**

Keywords— Emotion, recognition, computer vision, face detection, decision tree

I. Introduction

Emotions are an integral part of human being's behavior and are essential in his non-verbal communication. By showing emotion, a person can express their current mood, mental state, or feelings. When communicating, the other person can identify these expressed emotions and determine the current mood or feelings of the person, which can be helpful. However, not all people can do this. For example, people with autism spectrum disorder have difficulty recognizing people's emotions and the system could help them with communication. Although the primary purpose of the system is already defined, it should not be a problem to use the system for emotion recognition for other purposes.

II. Features extraction

In the area of computer vision, features can be imagined as pieces of information obtained from the analyzed picture. Features extraction for emotion recognition can be done in two main areas.

A. Face

Features extraction from the face is a natural approach because emotions are expressed as a face muscular activity. According to Facial Action Coding System (FACS) [1][2][3], every emotion can be expressed as a set of Action Units (AUs). Action unit refers to the muscular activity that is triggered by a change in facial expression. There is a defined set of muscles for each AU. When these muscles are contracted, AU is active, when relaxed AU is inactive. Besides AU presence, FACS also defines AU intensity. Action unit presence and intensity can be used as features for emotion recognition.

B. Gestures and posture

Both gestures and posture are part of the body language area, and they play a significant role in expressing emotion. People use gestures in their non-verbal communication and often to express a certain feeling. Extracting features from posture and gestures can be used as an additional approach for emotion recognition when there is not possible to extract features from the face. For example, the face is covered or censored. A suitable example can be micro-gestures [4], which are expressed unintentionally, e.g., face covering, eyes rubbing, nails biting.

III. Datasets

The dataset represents a database of annotated pictures, which are crucial for training and validation. There are many available datasets for emotion recognition, but the quality is often very poor. The following datasets were tested and used for designing an emotion recognition system.

A. FER2013 [5]

Data consists of grayscale images with a size of 48x48 pixels. The faces are already centered, so each face occupies approximately the same area in the image. The database contains about 30,000 images, which are categorized into seven groups corresponding to the expressed emotion. The problem with this dataset was its small resolution. It was almost impossible to extract AUs from the faces.

Example of FER2013 dataset is shown in the picture below. Each column contains four samples of images. Each column represents one emotion class with the following order: anger, disgust, fear, happiness, neutral, sadness and surprise.



Figure 1 - FER2013 data examples [5]

B. AffectNet

It is a vast database of colored images with much higher image resolution than the FER2013 dataset. The typical image size in the AffectNet database is 512x512 pixels. The database consists of approximately one million images, 440,000 of which have been manually annotated and added into 8 emotion classes (same as FER2013 + emotion class contempt). Unfortunately, the data is often inaccurately annotated. Therefore, there was a need to check it and delete or move

incorrectly annotated pictures. Only 300 pictures per class were selected due to the high time demands. The following picture shows some of the selected pictures. Each column contains four samples for each emotion. [6][7]

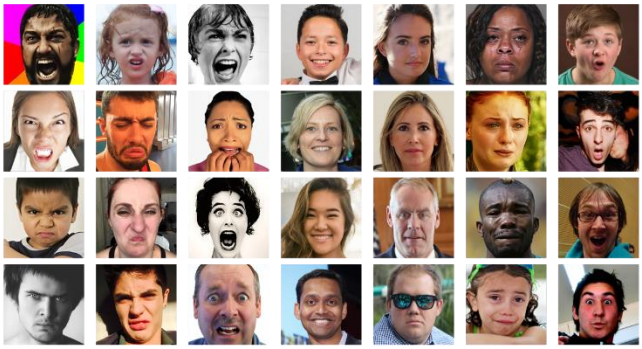


Figure 2- AffectNet images [6][7]

IV. METHODS

Emotion recognition is a popular area of computer vision. Therefore, many methods and algorithms that can be used for the extraction of relevant features either from face or posture already exist. The methods mentioned below were tested and modified if needed for getting the best possible results.

A. *DeepFace*

DeepFace is a lightweight python library for face recognition and analysis of facial attributes such as age, gender, emotion and race. It also allows the user to select one of the pre-trained models. The DeepFace emotion model can classify data into one of the seven emotion classes that were mentioned before. [9][10]

B. *OpenFace*

OpenFace [12][13] is a computer vision toolbox, which provides human face analysis, specifically, it can detect facial landmarks, estimate head position, recognize facial actions (AUs) and also estimate the direction of a person's gaze. It can detect the presence and intensity of most of the AUs described in FACS.

C. *OpenPose*

OpenPose [11] is a toolbox that allows you to detect the positions of points of the human skeleton, hands and face in real-time. OpenPose allows user to detect the points of interest for multiple people at once in a single frame, and in total it can track up to 137 different points for each person.

V. RESULTS

Both DeepFace and OpenFace were tested on AffectNet and FER2013 dataset. Performance was evaluated using a confusion matrix. Unfortunately, FER2013 dataset was found unsuitable due to the low image resolution, which led to unsuccessful feature extraction using OpenFace. In addition, FER2013 was used for training of the DeepFace emotion model. Therefore, DeepFace performance results on FER2013 dataset were most probably biased. For comparison, the overall accuracy of DeepFace emotion model for the FER2013 dataset was 83.81% while the overall accuracy for AffectNet dataset was only 42%. DeepFace built-in functions were used for the emotion prediction. Results of DeepFace performing on the AffectNet dataset are in following confusion matrix.

Overall accuracy 42.00 %		**Reality**								
		Angry	Disgust	Fear	Happiness	Neutral	Sadness	Surprise	Total	Precision
Prediction	Anger	22	14	3	1	5	3	5	53	42 %
	Disgust	2	6	0	0	0	0	0	8	75 %
	Fear	9	4	18	3	4	9	6	53	34 %
	Happiness	4	8	10	40	2	10	3	77	52 %
	Neutral	6	10	4	4	26	12	11	73	36 %
	Sadness	6	8	3	2	13	15	5	52	29 %
	Surprise	1	0	12	0	0	1	20	34	59 %
	Total	50	50	50	50	50	50	50		
	Recall	44 %	12 %	36 %	80 %	52 %	30 %	40 %		
	Accuracy	83 %	87 %	81 %	87 %	80 %	79 %	87 %		
	F1 score	43 %	21 %	35 %	63 %	42 %	29 %	48 %		

Figure 3 - DeepFace + AffectNet confusion matrix

OpenFace has no functionality to directly predict emotion from a given picture. Therefore, some modifications needed to be performed. Action unit presence and intensity ware used as extracted features and decision tree classifier was trained for emotion classification [14][15]. A fraction of the AffectNet dataset, which was manually revised and reannotated, was used. This fraction was split in 4:1 ratio with the larger portion being the training set and the smaller the testing set. The resulting confusion matrix can be found in following picture.

Overall accuracy 62,72 %		**Reality**								
		Anger	Disgust	Fear	Happiness	Neutral	Sadness	Surprise	Total	Precision
Prediction	Anger	14	9	0	0	1	1	0	25	56 %
	Disgust	5	21	0	0	2	0	0	28	75 %
	Fear	1	0	15	0	0	0	5	21	71 %
	Happiness	0	0	4	49	0	1	0	54	91 %
	Neutral	6	4	10	1	38	16	5	80	47 %
	Sadness	12	8	1	0	2	18	2	43	42 %
	Surprise	0	0	4	0	7	0	25	36	69 %
	Total	38	42	34	50	50	36	37		
	Recall	37 %	50 %	44 %	98 %	76 %	50 %	68 %		
	Accuracy*	87,8	90,24	91,29	97,91	81,18	85,02	91,99	*(%)	
	F score	44 %	60 %	55 %	94 %	58 %	46 %	68 %		

Figure 4 - OpenFace + AffectNet confusion matrix

Classifier was very successful in predicting happy emotion. Happiness was recognized with 91 % precision and 97.91 % accuracy. F score for that emotion class was the highest of all classes, with the value of 94 %. By percentage, the second most precisely classified emotion was disgust, with 71 % of precision value. However, the recall of that emotion was only 50 %, that means that only half of the pictures with disgust emotion were recognized. Interesting thing is that disgust emotion was very

often confused by angry emotion. From totally 25 predicted angry emotion, 9 were actually disgust emotion. A better parameter for comparison is F score, which combines both recall and precision. The second highest F score has emotion surprise with the value of 68 %. Still, there is a quite big difference comparing F score of happy emotion.

Classifier performed with overall accuracy 62.72 % on static images. It was trained on approximately 200 pictures per class, however, features were not successfully extracted from every picture, so in reality, there were less than 200 pictures for each class. The same problem occurred with validation, not all 50 pictures of the testing set were successfully analyzed. Better quality results require a bigger dataset.

OpenPose was also tested on both datasets, unfortunately, it could not extract any valid data. Datasets already contain cropped images, so only a face is present. Therefore, OpenPose cannot detect any points related to human posture. However, for the emotion recognition system is assumed, that person will be recorded from the waist up. At least arms and hands will be visible and some relevant data could be extracted from OpenPose.

A. *Classifier design*

Using the scikit-learn library, Python code was written to train the decision tree on the training data and evaluate its classification on the test data. The scikit-learn library [14] uses the optimized CART (Classification and Regression Trees) algorithm and allows setting the criterion function, splitting type, maximum tree depth, minimum number of samples in the tree leaf, etc. With the help of the matplotlib library, the generated decision tree can also be rendered as an image.

The overall classification accuracy was monitored during the decision tree learning process. Learning was executed with the following constraints. The depth of the tree was constrained in the range of 1-15 and the minimum number of samples in the leaf of the tree was constrained in the range of 5-40. All combinations were tested for the given constraint conditions and the best overall accuracy was achieved for a tree with an entropy criterion function, depth of 7 and a minimum number of samples in the leaf of 37.

The decision tree was trained and tested on static pictures using detected action units by OpenFace. However, OpenFace can analyse real-time data from a webcam for example. Authors of OpenFace also claim that action units' detection is more accurate when applied to dynamic data [13]. Decision tree classifier was implemented into OpenFace graphical user interface. Emotion prediction was visualized to the user in textual form. The result is shown in the following picture.

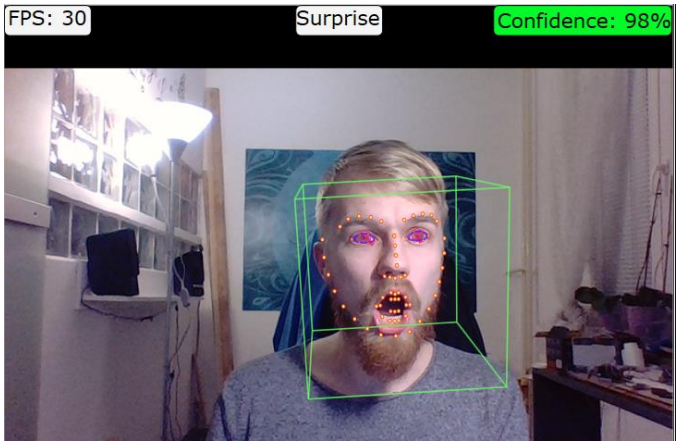


Figure 5 - Example of modified OpenFace GUI

The correct classification of surprised emotion can be seen in Figure 5. OpenFace GUI also informs the user about actual frames per second and confidence of extracting features from the face. Confidence significantly drops when a person covers his face with hands. In that situation features extracted from arms position would be beneficial. The combination of OpenFace and OpenPose could be promising.

OpenFace GUI also indicates the presence and intensity of each AU. Detected action units are AU1, AU2, AU4, AU5, AU15 and AU26. According to the FACS, surprised emotion can be a combination of AU1, AU2, AU5, AU26/AU27 [2][3]. All of them are successfully detected by OpenFace as shown in following picture.

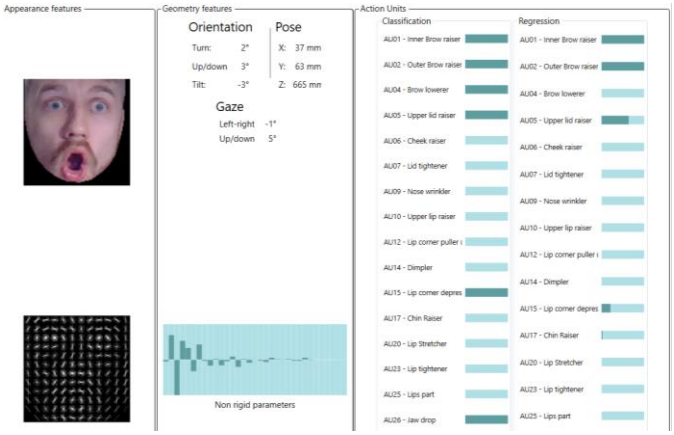


Figure 6 - OpenFace and detected AUs

B. *Challenges*

Obtaining a sufficiently large and good-quality dataset is not an easy task to do. There are many public datasets, but the quality of annotation is very poor. This can be solved by manual reannotation, however, this is time-consuming task.

Another issue is that people can recognize emotions in different ways. For some people, the person in the picture may seem angry, but for some, it can have a neutral expression. Recognizing emotion is a very subjective task.

For the best results, it will be necessary to implement more than one of the mentioned methods to analyze real-time data

from the camera. The problem is that some of them are written in C++ and some of them in Python. Merging solutions into one can be challenging.

VI. Conclusion

Two approaches for features extraction were described within the paper. The first one extracts features from face, the second one extracts features from gestures and posture. DeepFace and OpenFace were tested as methods of extracting features from the face and classifying the emotion. DeepFace and its learned emotion model classified emotion directly from the image. OpenFace was used as tool that detects muscular activity in the face – action units. The decision tree was then trained and used as a classifier. Detected action units were used as input data for decision tree. Comparing these two methods, decision tree with OpenFace had better results, which are presented in Figure 4 - .

OpenFace GUI was then slightly modified and decision tree was implemented here. Result can be seen in Figure 5 - Example of modified OpenFace GUI. OpenFace processes real-time data, detects action units and output of the classifier is presented to the user.

OpenPose was tested as a method of extracting features from gestures and posture. It successfully detected joints and skeleton of the person in the picture, but no more detailed results are present. Extracted data from OpenPose could be simultaneously used with OpenFace and decision tree output. Combining these two approaches could improve the classification performance.

All of the methods were tested on both datasets, FER2013 and AffectNet. FER2013 could not be analyzed via OpenFace, probably due to the low image resolution. AffectNet dataset had better resolution than FER2013, but problem was with incorrectly annotated data. Fraction of AffectNet dataset was then manually reannotated, but size of reannotated dataset is not sufficient. Generated Images contains only two emotion classes, which is not enough, but it could be used to expand the AffectNet dataset.

All computational operations were performed on a laptop with an Intel i5 7300HQ processor and an Nvidia GeForce GTX1050 graphics card.

ACKNOWLEDGEMENTS

I would like to thank my diploma thesis supervisor Ing. Ilona Janáková Ph.D for valuable advices and support.

References

[1] EKMAN, Paul, Wallace V. FRIESEN a Joseph C. HAGER. *Facial Action Coding System: The manual*. United States of America: Research Nexus division of Network Information Research Corporation, 2002. ISBN 0-931835-01-1.

[2] Facial Action Coding System (FACS) – A Visual Guidebook. *Imotions* [online]. 2022 [cit. 2022-10-26]. URL: https://imotions.com/blog/facial-action-coding-system/#emotions-action-units

[3] The Facial Action Coding System for Characterization of Human Affective Response to Consumer Product-Based Stimuli: A Systematic Review. *Frontiersin* [online]. United States, 2020 [cit. 2022-10-26]. URL:https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00920/full#SM1

[4] LIU, Xin, Henglin SHI, Haoyu CHEN, Zitong YU, Xiaobai LI a Guoying ZHAO. *IMiGUE: An Identity-free Video Dataset for Micro-Gesture Understanding and Emotion Analysis* [online]. 1-8 [cit. 2022-11-14]. URL: https://openaccess.thecvf.com/content/CVPR2021/papers/Liu_iMiGUE_An_Identity-Free_Video_Dataset_for_Micro-Gesture_Understanding_and_Emotion_CVPR_2021_paper.pdf

[5] *Challenges in Representation Learning: Facial Expression Recognition Challenge* [online]. 2013 [cit. 2022-11-29]. URL: https://www.kaggle.com/competitions/challenges-in-representation-learning-facial-expression-recognition-challenge/data

[6] MOLLAHOSSEINI, Ali, Behzad HASANI a Mohammad H. MAHOOR. *AffectNet: A New Database for Facial Expression, Valence, and Arousal Computation in the Wild* [online]. In: . IEEE Transactions on Affective Computing, 2017 [cit. 2022-12-06].

[7] *AffectNet-HQ* [online]. [cit. 2022-12-06]. URL: https://www.kaggle.com/datasets/tom99763/affectnethq

[8] *Academic Dataset by Generated Photos* [online]. [cit. 2023-03-08]. URL: https://generated.photos/datasets

[9] SERENGIL, Sefik Ilkin. *Face Alignment for Face Recognition in Python within OpenCV* [online]. 2020 [cit. 2022-10-31]. URL: https://sefiks.com/2020/02/23/face-alignment-for-face-recognition-in-python-within-opencv/

[10] *DeepFace: GitHub* [online]. 2022 [cit. 2022-11-29]. URL: https://github.com/serengil/deepface

[11] *OpenPose* [online].[cit. 2022-12-30]. URL:https://cmu-perceptual-computinglab.github.io/openpose/web/html/doc/md_doc_02_output.html

[12] BALTRUSAITIS, Tadas, Amir ZADEH, Yao Chong LIM a Louis-Philippe MORENCY. *OpenFace 2.0: Facial Behavior Analysis Toolkit* [online]. Xi'an, Čína, 2018 [cit. 2022-11-29]. URL: doi:10.1109/FG.2018.00019

[13] *OpenFace: GitHub* [online]. 2021 [cit. 2022-11-29]. URL: https://github.com/TadasBaltrusaitis/OpenFace

[14] *Scikit-learn: Decision trees* [online]. [cit. 2022-12-29]. URL: https://scikit-learn.org/stable/modules/tree.html

[15] *Python - Decision tree implementation* [online]. [cit. 2022-12-29]. URL: https://www.geeksforgeeks.org/decision-tree-implementation-python/

Design of data logger for train safety management unit

Michal Dvorak
Faculty of Electrical Engineering and Communication
Vysoké učení technické v Brně
Brno, Czech Republic
xdvora2i@vutbr.cz

***Abstract*—Even though the railway technology is more than two hundred years old, today it is still a vital mode of transportation, be it for transport of goods, or people. Even though hayday of railway and trains of late 19th century seems all but over, in some countries, namely in east Asia or in western Europe railway is bedrock of public transport over great distances. In other parts of the world, where trains were thought obsolete by invention of car for public transport, slow but noticeable change of approach can be seen (United states, Australia…). Even though invention of early 19th century, train seems to be a cure for a lot of 21st century problems and ailments, such as city centers congested by car traffic and huge CO2 production by personal transport. To transport safely passengers at speeds of 350+km/h safely, comfortably and on time, modern trains are incredibly complex systems employing vast array of safety electronical systems.**

I. Introduction

With Europe gradually integrating throughout the second half of the 20th century, its railways stayed hardly compatible. Vast array of railway gauges, safety systems and traction voltage meant that while travel inside a country was comfortable and reliable, traveling between countries poses somewhat a logistical challenge, rather than common convenience. Scope of the problem is seen in the Figure 1,Figure 2 and Figure 3. This means that a trip from Vienna to Paris takes 12 hours, with two change and costs somewhere between 3000 and 5000CZK. For a contrast, trip from Beijing to Shanghai, roughly same distance from Vienna to Paris, takes less, then 4 hours [6]. This means that most of the travelers select a plane as their preferred method of transport. Be it because they want to save time, or money. This create huge strain on the environment, since the plane transport produces largest amount of CO2 per passenger/kilometer (Figure 4). To tackle this serious problem, countries of European Union agreed on introducing one single safety standard – European train control system, or ERTMS/ETCS. Proposed systems of ERTMS and ETCS are designed to work in tandem with safety systems already employed without conflicts. To mirror the possibilities of the railway where train is currently going ETCS is employed in four levels. This level is an indicator to "how much" control is done by ETCS. The development of ETCS has matured to a point that cross-border traffic is possible and some countries have announced a date for the end of older systems. However, the effort is very far from done. For example Germany wants to have 80% of its railway traffic controlled by ETCS by 2050 [5].

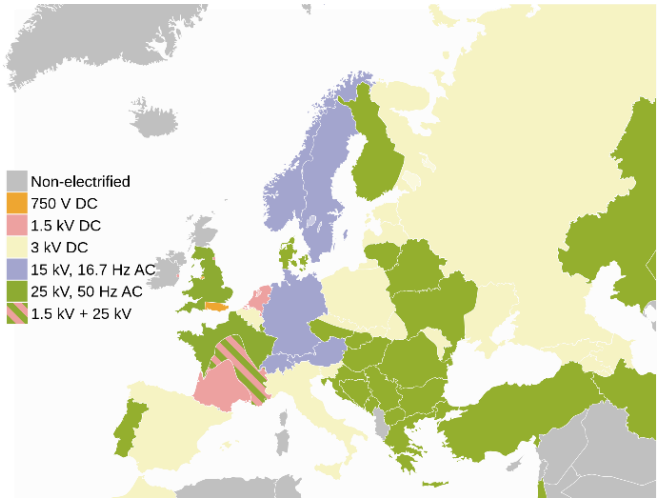


Figure 1 - Europe rail electrification system [1].

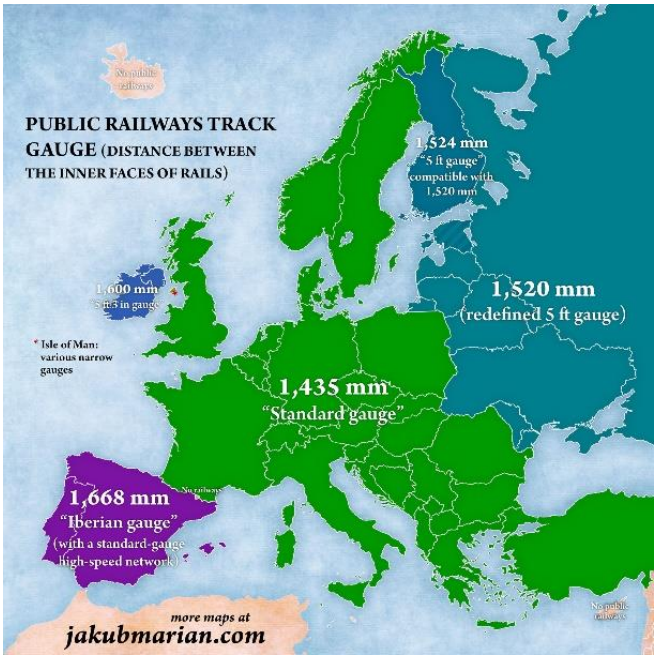


Figure 2 - European rail gauges [2]

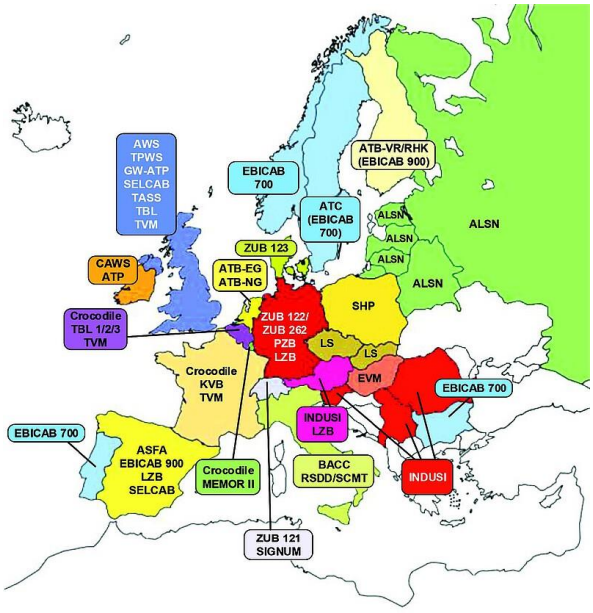


Figure 3 - Safety systems used in European railway transport [3]

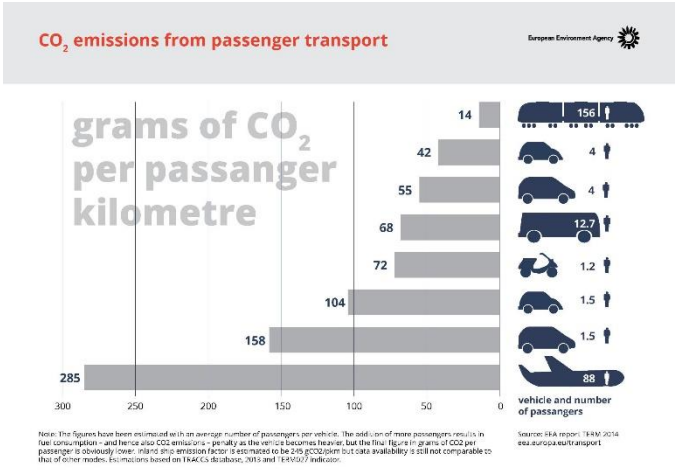


Figure 4 - Modes of transport by Co2 per passenger/kilometer [4]

II. Railway Safety Systems

A. Current railway safety systems

Plan of European Union counts on 2500km of new ETCS operated railway by 2030. This means, that vast majority of rail lines will still operate the previously used systems. As an example of an old system in use, let's use Czech LS (Current version LS06 or LS07).

In this specific case a given part of the rail is equipped with a circuit that can transfer a code for specific instruction to the rail car. These codes are giving information to train conductor about the status on the rail before him. However, this system is very rigid and does not allow for a smooth operation of the railway.

B. ERTMS/ETCS

The „European rail traffic manager", is a safety system that seeks to not only unify the various European train safety systems already in place, but to also expand upon them. It provides train conductor and crew with much more information, than systems previously in use. It communicates much more data in real time through GSM and other means to ensure smoother operation of the railway. This comes with a challenge to design electronics more capable in train vehicles.

ETCS can be divided into four levels depending on a railroad equipment and how the information is transmitted into the train [].

- Level 0: When a ETCS train is used on non ETCS rail, the train-borne equipment monitors the train for maximum speed of that type of train. Furthermore, the train driver has to observe the trackside signals.
- Level 1: Level 1 of ETCS can be superimposed on existing signaling system. The main communication element is **eurobalise**, which is equipped with radio beacon. This eurobalise will pick up trackside signals and transmit them as a **vehicle movement authority (**permission to cross a certain point)**.** Onboard computer continuously calculates the maximum speed and breaking distance from this data provided [].
- Level 2: Is a radio-based signal and protection train system. For this system level is possible to work without trackside signaling. Train movements are monitored continually by the Radio block centre. Any movement authority is transmitted to the train continuously via GSM-R together with speed information and route data. Eurobalises in this level are used as passive milestones [].
- Level 3: Offers a full control of the train and spacing. Fixed track signaling is no longer required.

III. Device Requirements

Electronics used in trains and railway is subjected to standard that it must uphold. Most importantly EN 50124-1 and EN 50155, but in reality many more. With this comes what is actually expected from the logger from a functional perspective.

A. Standard requirements

Standard EN50155 quotes that the device must operate with temporary power supply drop to $0.6V_{PWR}$ and it must survive without a damage to the system a sudden power surge to $1.4V_{PWR}$. Also, used components must be selected as such, that designed device must operate in temperature range from -45℃ to 85℃.

Standard EN 50124-1 gives a closer on look how wide isolation barriers on a designed PCB will be used.

B. Functional requirements

Requirements for the data logger are following:

- Ability to facilitate and drive two independent 1GB/s ethernet ports with industrial connector.
- Provide USB-C connector for interface and data access.
- Save the data on the SD card, which should be accessible to the operator.

Identify applicable funding agency here. If none, delete this text box.

- Be equipped with LED signaling, informing the operator about the status of the vehicle.
- Dimensions of data logger must fit euro card dimensions (220x100mm).
- Be able to be safely supplied from the power supply of $V_{PWR} = 24V$.
- Longevity of the components, meaning high rating of safety and ability to operate in harsh environments and long term support by supplier.

IV. Design of Data Logger

A. Computational unit SoM

Main component that the design of the data logger was based around was the computational module DART-MX8M-PLUS V1. This system on module (SoM) was selected because it meets all the necessary criteria for the data logger. It provides sufficient computational power meeting requirements, it is manufactured in industrial standard meeting temperature and ruggedness requirements and it is of very small footprint and reasonable price. Moreover, manufacturer is guaranteeing its manufacturing until 2036. Manufacturer also guarantees maximum power consumption of $P_{MAX} = 2.79W$



Figure 5 - Computational unit for data logger [5]

B. Ethernet interface design

One of the main requirements is for the SoM to be able to facilitate two 1GB/s Ethernet ports ended with industrial standard M12 connectors. One Ethernet port fully equipped with physical layer is present on the SoM module. Task therefore is to equip the DART with other physical layer interface.

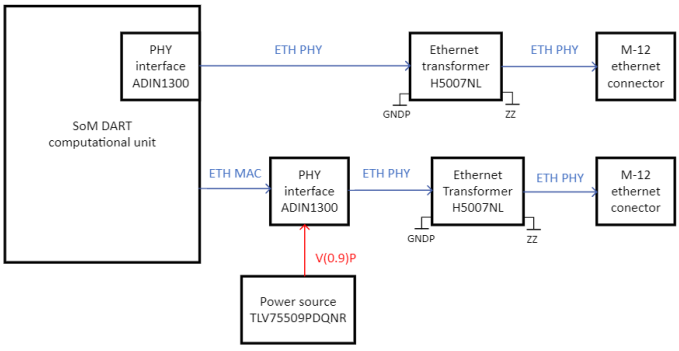


Figure 6 - Ethernet configuration block diagram

C. Power delivery

The main power source to the data logger is provided by the safety unit. This voltage $V_{PWR} = 24\,V$ and can vary by ±40%. Therefore a power unit THM15-2411WI, with offers a wide input voltage scale will be employed and maximum power draw of $P_{MAX} = 15W$. This unit will also act as a galvanic protection of the more sensitive circuitry. This power source is providing voltage $V(5.0)P = 5\,V$. This voltage feeds necessary peripherals or is transformed to lower voltage values.
Computational unit SoM has operational supply voltage range $V_{BAT} = 3.3 - 5V$. Both 3.3V and 5V is used in the design of the data logger. However, it is undesirable to supply main computational unit with either voltage extreme. We want to prevent overstress of the component and also we don't want power cuts in case of slight voltage drop. It was decided that its own power source $V_{DART} = 3.8\,V$ will be used. This voltage with be provided by power source LM2640. This power source is double channeled and allow for power draw of $I_{OUT} = 2\,A$, so it will reduce the space on the board.

D. User interface USB

To access data stored to the either eMMC or SD card data logger will be equipped with accessible USB-C connector equipped with a safety screws. Used SoM provides all the necessary peripherals to employ USB 3 with minimum amount of additional components.

E. SD CARD

In a case of a critical failure the system is equipped with a SD card. In a case of an emergency signal the system will start rewriting data from eMMC to SD card. SoM also allows to boot either from the SD or eMMC. Used micro-SD card will be placed in accessible holder.

V. Conclusion

Design considerations were put into the context of future aims for railway integration in European Union and beyond. Requirements and components used were introduced and their future use in design of final product was established.

A. Authors and Affiliations

Michal Dvořák – Faculty of Electrical Engineering and Communication, Brno University of Technology

Acknowledgment *(Heading 5)*

This work is conducted under supervision of the company AŽD Praha and namely Ing. Libor Šimek as a professional supervisor. I would like to thank him for the guidance and help with this work

References

[1] https://en.wikipedia.org/wiki/Rail_transport_in_Europe
[2] https://jakubmarian.com/wp-content/uploads/2018/11/railway-gauges-europe.jpg
[3] https://www.trackopedia.com/en/encyclopedia/infrastructure/european-train-control-system-etcs
[4] "Die Finanzierung der Eisenbahn des Bundes," Deutsche Bahn AG, 2015

Instrument power supply unit

Bc. Petr Navrátil
Department of Power Electrical and Electronic Engineering
Brno University of Technology, FEEC
Brno, Czech Republic
xnavra63@vutbr.cz

Ing. Radek Tománek
Department of Power Electrical and Electronic Engineering
Brno University of Technology, FEEC
Brno, Czech Republic
Radek.Tomanek@vut.cz

***Abstract*—This paper introduces the reader to the issue of switched-mode power supplies, specifically flyback topology sources for powering a private branch exchange. It also introduces the reader to the results of the thesis so far, where the design is analyzed in more detail.**

Keywords—flyback converter, pulse width modulation, power supply, pure sine wave, H-bridge, low pass filter

I. Introduction

Most complex electronic devices require a mains power supply with multiple outputs of different parameters. Apart from the requirements for different combinations of output voltages, current dimensioning, stability and output noise, there are often requirements for the ability to switch off some of the power supply, ensuring minimal power consumption from the network in standby mode. The optimal design of such a power supply unit usually needs to be tailored specifically to the given application, including the design of individual transformers and inductors.

II. Power Supply Unit

Overall the power supply unit comprises several switching sources, each serving a specific function. These sources are supplied by mains voltage rectified and filtered using the first block in Figure 1. The power supply unit was designed to minimize standby power consumption as possible. The standby power consumption was reduced by using the resistor for the initial charging of the UCC28C44 power supply capacitor only for A source. The other sources are powered by the auxiliary winding of source A. Two ultra-fast diodes were employed to rectify the auxiliary power supply of individual sources to prevent needlessly discharging the supply capacitor of A source by B and C sources during standby mode.

All three power sources use the flyback converters topology, but each operates at a different switching frequency. They were all designed with a UCC28C44 controller operating in current mode PWM. The control circuits are based on the sample circuit in the DataSheet [3], with additional switching logic for individual sources, the additional LC filter on the output, or adjustments for the maximum overall efficiency of the power supply unit.

If a flyback topology has multiple secondary windings, the individual voltages may become unbalanced with uneven loading. This situation isn't possible to regulate by the central controller of the flyback converter, so two out of three windings use output linear regulators, and the voltage feedback is taken from the output without the linear regulator. The winding for voltage feedback is chosen for the highest output current because this is where the linear regulator would have the highest losses. All sources have a filter capacitor at the output, and A and B sources also have an LC filter at the output to achieve the lowest possible ripple and noise of the output voltage.

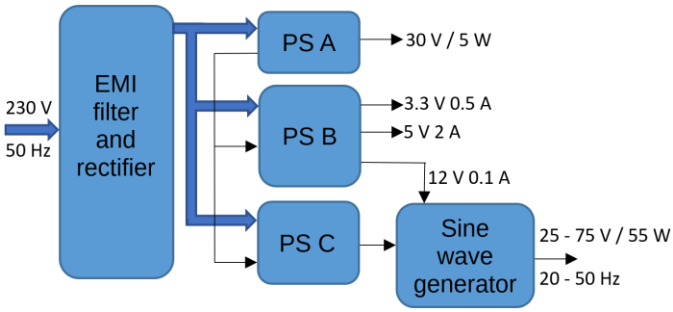


Fig. 1. Block diagram Power Supply unit for PBX

A. Power supply A

The power supply A is used to power the DC-loop during the off-hook handset. When the handset is off-hook, the current detector of A source turns on B source, which powers the auxiliary winding. Power supply A was designed to output 30 V and deliver 5 W. It operates at a frequency of 52 kHz. Source A contains an auxiliary winding to power the UCC28C44 control circuit on the primary side of the inverter. Other sources are also powered by this auxiliary winding to save energy.

B. Power supply B

Power supply B has three DC voltage branches with the following parameters: 3.3 V/0.5 A, 5 V/1 A and 12 V/0.1 A.

C. Power supply C

Power supply C has an output DC voltage range of 40 V to 120 V and output power up to 55 W. The voltage control of this source is achieved using a microprocessor from the STM series. The output voltage of the sinusoidal inverter is controlled by varying the DC output voltage of the B source.

D. Sine wave converter

The sine wave source has a sine output voltage range of 25 V to 75 V at a frequency range of 20 Hz to 50 Hz up to power 50 W. The change in the output frequency will be controlled by the same microprocessor that regulates the output voltage of the C source. In order to achieve a constant PWM resolution, the change in the effective output voltage of the sine wave source is regulated by the output voltage of source C. The sine wave source was used instead of a mains transformer due to limited space in the Tesla Almes cabinet.

III. Switching on source B and C

Waking up the private branch exchange from standby mode is handled by a current detector included at the output of source A.

Off-hooking any telephone handset closes the DC loop, which starts to consume DC current from source A. This is detected by a current detector, which ensures that source B is switched on.

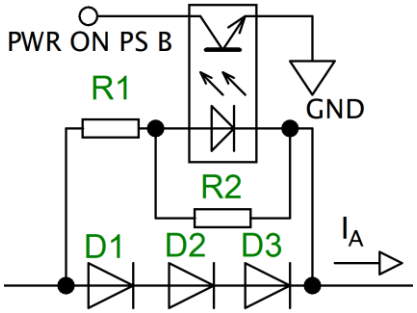


Fig. 2. Circuit diagram of the current detector

The circuit diagram of the current detector is in Fig. 2. The current drawn from power supply A causes a voltage drop across resistor R2. If the drawn current increases to a limit such that the threshold voltage of the optocoupler diode is on the resistor R2, the optocoupler will be switched on. Resistor R1 serves to limit the current through the optocoupler. If the current consumption from power supply A rises above a certain limit, this current will be bypassed using a series combination of three rectifier diodes.

To prevent switching off of power supply B during fluctuations, there is a delay RC element in the control circuit, which controls an N-type MOSFET to negate the signal logic. Additionally, one more optocoupler was added for switching power supply B in case of maintenance or upgrading a new firmware in the main microcontroller.

Power supply C is switched on by the controlling STM microcontroller which is powered by the B power supply. Source C is switched on only when it is required ringing of some phone.

IV. Filters design

A. Design of line voltage filter capacitors

Filter capacitors for line voltage were designed using the basic differential equation for capacitor capacitance (1). The maximum voltage drop at the average current was chosen, and the angle was determined from this drop using the sine function. The discharge time was then calculated from the angle [2].

$$C_f = I_{AV} \cdot \frac{t_{dis}}{\Delta U} \tag{1}$$

B. Design of output filter capacitors

The output voltage filter capacitors were designed using the basic differential relation for capacitor capacitance (1). The maximum voltage ripple at medium current was chosen, from which the discharge time was calculated assuming that the current through the secondary winding decreases linearly to zero. The discharge time was considered from the point when the current through the secondary winding is equal to the average current drawn, up to the beginning of the next current pulse from the secondary winding [1]. Ceramic capacitors were used for the filters in sources A and B because they have negligible ESL. The combination of multiple capacitors in parallel was used to minimize ESR.

C. Design of additional LC filter for outputs of power supplies A and B

The voltage transfer was calculated using the voltage ripple across the first filter capacitor and the selected output voltage ripple. The required resonant frequency of the LC filter was then calculated from this voltage transfer (2). After that, it was necessary to choose the inductance of the filter and then the capacitor capacitance can be calculated using Thomson's formula [1].

$$f_r = \frac{f}{10^{\frac{A_{u,dB}}{40}}} \tag{2}$$

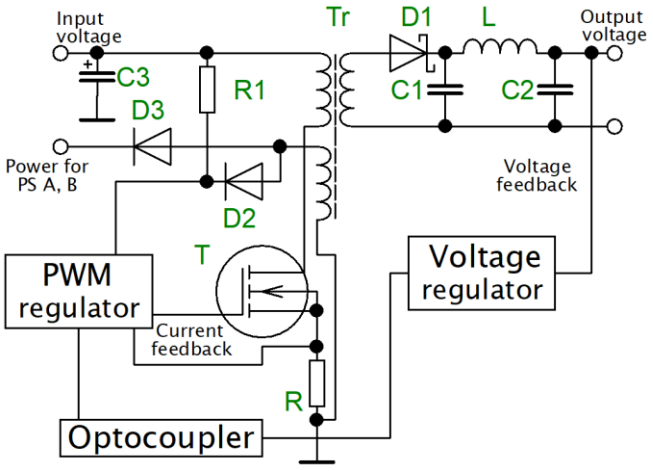


Fig. 3. Principal schema of power supply

V. Design sine wave converter

The sine converter consists of a controlling microcontroller from the STM family. This microcontroller internally generates the sine function using a lookup table and, depending on the time, writes a specific value from the lookup table to the register for the PWM output cycle. The PWM output then controls the 2ED2103 MOS-FET drivers [4]. The MOS-FET drivers further

excite the transistors in the H-bridge, the output of which is an LC filter. This LC filter was designed so that its resonance frequency lies in the logarithmic half (3) between the output frequency of the sine converter and the PWM frequency, which is 30 kHz.

$$f_r = \sqrt{f_{\sin,\max} \cdot f_{PWM}} \quad (3)$$

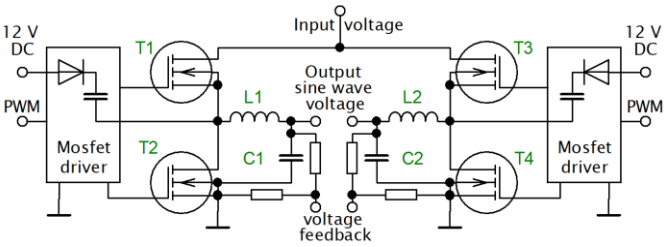


Fig. 4. Principal schema of sine wave converter

A. *Mosfet driver*

The used driver, the 2ED2103, has several key features that led to its selection. The driver has a built-in deadtime generator, which is the time delay between turning on the transistor and turning off the second transistor in the branches. Deadtime prevents the switching of both transistors at the same time, which would cause a short circuit and destroy the inverter.

Furthermore, it is important to provide a power supply for driving of the upper N-channel transistor, since a higher voltage is needed to turn it on than the supply voltage of the H-bridge. This higher voltage is produced by a charge pump consisting of an external capacitor and diode, which is part of the 2ED2103 circuit.

The last important feature is under voltage protection [4]. This kind of driver was chosen because it is the simplest, cheapest and fully sufficient for this application. There are drivers composed of discrete components or drivers that contain additional protections, but for these drivers, it is necessary to use a power transformer.

VI. TRANSFORMER CONSTRUCTION

Kapton tape was used for the electrical insulation of the transformer due to its heat resistance. The insulation on the sides, which prevents the winding from shorting, is made of PET material. To limit the skin effect, some windings were made using multiple thinner wires. To improve the coupling between the primary and secondary windings, the primary winding was split into two sections with the secondary winding inserted between them [1].

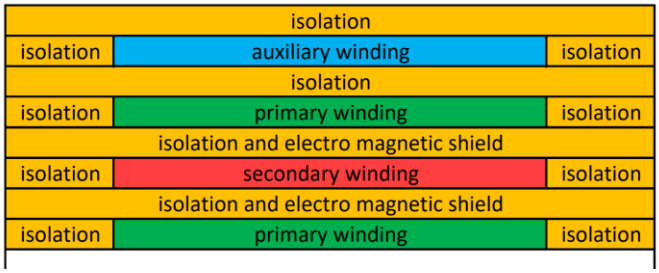


Fig. 5. Section of switching transformer

VII. PHYSICAL IMPLEMENTATION OF POWER SUPPLY UNIT

Transformers, coils and printed circuit boards have been designed and manufactured already. The PCB on the left in Fig. 6 is an EMI filter with a rectifier. The PCBs on the right show individual power supplies arranged in order from power supply A to the sine wave converter. Power supply A is at the bottom and the sine wave converter is at the top. The individual PCBs of the power supplies were designed to be modularly inserted into the Tesla Almes cabinet rails or screwed together with spacers.

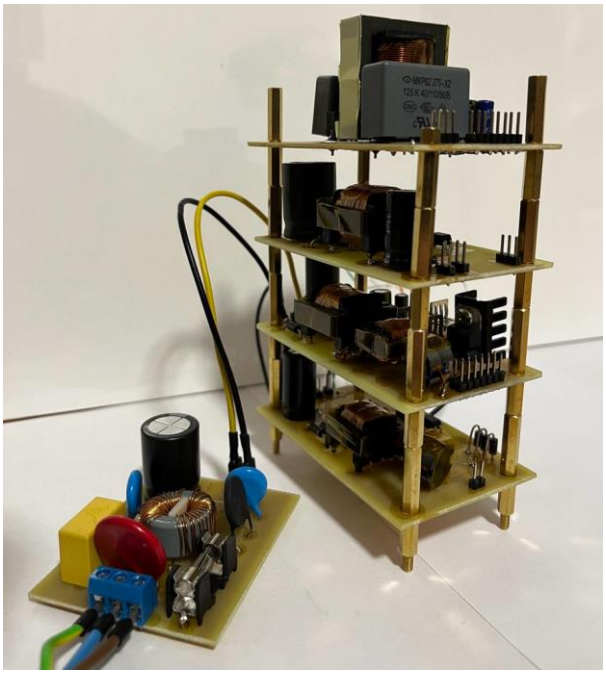


Fig. 6. Physical implementation of power supply unit

VIII. MEASUREMENT

The measurement at nominal parameters was performed only with a sinusoidal source. The waveforms of the PWM signal from the microcontroller and the output voltage can be seen in Fig. 7.



Fig. 7. Waveforms of output and excitation voltages of sine wave converter

The output of each branch of the H-bridge of the sine converter was measured using a different channel of the oscilloscope. These two signals were subsequently subtracted from each other to obtain the voltage waveform that will be present on the load connected to the H bridge. The last voltage waveform is the driving PWM signal from the microcontroller, which generates the signal with alternating polarity based on an equation that uses a sine function and instantaneous time.

Originally, the hardware was designed to allow for feedback with regulation to achieve a sinusoidal waveform for the current at various loads. However, during the measurement, it was found that the source is hard enough and therefore there is no need to regulate the shape of the sinusoidal output.

TABLE I. Measured value for sine Wave converter

V_C (V)	I_C (A)	P_C (W)	V_{ef} (V)	I_{ef} (A)	P_{sin} (W)	η (%)
116	0.5	58	76.7	0.72	55.15	95

The input and output power have been measured for the sine converter. Measured values are registered in Table 1. An efficiency of 95% has been calculated from these measurements. However, this efficiency does not take into account the power consumption required to power the exciters, which is only about 36 mW.

IX. Conclusion

The power supply unit is still under development. Currently, the printed circuit boards are populated and the transformers are wound. The efficiency of the sinusoidal converter was measured at a nominal load, and the output voltage waveform was also measured. Power supply A was tested with a load of four watts for a short period. During the operation of power supply A, high-frequency whistling of the transformer and cycling of the power supply were detected when it is not loaded. These issues must be addressed in the future, and sources B and C must be put into operation. Then, it will be necessary to measure the parameters of each source, such as voltage ripple, efficiency, and waveforms of the flyback converter.

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-23-8430).

References

[1] PATOČKA, Miroslav. Magnetické jevy a obvody ve výkonové elektronice, měřicí technice a silnoproudé elektrotechnice. Brno: VUTIUM, 2011. ISBN 978-80-214-4003-6.

[2] PATOČKA, Miroslav. *Vybrané statě z výkonové elektroniky*. 3rd edition. Brno: Zdeněk Novotný, 2004. ISBN 80-214-2742-6.

[3] Datasheet of UCCx8C4x BiCmos Low-Power Current-Mode PWM Controller. [online] [accessed 2021-05-10]. Available from: https://www.ti.com/lit/ds/symlink/ucc28c45.pdf?ts=1620488601131&ref_url=https%253A%252F%252Fwww.google.com%252F

[4] INFINEON. 2ED2103S06F Datasheet. [online]. v02.03. May 2019 [cit. 2023-03-10]. Available at: https://www.infineon.com/dgdl/Infineon-2ED2103S06F-DataSheet-v02_03-EN.pdf?fileId=5546d462766a0c1701766af58dde0551

Evaluation of the Possibility of Adapting the Public Lighting Network in Brno for Electric Vehicles Charging

David Hýsek, Martin Paar,
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
xhysek05@vutbr.cz, paar@vutbr.cz

***Abstract*— The article deals with the use of the existing public lighting network for the development of charging infrastructure. This synergy falls within the concept of the smart city strategy. The basic motivation for exploiting the synergy between the charging station network and public lighting is described. The projects already implemented in Prague are also presented and compared with the rest of Europe. The current state of public lighting in Brno is then presented, followed by an evaluation of the possibility of connecting charging points. Finally, three options for future synergies are identified.**

Keywords*— *Smart city, charging station, public lighting, Brno

I. INTRODUCTION

Smart city is a global term that characterizes the concept of strategic management of a city to improve the quality of life of its inhabitants and achieve social and economic goals using modern technologies. The smart city concept is characterized by synergies between different services and activities related to the functioning of the city (logistics, energy, security, mobility, building management, etc.). Smart Cities projects, which have already been implemented in a number of cities in Western Europe, involve using existing public lighting infrastructure to introduce charging stations.

The transition of the light source of the luminaires to LED technology has resulted in a reduction of the load on the public lighting network during night time periods. A further reduction in energy consumption comes with smart luminaire intensity control technology. The study "A Microgrid Architecture for Integrating EV Charging System and Public Street Lighting" conducted in 2019 shows that due to the transition from sodium lamps to LEDs, controlling the lumen output of the lamps and optimally managing the available energy sources, there is an increase in the hosting capacity in the public lighting network [1].

These facts lead to the question of the use of power lines for public lighting for other purposes. Charging points for charging electric vehicles can be used. These can conveniently be installed on public lighting poles on housing estates and other residential units where not every electric vehicle (EV) owner will be able to charge their car in a garage or have their own parking space.

Using the public lighting network to charge electric vehicles has a number of advantages and disadvantages, some of which are listed below in TABLE I.

TABLE I. ADVANTAGES AND DISADVANTAGES OF USING THE SYNERGY OF CHARGING STATIONS AND THE PUBLIC LIGHTING NETWORK

Advantages	Disadvantages
Recharge points can be installed at any pole height	Column may be disturbed (possible new revision required)
Less space requirements around charging points	Installations are limited to the locations of the existing network
Lower cost of implementing a charge point using the existing network (possible excavation work can be combined with the restoration of the lighting network)	Restriction to light pole standards

All installations are done exclusively with Level 2 charging. This is standard charging with the phase voltage available on the low voltage grid.

The development of public charging infrastructure with lower charging capacities will be accompanied by efforts by larger cities to reduce the presence of private vehicles in city centers. This would lead to the construction of park and ride (P+R) car parks on the outskirts of cities and the subsequent development of charging infrastructure, because according to current European legislation, for every ten parking spaces built, one parking space with a charging point must also be implemented.

Prague is the most innovative city in the Czech Republic in terms of synergy between charging points and the public lighting network, and therefore the following chapter presents the strategy of Prague companies in this synergy.

II. PRAGUE'S PUBLIC LIGHTING CHARGING INSTALLATIONS

The lifetime of public lighting is 40 years, but only 20 years in the case of luminaires. Out of the total number of over

136 thousand luminaires in Prague, one third are already over forty years old. For this reason, it is expected that most of the current public lighting will not be operational by 2025. Technology of the Capital City of Prague "Technologie hlavního města Prahy" (THMP) plans a complete replacement of the poles within 10-13 years. This represents a great potential for the implementation of "smart cities" technologies, such as lighting columns with installed charging points [2] [3].

The Prague City Council expects a concentration of 100 to 140 thousand electric vehicles on Prague roads in 2030. In order to strengthen the charging infrastructure, charging points will be implemented on the territory of Prague in public lighting poles with an investment of CZK 880 million [4].

By 2026, the project "Upgrade of the public lighting network for the purpose of e-mobility charging and other synergies", implemented by THMP, plans to reconstruct 300 km of cable routes and public lighting and distribution system routes in Prague, during which up to 3,000 EV-ready lamps will be implemented. The expected savings in investment costs are estimated at CZK 500 million. The vast majority of the installations are planned outside the wider city center [5].

The charging points installed in Prague in synergy with the public lighting are implemented directly on the lighting poles, on which there are two 400 V sockets and one 230 V socket. Two electric vehicles up to a power of 22 kW can be charged at the same time on these points. PREdistribuce prefers a charging capacity of 22 kW, although most current EVs support charging capacities of 7.4 and 11 kW. For this reason, the projects are being built with a new cable network for the charging points, as the existing one would be undersized. The new cable network must be able to supply power for both the charging point and the lighting source [6]. The vast majority of EV-ready lamp installations in Prague are carried out as part of pole replacement or possibly pavement reconstruction. An example is the installation on Slezská Street in Vinohrady, where 47 LED public lighting poles were erected as part of a sidewalk reconstruction, 13 of which were EV-ready lamps [7].



Fig. 1. EV-ready lamp [7]

Figure (Fig. 1) shows the implementation of the charging point by PREdistribuce a.s. The distribution network cabling is connected directly to the lighting pole. The lighting fixture is equipped with a wallbox and a pole socket box. The pole box consists of two cabinets connected back-to-back: a junction box (part of the distribution network) and a meter box. The junction box mediates the connection of the lighting fixture and the charging station. In PREdistribuce's implementations, there is only one power cable to the luminaires for both the charging station and the luminaire. For this reason, the luminaires are switched individually. A combined LV power cable (KSK LV) with a microtube is used, which can be used for other installations related to the smart cities strategy (air quality sensors, camera systems, etc.).

The above information shows that the installations carried out by PREdistribution in the city of Prague are implemented as alternating three-phase charging with two charging points per lighting column with an output of 22 kW. The installation of these points is connected with the renewal of the existing lighting network and new cabling is in place to supply the charging points. PREdistribuce's strategy is to introduce cabling for higher power outputs than are currently planned, so that charging outputs can be increased in the future without the need for additional excavation work [8].

On the other hand, installations made in other parts of the world mostly use existing public lighting cabling and are preferred as single-phase AC single point charging with outputs of 2.3-7.4 kW. These charging powers are sufficient to charge more than half of the battery capacity within 8-10 hours.

III. Current Status of Public Lighting in the City of Brno

The maintenance and operation of public lighting in the city of Brno has long been carried out by the city company Technical Networks Brno “Technické sítě Brno” (TSB). The aim of the company is not only to provide lighting services to the public but also to indicate faults, their possible elimination, efficiency, economy and modernization of the network.

At the end of 2021, the public lighting network in Brno consisted of 41 735 lighting points. The average power consumption per light point is given as 79 W. The total annual electricity consumption for the operation of the public lighting network in 2021 was 14.1 GWh.

There are 473 meters switchboards installed in Brno to control public lighting. The vast majority (444) of the switchboards are controlled by a radio-data network and only 29 by time switches. A total of 86 switchboards have voltage regulators and a further 234 have voltage stabilization, which aims to save electricity (about 26%) in the area.

The lighting switching (on/off) switchboards are controlled by radio signal directly from the TSB control room (Fig. 2). The control room is capable of switching the switchboards with sodium lamps and controlling the switchboards with LED technology lamps by radio communication [9].

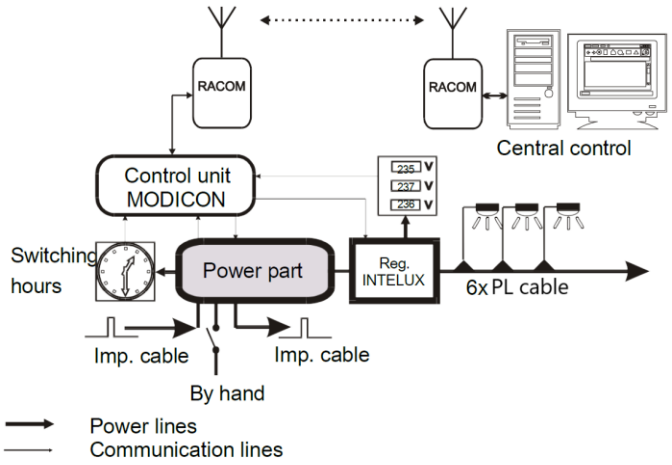


Fig. 2. Block diagram of public lighting switchboard [9]

Furthermore, as can be seen from the block diagram of the lighting switchboard, there are a maximum of 6 branches of lighting from each switchboard, each protected by a type B 32 A circuit breaker.

The public lighting distribution cables in the ground shall be provided with a protective conductor of the same cross-section as the phase conductors. Cable lines installed in the territory of the City of Brno must be at least CYKY-J 4 x 10 mm^2. Currently, the trend in reconstruction projects is to install CYKY 4 x 16 mm^2 or CYKY 4 x 25 mm^2 cable [9].

IV. Evaluation of the Connectivity of Charging Stations for EV to the Current Network in Brno

One of the biggest problems in the connectivity of charging points to the current lighting network is the legislation given by the Energy Regulatory Authority. Specifically, the C62 tariff in place for public lighting. This tariff is intended exclusively to power public lighting (illuminating public spaces) and to power the cameras of the integrated rescue system and cannot be used to power charging stations or to sell energy. TSB is also not licensed to sell electricity [10].

Possible solutions for deploying charging stations on the public lighting network through the C62 tariff are:

1) Running the Power Cable (in parallel) in Front of the Lighting Distribution Board

- The charging station must have its own electricity meter.
- Implementation limited to public lighting renovations only (cable laying)

2) Management of the Public Lighting Network Owned by the Distribution System Operator (DSO)

- DSO have a license to sell electricity and can use a different tariff than C62 for the public lightning network.
- Projects implemented in Portugal and Prague

3) Re-invoicing of Electricity Sold

- Currently, charging stations are implemented with their own electricity meter (the charging station reports the amount of energy consumed). It is therefore possible to back-bill the amount of energy consumed by all stations and to assign a different tariff to this value.

Another challenge of the synergy between charging points and the lighting network is that the network is only powered during the night period. The solutions are:

1) The Aforementioned Cable Routing in Front of the Power Distribution Box

2) Constant Power Supply to the Network.

- The problem with switching the lighting fixtures – the fixtures can be switched individually (all must contain a communication module)
- TSB currently use a radio signal to the switchboard to switch luminaires, which switches the lighting branches in bulk.

For a universal network function, a single communication technology for control and regulation of both luminaires and charging stations would be adequate.

Another problem is the limitations on the maximum power of the charging stations. This is limited by the cross section of the cable used by TSB (CYKY 4 x 16 mm2 - the current carrying capacity in the ground is 95 A) and the rating of the main Type B circuit breaker at the switchboard of 63 A and at each of the connected branches of 32 A. TSB has reserved a maximum of 63 A per switchboard with the DSO and there may not be enough spare capacity in each area.

The last issue mentioned is the location of the lighting fixtures. In order to use the public lighting network for charging electric vehicles, the luminaires must be located directly at parking spaces. Therefore, they cannot be located on the other side of the street or across the pavement (charging an EV would obstruct the pedestrian walkway).

V. Design of the Network

From the above, it is clear that public lighting networks in the Czech Republic are not suitable for charging electric vehicles. In order to exploit the synergy, technical or legislative mechanisms must be put in place.

One possible solution to the problems and constraints related to the public lighting network is to build a network for multifunctional use – smart grid (Fig. 3). This concept here represents the interconnection of the lighting network with the charging station infrastructure for the installation of photovoltaic sources. The aim is to create a network where a single line is used for all three purposes. The management of the smart grid would have to be handled by a single operator for the highest operational efficiency.

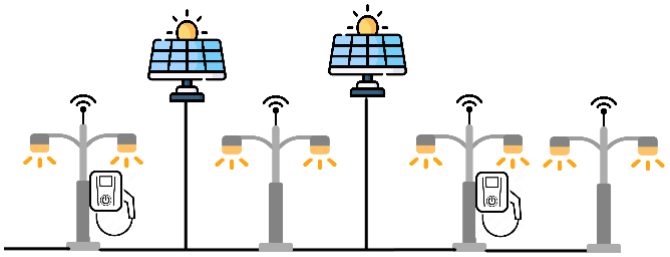


Fig. 3. Smart grid design

The photovoltaic sources would be used to smooth out peaks in demand from charging stations. Depending on the size of the installed PV capacity and the time of year, the excess electricity generated could be stored and then used to smooth out the peak switching currents of LED luminaires.

As a result of the installation of the charging stations, the grid must be permanently powered. The grid would have to be significantly oversized compared to the current situation, as the load on the grid would increase due to the charging of electric vehicles and the connection of sources.

Permanent power supply to the grid would require a different switching solution for luminaires than is currently the case. A solution is proposed to control via a suitable communication technology and switch the luminaires separately, where each luminaire would have to have its own communication module for disconnection and connection to the grid. The selected communication technology would not only control and regulate the luminaires, but also the photovoltaic sources and the charging stations for the convenience of the network manager.

VI. Conclusion

The paper evaluates the possibility of connecting charging stations to the public lighting network in Brno. To enable the evaluation, the current state of the technical and legislative framework of the public lighting network in Brno is presented, as well as the already implemented projects of synergy of the public lighting network and charging stations in Prague, as a model area for Brno. A comparison of the current trend of synergy development in Prague and Western Europe is mentioned.

Subsequently, problem and limitation points for the current public lighting network in Brno were identified. These points show that it is currently not possible to operate charging stations on the public lighting network in Brno. Only the use of lighting poles and excavation works is possible. The reason for this is the existence and use of the C62 tariff by TSB.

For the limiting points, proposals for solutions that can be implemented in the future have been identified. From these solutions, three options for the installation of charging stations on the lighting network emerge:

1. Use of the current state of the public lighting network (without the existence of the C62 tariff)
2. Reconstruction of the public lighting network
3. Building a smart grid

In addition, the concept of a smart grid, which includes charging stations, luminaires and photovoltaic sources, is presented as a possible solution. The network would therefore provide services to provide capacity for charging electric vehicles, lighting public spaces and mediating overflows to the upstream network or reducing the consumption of the network.

References

[1] BRUNO, Sergio, Giovanni GIANNOCCARO, M. SCALA a Rodio CARMINE. A Microgrid Architecture for Integrating EV Charging System *and Public Street Lighting. 2019, 1-5.: doi:10.1109*/EEEIC.2019.8783640

[2] *Renewal of public lighting in Prague to cost 2.4 billion* [Obnova veřejného osvětlení v Praze bude stát 2,4 mld.]. Smart light solution, [in Czech], 2018 [online]. Available: https://www.smartlightingsolution.cz /upload/files/sc-praha-2mld.pdf

[3] *Prague's public lighting administrator renews up to thousands of luminaires and masts annually* [Správce pražského veřejného osvětlení obnovuje až tisíce svítidel a stožárů ročně], Technologie hlavního města Prahy a.s. [Technologies of the City of Prague a.s], [in Czech], 2020 [online]. Available: https://www.thmp.cz/hlavni-strana/aktuality/spravce -verejneho-osvetleni-v-praze-rocne-obnovi-az-tisice-svitidel-a-stozaru-na-podnety-mestskych-casti-prisviti-treba-detska-hriste

[4] *In Prague, electric cars will be charged directly from the public lighting network* [V Praze se budou elektromobily nabíjet přímo ze sítě veřejného osvětlen]. Magistrát hlavního města Prahy [Prague City Council], [in Czech], 2021 [online]. Available: https://www.praha.eu/jnp/cz/o_meste/ zivot_v_praze /sluzby/v_praze_se_elektromobily_budou_nabijet$5465-export.htm

[5] *Annual Report* [Výroční zpráva]. Technologie hlavního města Prahy a.s., [Technologies of the City of Prague a.s], [in Czech], 2021 [online]. Available: https://thmp.cz/images/ THMP_VZ_2021.pdf

[6] *Prague installs smart lamps, recharges cars or monitors smog* [Praha instaluje chytré lampy, dobíjí auta nebo monitoruje smog], MAFRA, a. s., [in Czech], 2018 [online]. Available: https://www.idnes.cz/technet/ technika/chytre-lampy-praha-karlin.A180502_151509_hardware_vs

[7] *Thirteen lighting poles were erected in Vinohrady* [Na Vinohradech vyrostlo třináct stožárů veřejného osvětlení]. Magistrát hlavního města Prahy [Prague City Council], [in Czech], 2020 [online]. Available: https://portalzp.praha.eu/jnp/cz/tiskove_zpravy_z_mesta/na_ vinohradech_vzniklo_trinact_stozaru.html

[8] *Will residents of housing estates have a place to charge their electric cars?* [Budou mít obyvatelé sídlišť kde nabíjet elektromobily?], Spotify, [in Czech], 2022 [online]. Available: https://open.spotify.com/episode/ 1WPrcljWQ9Z8g01bmQz6Eo? si=588343ad2e6a4973&nd=1

[9] *Municipal standards and methodologies*.[Městské standardy a metodiky] Statutární město Brno [The Statutory City of Brno], [in Czech], 2022 [online]. Available: https://www.brno.cz/ w/oddeleni-majetkopravni-a-ekonomicke

[10] *ENERGY REGULATORY BULLETIN 11/2021* [ENERGETICKÝ REGULAČNÍ VĚSTNÍK 11/2021], Energetický regulační úřad [Energy Regulatory Office], [in Czech], 2023 [online]. Available: https://www.eru.cz/energeticky-regulacni-vestnik-11-2021

WORK NOW ON THE VEHICLES OF TOMORROW

GROW YOUR CAREER WITH US. JOIN NOW!

Garrett Motion is a cutting-edge automotive technology company that has been pioneering turbocharging for more than 65 years. Now, Garrett is leading exciting breakthroughs in electric & hybrid solutions and vehicle software, contributing to a cleaner, safer, and smarter world on the move.

Garrett's R&D centers in the Czech Republic are driving innovation that is shaping the future of mobility. No matter the experience or background, talent will always find a voice at Garrett.

Join us and make your mark!



Garrett
ADVANCING MOTION

Garrett
ADVANCING MOTION

Už více než 65 let představuje značka Garrett to nejlepší z inovativních technologií přeplňování spalovacích motorů. Ať už v oblasti motorsportu, osobních nebo nákladních vozidel, Garrett dal světu řadu průkopnických technologických řešení pro zvyšování výkonu a současně snižování emisí při dosažení optimální spotřeby paliva.
Dnes se Garrett zaměřuje také na elektrifikované systémy přeplňování a softwarová řešení pro konektivitu vozidel. V portfoliu výrobků najdeme inovativní řešení pro širokou škálu aplikací od benzínových přes naftové, plynové a hybridní pohonné jednotky včetně plně elektrických pohonů založených na použití vodíkových palivových článků

Ve vývojovém centru v Brně nabízíme dlouhodobé brigády i pozice na hlavní pracovní poměr, exkurze a workshopy.

Koho hledáme?

- Studenty a budoucí absolventy oboru elektrotechniky a mechatroniky
- Návrháře obvodů a DPS
- Testovací inženýry
- Projektové inženýry, kteří řídí vývoj aktuátorů a senzorů u dodavatelů
- Návrháře elektrického motoru a výkonové elektroniky
- Programátory embedded systémů

Co v Garrettu najdete?

- Šanci podílet se na unikátních vývojových projektech, které zahrnují elektromechanické soustavy, elektrické stroje, řídicí elektroniku a software, a to vše pro elektrifikované přeplňování spalovacích motorů nebo elektrická vozidla
- Spolupráci se špičkovými odborníky v ČR i zahraničí, od kterých se můžete učit
- Projekty s reálným výsledkem a nejmodernější technologie
- Možnost poznat vývojová centra i v zahraničí
- Vedoucí, kteří Vám dají šanci růst
- Přátelské pracovní prostředí
- Rozvoj svého potenciálu

Kontakt:
Jarmila Uherová
jarmila.uherova@garrettmotion.com
Tuřanka 100, 627 00 Brno – Slatina
www.garrettmotion.com/careers

Garrett Internal

Secure Transfer of Digital Messages

Matyáš Horký
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
211790@vut.cz

Adrián Tomašov
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
tomasov@vut.cz

***Abstract*—E-mail technology was created more than forty years ago, back in the time of the ARPANET. Over the years, there have been additions and updates to the Request for Comments, but neither the format nor the protocol have changed in a significant way. The paper presents a proposal for a new message exchange protocol addressing some of e-mail's issues, namely (the lack of) confidentiality, secrecy, and reliability.**

***Index Terms*—digital communication, e-mail, certificates, federation, cryptography**

I. Introduction

Nearly every person, adult or child, who has been using internet for at least a bit, has one e-mail account at minimum registered to some of many e-mail providers. E-mail inbox is a predisposition to having an user account on nearly every other platform on the internet, be it social media, instant messaging solutions, streaming services, or remote storage, to name a few.

This first chapter describes the e-mail protocol stack and its technical limitations, which were the main starting point of this research, and briefly describes other protocols that are currently used for online communication. Second and third chapters offer a description of a new file format and message distribution protocol that aims to solve some pain points of e-mail.

A. The e-mail message format

E-mail as we know it has been properly defined in 1982 in RFC (*Request for Comments*) 822 [1], over forty years ago. While there have been discussions about message distribution since the first computers were connected together – A Mail Box Protocol was proposed in 1971 [2], FTP (*File Transfer Protocol*) was partially supporting text message distribution [3], and limited mail-like functionality existed since the 1960s [4] –, this is the year the format we are still using today was outlined and defined.

The file format itself does not have an officially associated file extension to go along with. Both its name and file extension depend on the software that deals with it: Microsoft tends to name the e-mail files *.eml* (from *Electronic Mail*), but other vendors use *.msg* (*Message*), *.mbox* (*Mailbox*), *.imf* (*Internet Message Format*, which is also a name of a later RFC deprecating 822 [5]) or some other extension.

An important part of today's e-mail structure is something called MIME. *Multipurpose Internet Mail Extensions* is a set of multiple RFCs allowing to send anything, not just ASCII 7-bit text, such as Unicode text (commonly in a form of UTF-8), rich text (HTML-formatted text most of the time) and for adding binary attachments to the messages [6], [7].

The 7-bit aspect is more or less baked into both the message format and SMTP (*Simple Mail Transfer Protocol*) and POP3 (*Post Office Protocol, version 3*)/IMAP (*Internet Message Access Protocol*) protocols, which are used to transfer the messages into and from mailservers [8]–[10].

All those decades ago, messages were being transferred in plaintext over the network, hopping from server to server, being read over and over, until they reached a destination. This forced the e-mail protocol stack to be quite strict on the contents of the files, as there were mailservers in the wild without 8-bit support. As they would not be able to receive and transmit such messages, the stack disallows bytes with values greater than 127.

To avoid that, MIME defines several headers that allow the sending software to transform 8-bit data into a valid 7-bit stream. So-called *Content-Type* can be defined on either mail-global or block-local level, and tell the reading device (an e-mail client) how to interpret contained data. For example, header *Content-Type: text/plain; charset=utf-8* tells the client to interpret the data as nonformatted Unicode text. However, to prevent 8-bit characters, it must be accompanied by another header, *Content-Transfer-Encoding: base64* [7]. The use of base64 encoding makes it possible to include any byte blob in the text-only file format, for the price of making the encoded data about 33% larger. Another disadvantage of base64 is that it makes it impossible for a human to read the file without using a client, but that would be a problem in 1970s and 1980s, when the e-mails were (sometimes) literally printed in their raw form and delivered on receivers' desk, which is something that is no longer a usecase [2].

B. The e-mail message transmission

The transport protocol that does most of the work is SMTP. It was published alongside the original file format, making it older than HTTP. For that reason, it has completely different status messages: where HTTP uses *200 OK*, SMTP has *250 OK*, where HTTP defines client- and server-side errors (*4xx*, *5xx*) SMTP notifies of temporary and permanent errors [8].

As already hinted, this protocol is fully text-based. The exchange can be read as a polite conversation: introduction

(using keyword *HELO*, later *EHLO*), announcement who and to whom is sending a message (*SOML FROM*, *RCPT TO*), followed by data, an empty line containing just a period character, and a keyword *QUIT*.

SMTP transfers messages over the network using forwarding. Sender's server sends the message using the same SMTP protocol to the next server and that one to the next, until the receiver's server is reached. By default, this journey is completely unencrypted and unverified. However, since the late 1990s and early 2000s, all SMTP traffic has been carried over SSL/TLS (*Secure Sockets Layer, Transport Layer Security*). The messages themselves still lack any kind of integrity protection, and while solutions exist (OpenPGP, S/MIME), they are far from usable by the general public [11].

In general, the field has not advanced in decades. The only recent changes to the email protocol stack have been made via new DNS entries, such as DKIM or DMARC, but they only address authentication and authorization, not confidentiality or reliability [12], [13].

C. State of the Art: Modern approaches to digital communiation

Current (instant) messaging protocols and solutions build on the protocols that started appearing in late 1980s and 1990s, IRC (*Internet Relay Chat*) being the most famous one, along with XMPP (*Extensible Messaging and Presense Protocol*, formerly Jabber). After the internet became widely available to people at home, and especially after the smartphone "revolution" around 2010s, dozens of various solutions have been created by technological companies.

These protocols, mostly proprietary, do feature transport security and integrity (as the usage of TLS became the default, or even a requirement), but only in recent years have they started to add security features like E2EE (*end-to-end encryption*).

Current instant messaging applications, such as Signal or Wire, feature security certainties such as E2EE, (perfect) forward secrecy, post-compromise security or (non)repudiation [14]. Most of them share one core aspect: they are provided as-a-service, with one official server to distribute messages over. In other words, they are not federated.

Federation is a technological term for allowing multiple services (either multiple instances or multiple implementations) to communicate in between themselves. Several federated protocols are used "in the wild," with more or less differing goals.

Matrix is a protocol inspired by XMPP and focuses on real-time communication [15]. It uses HTTP APIs and SMTP-like principles to transfer messages and any other kinds of data inbetween servers. It does support E2EE. It is fairly complicated protocol and the specification is easy to get wrong [16].

Another protocol, which has received a lot of attention lately, is ActivityPub. It also uses HTTP API for distributing content. However, its focus is to provide an alternative to global social media products, and does not offer the security properties that would be required from an e-mail's successor – messages are stored in plaintext and the main focus is on content distribution through the network [17].

In conclusion, no research is being performed in highly asynchronous message distribution, and the field is stale. While there are secure solutions for instant messaging, there are none for long-form messages.

D. The paper's contribution

This document proposes a message exchange protocol called Letter. It is strongly inspired by e-mail with modern, privacy-first protocol sparkles. It aims to address some of the pain points of e-mail: file format, distribution and the lack of easy to use security and privacy.

II. Protocol proposal

The Letter protocol is a binary protocol for message construction and distribution, supporting federation, built using trusted cryptographic primitives.

It tries to closely mimic the e-mail protocol stack, as it is something everyone is familiar with. By not introducing major workflow changes for account and message management, users will not need to be trained to use it.

A. Definitions

A User is a representation of an actual person or an automated system. A Client is an entity (software) that the User uses to connect to the Server. A Server is an entity (software) that connects the User to the message distribution network. A host is the hardware and software (bare metal or virtualized operating system) that runs the Server.

B. Description of the initial system

Each user can be uniquely identified by their address: *alice@athens.example.org*, *bob@brussels.example.org*. Each server can be uniquely identified by their domain name: *athens.example.org*, *brussels.example.org*. In this example, Alice has a user account on a server called Athens, and Bob has one on a server called Brussels. The server the user is registered to is refered to as their home server.

Each server has a X.509 certificate, which can be either signed by some entity in a PKI (*Public Key Infrastructure*) chain of trust, or it can be self-signed (which is supported but discouraged, see chapter II-C). At a given time, a server can only have one valid certificate. After the end of the certificate lifecycle (which can happen by either expiration or revocation), another certificate becomes valid to take its place.

The same goes for the client: they can also have only one valid certificate at any given time. Once a user account is created on a server, they are given an initial certificate with its secret key, which is transported by unspecified means to the account holder (an example of that could be downloading it from a web server via TLS-secured connection).

Once the user is in possession of a valid initial certificate, they generate a certificate signing request. They sign it using their initial certificate and send it to their home server. The server validates the request, issues a new certificate, and

marks the initial certificate as revoked, to prevent it from being used again. The client saves the issued certificate, along with its secret key, in a secure local place.

C. Trust in the system

The trust is established by extending the chain of trust of the Public Key Infrastructure. The idea is the following: the Server's Host obtains its X.509 certificate for its hostname, using any CA (*Certificate Authority*). The certificate may already be present on the system, if the host also serves a website using HTTPS. This certificate, signed by the CA, is used as a signator for Server's own X.509 certificate. This newly created certificate can be used independently of the PKI, allowing the Server software to issue certificates for any amount of time.

The Server certificate is used as a signator for Client certificates (e.g., their Certificate Signing Requests).

If the Server only has self-signed certificate, the Client must be explicitly allowed to connect to it and to trust it. Such connections are prone to Man in the Middle attacks on the first connection, and must be verified by other means.

The Servers also need to establish trust with the Clients. This is done using mTLS (*mutual TLS*), where the Client authenticates by using a private key, which is associated with a certificate signed by their home server. The Server verifies the connection and determines if it is trustworthy.

D. The file format

The Letter format is binary and follows the structure header–length–data — header–length–data — .... As all the fields have variable length, this allows for simple parsing rules while keeping the message metadata overhead reasonable. The format is Big Endian.

The smallest entity is called a Packet. It holds specific information, and can be dynamic or static. The only static Packet is Timestamp: it is exactly four bytes long and represents the UNIX timestamp. The dynamic Packets can have either a length header with a size of two bytes (meaning the data can be at most $256^2 - 1 = 65\,535$ bytes long), of four bytes ($256^4 - 1 \approx 4\,\text{GB}$), or of eight bytes ($256^8 - 1 \approx 18\,\text{EB}$).

The maximal allocation size does mean that the servers and clients will have to support such gigantic filesizes. Both the Letter protocol and the implementations set their own, much more sane, limits.

Then there is the Client Request. A request is sent from Client to Server as a first message of an exchange. The client asks the server for an information or permission: user/server cretificate request, request for a list of messages in inbox, or a permission to send a message to the server.

The third payload option is a Response. It consists of an integer in the first two bytes and a Data Packet. The integer is equivalent to one of HTTP status codes, which the Letter protocol uses to tell the Clients the result of an action without the need to parse some message.

Packets and Requests are placed in Wraps. A Wrap is a high-level object with predefined structure of mandatory and optional Packets or Requests. An equivalent of RFC 822 message is a wrap called *Plaintext Message*, identifiable by its header *p*.

```
4c5452017000000000000000bc46001066726f6d406578616d706c652e6f72675
40011616c696365406578616d706c652e6f726754000f626f62406578616d706c
652e636f6d54001363686172 6c6965406578616d706c652e636f6d53000c54657
374206d657373616765 4d0000000000000001843000a746578742f706c61696e4
40000000068656c6c6f2149001094fc37af39e543fc8d4336fb140f1b2457 64026
d6b
LTR1p188
F16from@example.org
T17alice@example.orgT15bob@example.orgT19charlie@example.org
S12Test message
M24C10text/plainD6hello!
I1694fc37af39e543fc8d4336fb140f1b24
W20230303T225803

4c5452016300000000000000077273576 4027db1
LTR1c7rsW20230304T000729

4c5452017300000000000000049019d44000000042596f7572206d657373616765
206973203733206279746573206c6f6e672c207768696368206973 2039206279
746573206d6f7265207468616e20616c6c6f7765642e
LTR1s73413D66Your message is 73 bytes long, which is 9 bytes more
than allowed.
```

Fig. 1. An example of message, request and reply: raw binary blob and explanation. Specifically, those are a plaintext message (as stored on a client, unencrypted), server certificate request and error 413 stating that the client has sent a message that is too long to be processed.

Examples of binary messages are shown in Fig. 1. Lists of Packets, Client requests and Wraps is in Appendix A.

E. Confidentiality

Once a Plaintext Message is constructed, it is symmetrically encrypted. The symmetric key is encrypted using receiver's public key and prepended before the message. The result is hashed and the hash is signed by sender's public key of the valid certificate, and the signature is prepended to the message.

After that, the message is sent directly to the receiver's server.

F. Cryptographic primitives

The Letter protocol uses ChaCha20Poly1305 to symmetrically encrypt data [18], SHA3-256 to hash values [19], RSA 4096 to asymmetrically encrypt data and RSA 4096 PSS (with MGF1 padding) to sign data [20].

ChaCha20 is a PQC (*Post-Quantum Cryptography*) streaming cipher. It has been chosen over AES for its speed and simple design, although AES-GCM would provide the same level of safety. SHA-3 is a PQC Keccak-based hash function. It has been chosen over Merkle-Damgård-based SHA-2, although both schemes are considered safe; SHA-2 shares the same design with broken SHA-1, thus the alternative family was chosen. RSA is quantum-breakable scheme used for encryption and digital signatures. It has been chosen in its strongest form for the lack of PQC alternatives supported by X.509 and TLS. Once such replacement appears, it should be a priority for the next versions of the protocol.

III. The protocol implementation

An implementation of the protocol has been created. It acts as a proof of concept and as a reference implementation for the Letter protocol. It is written in Python 3.9 and is split into three modules: *letterlib*, *letterserver* and *letterclient*.

letterlib consists of several submodules: *crypto.sym* and *crypto.asym* (for symmetric and asymmetric encryption) and *file* (containing parsers and (de)serializers for the binary data). It has been used to generate all the examples in this paper.

letterserver and *letterclient* are example implementations written using Python's *asyncio* streams.

The implementation proves that the design of the protocol can act as an e-mail protocol stack nearly transparently, while allowing users to communicate in a private and secure manner, without the need to set up new software or learn new steps.

IV. Conclusion

The Letter protocol offers out-of-the-box confidential and reliable electronic mail stack while staying quite similar to e-mail protocols. It offers modern and future-proof cryptography solution while keeping the user experience really close to the workflow they are familiar with from using regular e-mail clients, making it possible replacement for the current e-mail system.

It replaces the text-based MIME extension with binary format and nontransparent SMTP with client-guaranteed transmission certainty. It uses modern TLS and PKI to establish trust and is easy to deploy with little to no requirements on the host system.

While still breakable by theoretical quantum computers (by partially relying on the RSA for encryption and signatures), it provides a minimal and testable protocol secure against current attacks performable using classical computers.

There are many ways for follow-up research. One of them is replacing RSA with a new scheme (or two different schemes for encryption and signatures) to make the protocol post-quantum secure. Another is the text content itself, for which a new format could be defined to support verifiable citations. A possible extension could be a system to synchronize keys and messages between client devices or a library to convert between Letter file format and RFC 822 format.

Appendix

TABLE I
Packet and Wrap types

name	identifier	name	identifier
From	*F*	Message ID	*I*
To	*T*	Reply-to Message ID	*J*
Timestamp	*W*	Subject	*S*
Name	*N*	Content Type	*C*
Bytes	*B*	Data	*D*
Message	*M*	Attachment	*A*
Extra	*X*		
Plaintext message	*p*	Request	*c*
Encrypted message	*e*	Status	*s*

Alice	Brussels
cert(Alice), sk(Alice)	*cert(Brussels), sk(Brussels)*
message	*cert(Bob)*
user certificate request →	
	← *cert(Bob)*
create $e :=$ encrypt$(message)$	
compute $h :=$ hash(e)	
have h signed by home server	
message submit request →	
	verify connection
	verify addressee
	← *CONTINUE*
full encrypted message →	
	verify, save
	← *OK*

Fig. 2. Message submission dialogue between a client and a server. Alice symmetrically encrypts her message. She hashes it and makes Athenes sign this hash. Then she joins the encrypted message, the encrypted key and Athenes' signature and sends it to Brussels, where it is stored until Bob picks it up and decrypts it.

References

[1] D. D. Crocker, "Standard for the format of ARPA internet text messages," RFC 822, Aug. 1982.
[2] R. W. Watson, "A Mail Box protocol," RFC 196, Jul. 1971.
[3] A. K. Bhushan, "Comments on the File Transfer Protocol," RFC 385, Aug. 1972.
[4] T. V. Vleck, "Electronic mail and text messaging in CTSS, 1965–1973," *IEEE Annals of the History of Computing*, vol. 34, no. 1, pp. 4–6, Jan.-Mar. 2012.
[5] P. Resnick, "Internet Message Format," RFC 2822, Apr. 2001.
[6] N. Freed and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) part one: Format of the internet message bodies," RFC 2045, Nov. 1996.
[7] K. Moore, "MIME (Multipurpose Internet Mail Extensions) part three: Message header extensions for non-ASCII text," RFC 2047, Nov. 1996.
[8] J. B. Postel, "Simple Mail Transfer Protocol," RFC 821, Aug. 1982.
[9] J. K. Reynolds, "Post Office Protocol," RFC 918, Oct. 1984.
[10] M. Crispin, "Interactive Mail Access Protocol – version 2," RFC 1064, Jul. 1988.
[11] M. Green, "What's the matter with PGP?" Aug. 2014. [Online]. Available: https://blog.cryptographyengineering.com/2014/08/13/whats-matter-with-pgp
[12] D. Crocker, T. Hansen, and M. S. Kucherawy, "DomainKeys Identified Mail (DKIM) signatures," RFC 6376, Sep. 2011.
[13] M. S. Kucherawy and E. Zwicky, "Domain-based Message Authentication, Reporting, and Conformance (DMARC)," RFC 7489, Mar. 2015.
[14] T. Perrin and M. Marlinspike. The Double Ratchet algorithm. [Online]. Available: https://signal.org/docs/specifications/doubleratchet
[15] (2023) Matrix specification. [Online]. Available: https://spec.matrix.org
[16] M. R. Albrecht, S. Celi, B. Dowling, and D. Jones, "Practically-exploitable cryptographic vulnerabilities in Matrix," unpublished. [Online]. Available: https://nebuchadnezzar-megolm.github.io
[17] C. Lemmer-Webber, J. Tallon, E. Shepherd, A. Guy, and E. Prodromou. ActivityPub. [Online]. Available: https://www.w3.org/TR/activitypub
[18] Y. Nir and A. Langley, "ChaCha20 and Poly1305 for IETF protocols," RFC 7539, May 2015.
[19] M. J. Dworkin, "SHA-3 standard: Permutation-based hash and extendable-output functions," Aug. 2015.
[20] K. M. Moriarty, B. Kalinski, J. Jonsson, and A. Rusch, "PKCS #1: RSA cryptography specifications version 2.2," RFC 8017, Nov. 2016.

Secure smart electrical socket for power metering

Vladyslav Shapoval
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
236349@vutbr.cz

Petr Musil
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
musilp@vutbr.cz

***Abstract*—In this paper is described process of development of secure intelligent electrical socket for power consumption metering and remote control (On/Off) of the connected device via radio channel. Also this paper provides short overview of the current problems with cybersecurity in IoT industry and short overview of the most common wireless technologies for smart metering, which were considered for the final product. At the end is demonstrated the current state of the smart socket development process at the time of publication as well as the first test prototype of the final device**

***Index Terms*—Smart socket, electrical metering, Thread, ZigBee, Z-wave, security, IoT**

I. Introduction

With the development of information technology, IoT devices designed to automate various processes of human life are becoming more widespread. At the same time they create new cybersecurity threats. According to SAM Seamless Network report [1], more than 1 billion attacks on IoT infrastructure took place in 2021. This pose greate risks for the end users and those risks creates demand for IoT devices resilient to most common cyberattacks.

In order to develop and implement such a device, in this paper was conducted a research of the current state of cybersecurity in IoT industry to define main requirements for the future smart socket. Also was conducted a short overview of the most commonly used wireless technologies, namely ZigBee, Thread and Z-wave in order to decide which one complies to the defined requirements and will be most practical in usage.

At the end are presented obtained intermediate results, the current state of development and is shown the created prototypes and describe plans for the future work.

II. General IoT security overview

Number of researches and articles were written on IoT security topic. Despite different approaches and classifications of vulnerabilities and attacks, a constant set of security problems is reoccurring in each of those works.

A. Inadequate authentication and hardening

Many manufactures neglect authentication, use unsecure versions of communication protocols or implementing authentication protocols with flaws, caused by bad programming techniques, such as storing preshared secret keys in plain text in flash memory, using weak default credential information , or sending keys unencrypted via wireless network. Often encryption or basic PIN protection is not used at all [2]–[4], [6].

B. Insufficient access control

Majority of IoT devices in use do not force a password of sufficient complexity and do not force end-users to change default credentials [2]. Also, due to lack of computational power, the majority of devices can not segregate access to processes, memory and sources on multiple users. Additionally in many use cases different smart meters are installed in unpopulated areas without any surveillance. This implies that in many cases anyone can get access to the device, which lead to device corruption and makes possible many of the attacks on the infrastructure, such as **False data injection, Firmware modification attack, Side-channel attack, Sybil, Sinkhole and Blackhole attacks** [8].

C. Active network attacks vulnerabilities

Eavesdropping is the simplest kind of those type of attacks. But it still has high chance on success. Some more advanced techniques like **Man in the Middle** (MitM) have some restrictions imposed on an attacker, because he needs to be in close proximity to the device under attack. Despite this it can still be successful [4]. Also the vulnerability to the MitM attack increases, if device can communicate with cloud or application for remote control over internet.

D. Denial of Service attacks

Denial of Service attack can be done by physically removing device from the network (device capture) or it's tampering and destruction; by jamming of communication channels; by exhaustion of recourse's of the device (memory, CPU time, disk space, modifying node configuration). One of the widely common Dos attack vectors is an attack on power consumption - **Battery draining attack** [2]. But attacks on power consumption does not limited by battery-powered devices only. Attacker can damage power cord, tamper battery, destroy power source, etc. [5].

E. Lack of secure update mechanism

Often manufacturers do not have automatic patch-update mechanisms for firmware or these patches are released quite rarely. If patch-update mechanisms are present then there

is may be no validation or signing of firmware updates or software updates without server authentication and file trust verification or no secure boot mechanisms. [5]

F. Application over-privilege

Application over-privilege is an often overlooked potential threat. Some vendors allows to install third-party applications for smartphones or PC to control smart devices, but this also creates additional risk because those apps can access data or devices they are not supposed to. For example, an app on Android with Bluetooth permission can access any Bluetooth device in the vicinity [3].

G. Problems, specific for smart metering at homes

In work [7] were described and given examples of some particular problems, specific for smart meters and smart plugs. One of them is a customer safety problem. Smart meters are used on critical communications such as electricity, gas, water, etc. Unintentional malfunction or cyber-attack can cause electrical surges and over voltage, which can damage and burn wiring and destroy appliances and electronics, cause fire and even explosion in case of gas leakage.

Another problem is unintentional privacy invasion. For example, by power consumption information attacker can deduce if there is some one at home right now or not, how many appliances installed at home, hours of sleep of the owners, etc. This issue was pointed out in [4] and [7].

III. Wireless technologies overview

In this section will be overviewed security mechanism of several most widely spread wireless technologies used in smart metering, namely ZigBee, Thread and Z-wave.

A. ZigBee

ZigBee defines 8 different security levels (from 0 to 7) available to the NWK and APS Layer. Only levels form 4 up to 7 provides both security and integrity with Message Integrity Code (MIC). For encryption is used AES-128. ZigBee introduces two types of keys: Network key for broadcast (the same for every node in the network) and the Link key (link key is unique for each par of nodes). For key establishment, standard relies on Symmetric-Key Key Establishment protocol (SKKE), preinstalled keys or install codes. In centralized networks Certificate-Based Key Establishment protocol (CBKE) introduced, which use Elliptic Curve Qu Vanstone (ECQV) implicit certificates, generated by Certicom certification authority. Authors of [12] and [17] pointed out some flaws in the protocol, which are inherited from 802.15.4 standard. The weakest point of the ZigBee is key management in the distributed security model and new device authentication since standard does not relay on asymmetric cryptography. In some use cases of SKKE if new device in the network do not posses preinstalled network key or master key, those keys can be transmitted unencrypted via radio, which creates a window for vulnerability. This means, that for maximal safety should be established centralized network with highest 7 security level and key management with install codes or CBKE2 protocol for key establishment. Detailed description and research of the ZigBee security provided by [9]–[12] and [17].

B. Z-wave

Z-wave provides authentication, encryption and data integrity check in distinct security layer. This layer has two security classes S0 for backward compatibility with older devices and S2 with stronger security [13]. Z-wave S2 protocol uses strong cryptological primitives for encryption, authentication and integrity control. Those primitives are Elliptic Curve Diffie-Hellman (ECDH) scheme, AES-128-CMAC and Pre-Agreed Nonces for integrity check. In [12] authors concluded, that no serious vulnerabilities were found in the protocol, but in the [17] was listed several works, which described practical attacks against Z-wave such as replay attack and Same Nonce attack and in [15] authors described procedure to force Z-wave device to use unsecure Z-Wave S0 protocol by sending unauthorized commands.

C. Thread

Analysis and description of the protocol in works [12] and [17] suggests that despite some potential flaws, inherited from IEEE 802.15.4 standard, Thread is proven quite resilient against the majority of the known attacks, since Thread mitigates those flaws by relaying on the end-to-end encryption on application layers. Also Thread has some advantages that are not obvious at first glance. Since devices in Thread use IPv6 addressing and rely mainly on security mechanisms at the OSI transport and application layer, the Thread node can establish a secure connection with devices located outside the PAN network. For example, such a device can be a cloud storage. And also there is no need for specialized gateways to parse, decrypt and reconstruct messages in order to transmit them outside the PAN. Another advantage is the use of already standardized and well-proven protocols TLS and DTLS. In works [12]–[14] and [17] was provided detailed description and analysis of protocol.

IV. Smart Socket development

A. General requirements

- The smart socket should be used in households or offices.
- To prevent battery draining or power source tampering attacks, the smart socket must be powered from the electrical grid of the Czech Republic. Connection to other devices should be provided through a standard wall outlet type E or F.
- The smart socket must be able to measure the parameters of the electrical grid (current, voltage, frequency, etc.) and send the measurement results via radio to the central home area network (HAN) controller.
- The smart socket must be able to connect / disconnect all other devices connected through it to the electrical grid on command form HAN-controller via a radio channel.
- The smart socket wireless protocol must provide node authentication, data security and integrity.

- The smart socket must be resistant against the firmware modification attacks.
- The smart socket must provide a secured way for firmware update.
- The smart socket must comply with electrical safety standards.

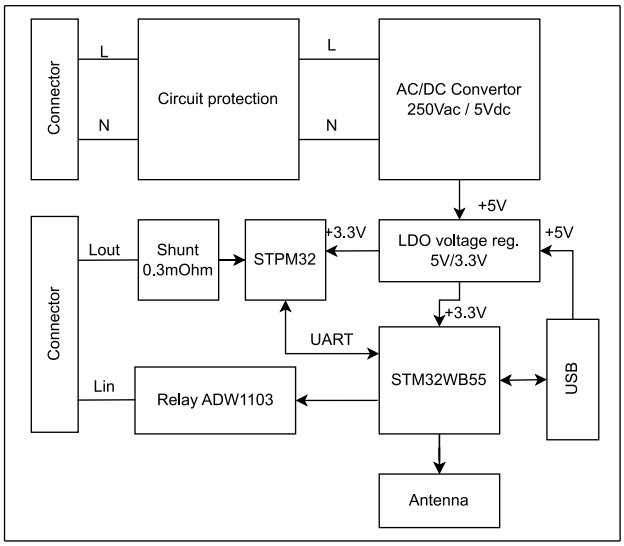


Fig. 1. Block scheme of the prototype.

B. Choice of the controller

to comply with requirements for the smart socket, in particular requirements of secure firmware update and resistance against the firmware modification attacks, as a MCU of the future device was chosen STM32WB55xx family controller. The biggest benefit of STM32WB is it's embedded memory protection capabilities and security features:

- Cortex-M0+ security - STM32WBxx controllers provide separation of the wireless stack and user application. Flash memory and SRAM are divided on two regions. Regions started with lower addresses is non-secure and used by Cortex-M4 core and user applications. Regions started with upper addresses is secure and used to store wireless stack firmware, provided by STMicroelectronics.
- Readout memory protection - provides protection of the memory and boot region from unauthenticated modification or read via debuggers interfaces. Readout Level 2 protection is permanent and can not be removed.
- Root Secure Services (RSS) - provides secure firmware update. The RSS is a protected firmware, provided by STMicroelectronics during microcontroller manufacturing. It executed by Cortex-M0+ core and used to securely install and update the wireless stack, provide integrity check and authentication of the firmware. RSS code can not be accessed by Cortex-M4 or debug port. when no wireless stack installed RSS is active by default. Wireless stack and RSS updates are provided by STMicroelectronics encrypted by AES and signed by ECDSA algorithms. Public key of the STMicroelectronics is stored in secure memory region and can not be accessed by user code or debug interface. RSS also provides Cryptographic Key storage, which allows user application store and use AES encryption keys without direct access to them

C. Choice of the wireless protocol

All the protocols discussed above provide an acceptable level of network security. However, the security of the protocol was not the only factor in the choice. According to [16], the geometric dimensions and layout of the PCB board have a very strong effect on the power radiated by the antenna. The general rule is that the lower the resonant frequency of the built-in antenna, the larger the PCB board is required to maximize the radiated power. Therefore, it was decided to stop at a frequency of 2.4GHz instead of 868MHz, otherwise the final smart socket could not meet the size requirements. This is the main reason why it was decided to refrain from using the Z-wave protocol since it is sub-Ghz only.

The second factor was the implementation of the wireless stack. The selected controller STM32WB55xx supports both technologies Thread and ZigBee because they are both based on the 802.15.4 standard. However, experiments shows, that STMicroelectronic's implementation of the OpenThread wireless stack at the time of publication of this paper does not support several key features - most notably boarder router and external commissioner roles for nodes. Without those, OpenThread loses its important advantage over ZigBee - the possibility of seamless routing on the Internet and the use of key infrastructure in conjunction with the TLS and DTLS protocols. In its current implementation on the selected controller, Thread protocol has no advantage over ZigBee with the right security settings but introduces redundant complexity to the system.

D. Intermediate results

For testing purposes and proof of concept two prototypes were designed and manufactured. Block scheme of the prototype is shown on the figure 1 Prototype is shown on figure 2. For functionality testing was created basic firmware, which implements simplest On/Off ZigBee cluster. During testing were found numeral design flaws, namely:

- Decoupling capacitors were chosen too big. Because of that after removing power source, on decoupling capacitors remain some charge from 0.3 - 0.4V. This voltage is too small for the controller to operate but it is enough to trigger brown-out detector, which prevents controller to reboot for the some time;
- Low-dropout regulator is used in the scheme to convert 5V from USB or AC/DC converter to 3.3V. Output current of this LDO is not enough to switch relay, which was connected directly to the pin of the MCU. A separate circuit with a driving transistor is required for relay control;

- The USB connector was chosen poorly, the selected one has a weak fastening, if user does not carefully connect / disconnect the cable, he can easily break it off;
- The device measures the voltage relative to the ground plane, so in case of measurement of voltage in mains L conductor, it will be measured relative to the ground plain in the device and not relative the neutral conductor N;
- The pins in the SWD/JTAG debugger connector were mixed up;
- Insufficient indication. Additional LEDs required for monitoring error and standby states and the state of the STPM32 measurement unit.

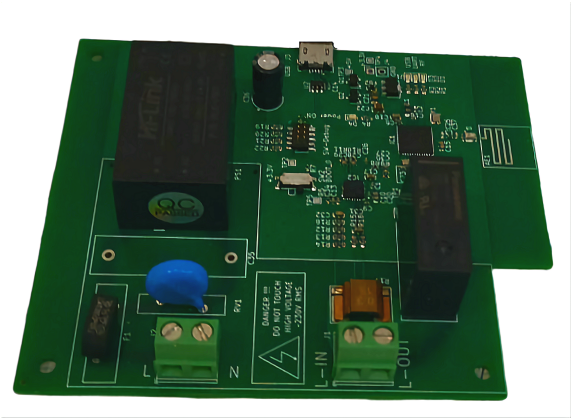


Fig. 2. First manufactured prototype.

V. Conclusion

The aim of this work was to create a device that could perform the functions of a smart socket, while meeting the latest requirements in the field of information security. To determine the requirements for the device being developed, a study was conducted of the main threats to information security that the device may encounter during operation. A comparative study of wireless technologies for smart metering was also carried out, as a result of which the protocol ZigBee was chosen. Several prototypes were made and a basic firmware for them was created. During testing, numerous design flaws were identified. At the time of publication of this article, work is underway on a new prototype, and a full-fledged firmware.

Acknowledgment

I would like to thank the advisor of my thesis, Ing. et Ing. Petr Musil for patience, professionalism, consultation and guidance during the work.

References

[1] *Landscape Security IoT 2021* [online]. New York: SAM Seamless Network, 2021 [cit. 2022-11-06]. Available from: https://securingsam.com/2021-iot-security-landscape

[2] N. Neshenko, E. Bouharb, J. Crichigno, G. Kaddoum and N. Ghani, *Demystifying IoT Security: An Exhaustive Survey on IoT Vulnerabilities and a First Empirical Look on Internet-Scale IoT Exploitations.* IEEE Communications Surveys & Tutorials. 2019, 21(3), 2702-2733. ISSN 1553-877X. Available from: doi:10.1109/COMST.2019.2910750

[3] Zhang, Nan & Demetriou, Soteris & mi, Xianghang & Diao, Wenrui & Yuan, Kan & Zong, Peiyuan & Qian, Feng & Wang, Xiaofeng & Chen, Kai & Tian, Yuan & Gunter, Carl & Zhang, Kehuan & Tague, Patrick & Lin, Yue-Hsun. *Understanding IoT Security Through the Data Crystal Ball: Where We Are Now and Where We Are Going to Be.* 2017, Available from: arXiv:1703.09809.

[4] I. Stellios, P. Kotzanikolaou, M. Psarakis, C. Alcaraz and J. Lopez, *A Survey of IoT-Enabled Cyberattacks: Assessing Attack Paths to Critical Infrastructures and Services.* IEEE Communications Surveys & Tutorials. 2018, 20(4), 3453-3495. ISSN 1553-877X. Available from: doi:10.1109/COMST.2018.2855563

[5] *Baseline security recommendations for IoT in the context of critical information infrastructures.* [online]. LU: European Union Agency for Network and Information Security, 2015.. ISBN: 978-92-9204-236-3, doi: 10.2824/03228. [cit. 2022-11-19]. Available from: https://www.enisa.europa.eu/publications/baseline-security-recommendations-for-iot

[6] S. Kumar, H. Kumar a G. R. Gunnam, *Security Integrity of Data Collection from Smart Electric Meter under a Cyber Attack.* 2019 2nd International Conference on Data Intelligence and Security (ICDIS). IEEE, 2019, 2019, 9-13. ISBN 978-1-7281-2080-5. Available from: doi:10.1109/ICDIS.2019.00009

[7] R. Marah, I. El Gabassi, S. Larioui and H. Yatimi, *Security of Smart Grid Management of Smart Meter Protection.* 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET). IEEE, 2020, 2020, 1-5. ISBN 978-1-7281-4979-0. Available from: doi:10.1109/IRASET48871.2020.9092048

[8] A. Rajan, J. Jithish and S. Sankaran, *Sybil attack in IOT: Modelling and defenses.* 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE, 2017, 2017, 2323-2327. ISBN 978-1-5090-6367-3. Available from: doi:10.1109/ICACCI.2017.8126193

[9] E. Yuksel, *Analysing ZigBee Key Establishment Protocols.* Kongens Lyngby: Technical University of Denmark, 2010, 1-39. Available from: doi:10.48550/arXiv.1205.6678

[10] V. Rudresh, *ZIGBEE SECURITY: BASICS (PART 2).* Kudelski Security Research [online]. 2017. [cit. 2022-11-26]. Available from: https://research.kudelskisecurity.com/2017/11/08/zigbee-security-basics-part-2/

[11] S. Nath, S. Aznabi, N. T. Islam, A. Faridi and W. Qarony, *Investigation and Performance Analysis of Some Implemented Features of the ZigBee Protocol and IEEE 802.15.4 Mac Specification.* International Journal of Online Engineering (iJOE). 2017, 13(01), 14-32. ISSN 1861-2121. Available from: doi:10.3991/ijoe.v13i01.5984

[12] S. Marksteiner, V. J. Exposito Jimenez, H. Valiant and H Zeiner, *An overview of wireless IoT protocol security in the smart home domain.* 2017 Internet of Things Business Models, Users, and Networks. IEEE, 2017, 2017, 1-8. ISBN 978-1-5386-3197-3. Available from: doi:10.1109/CTTE.2017.8260940

[13] I. Unwala, Z. Taqvi a J. LU, *IoT Security: ZWave and Thread.* 2018 IEEE Green Technologies Conference (GreenTech). IEEE, 2018, 2018, 176-182. ISBN 978-1-5386-5183-4. Available from: doi:10.1109/GreenTech.2018.00040

[14] I. Unwala, Z. Taqvi a J. LU, *Thread: An IoT Protocol.* 2018 IEEE Green Technologies Conference (GreenTech). IEEE, 2018, 2018, 161-167. ISBN 978-1-5386-5183-4. Available from: doi:10.1109/GreenTech.2018.00037

[15] A.Tierney, *Z-Shave. Exploiting Z-Wave downgrade attacks: BLOG: INTERNET OF THINGS.* Pen Test Partners: Security consulting and testing srvices [online]. [cit. 2022-12-04]. Available from: https://www.pentestpartners.com/security-blog/z-shave-exploiting-z-wave-downgrade-attacks/

[16] C. Newman and C. Tomlin, *Mission Impossible? How can an embedded antenna perform brilliantly within an IoT device with a very small PCB?: Designing IoT Devices using embedded antennas at sub 1GHz frequencies on very small PCBs.* Embedded World Exhibition&Conference. 2019, 1(1), 1-4.

[17] G. Kambourakis, C. Kolias, D. Geneiatakis, G. Karopoulos, G. M. Makrakis and I. Kounelis, *A State-of-the-Art Review on the Security of Mainstream IoT Wireless PAN Protocol Stacks.* Symmetry. 2020, 12(4). ISSN 2073-8994. Available from: doi:10.3390/sym12040579

Blind Issuance for Fast Keyed-Verification Anonymous Credentials

1st Kristián Klasovitý
Department of Telecommunication
Brno University of Technology
Brno, Czech Republic
xklaso00@vutbr.cz

2nd Sara Ricci
Department of Telecommunication
Brno University of Technology
Brno, Czech Republic
0000-0003-0842-4951

***Abstract*—Cryptographic credentials can be used to prove ownership of users' personal attributes, such as age, nationality, sex, or vaccine validity. Recently, a novel keyed-verification credential system designed for lightweight devices has been developed. In this paper, we propose an extension to this system that enhances the user's privacy by implementing blind issuance to the protocol. This allows for the creation of personalized credentials, that remain hidden from the issuer and anyone who would intercept the communication. To demonstrate the practicality of our extension, we tested the scheme on multiple devices and showed that the main algorithms are comparably fast as the original implementation.**

***Index Terms*—Anonymous Credentials, Blind Issuance, Two-party Computation, Homomorphic Encryption**

I. Introduction

Anonymous credentials allow access to internet-based services without necessarily needing to reveal all personal information. Specifically, users can anonymously prove the possession of required credentials and this proof cannot be linked to previous uses. This anonymization is not only towards outsiders but, if desired, also towards the service provider.

Depending on the services the information needed to be shown may vary. Anonymous credentials permit selecting which personal data to show depending on the request of the authorized service provider. This strategy is denoted as the data minimization principle, i.e., the personal data provided need to be "adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed". This principle is part of the General Data Protection Regulation (GDPR) [1]. A possible scenario could be the deployment of the European Union digital vaccine certificate. Here, the certificate holder does not need to show all his personal data such as name, date of birth, and place of residence when only needing to prove, for example, the validity of his vaccine.

A. Contribution

In this article, we propose a novel asymmetric anonymous attribute-based credentials scheme mainly designed for mobile devices such as smartphones and smartwatches. Our scheme enhances the privacy provided by the original Keyed-Verification Anonymous Credentials (KVAC) scheme [3] by adding blind issuance. This can be achieved by adding a Non-Interactive Zero-Knowledge Proof of Knowledge (NIZKPK) scheme in [2] in our `Issue` protocol. It is important to notice that, to our best knowledge, this is the first implementation complete with proof of knowledge of the aforementioned NIZKPK scheme.

In our proposal, the user is in possession of a secret key that is not shared with the issuer. Therefore, the issuer will be not able to impersonate any user. This property can also be used to protect the credentials from an attack during the issue phase as the real credentials can only be extracted by the user. Experimental results show the applicability of our scheme to the Internet of Things (IoT) environment.

Our proposal remains competitive, where `Show` and `Verify` algorithms that will be used regularly are almost unaffected with an increase of only 10% of the total computational cost. Therefore, the modified scheme still remains fast enough to be used in real-life IoT scenarios.

II. Preliminary

In this section, at first, we outline the used notation. Then we review the protocols on which our scheme is based, namely a non-interactive zero-knowledge proof of knowledge [2] and keyed-verification anonymous credentials [3].

From now on, we write $a \xleftarrow{\$} A$ when a is sampled uniformly at random from A. A secure hash function is denoted as $\mathcal{H} : \{0,1\}^* \rightarrow \{0,1\}^\kappa$, where κ is a security parameter. We describe the proof of knowledge protocols (PK) using the notation introduced by Camenisch and Stadler (CS) [4]. The protocol for proving the knowledge of discrete logarithm of c with respect to g is denoted as $\mathrm{PK}\{\alpha : c = g^\alpha\}$. Dec stands for homomorphic decryption.

A. Non-Interactive Zero-Knowledge Proof of Knowledge

The NIZKPK [2] allows two parties (a signer and a issuer) to jointly compute a Boneh-Boyen signature $\sigma = g^{1/(sk_m+m)}$ of a signer's private message m and the issuer's secret key sk_m. In our case, the signer's message is his private key sk_i The NIZKPK [5] scheme is divided into two parts: `Setup` and `2-party Signature` protocol as shown in Algorithms 1 and 2. Algorithm 1 generates the system parameters needed to run the signature. In particular, this algorithm requires the generation of 2 RSA-modulo $\mathbf{n}$ and $\mathbb{n}$, where one is used to create the signature and the other as a commitment that check

the validity of the computation. Note that $\mathbf{n}$ size depends on the order of the considered elliptic curve q_{EC}. Moreover, $\mathfrak{n}$ needs to be generated in a way that neither the sender nor the issuer knows its factors. We refer to [2] for more details.

Algorithm 1 `Setup` $(\mathbf{1}^{\kappa})$

Consider q_{EC} a prime of the right order and κ a security parameter.
Generate an RSA-modulus $\mathbf{n}$ of size at least $|2^{3\kappa}q_{EC}^2|$, where $\mathbf{n} = pq$, $\phi(\mathbf{n}) = (p-1)(q-1)$, and $|p| = |q|$.
Consider $\mathbf{h} = \mathbf{n} + 1 \in \mathbb{Z}_{\mathbf{n}^2}$
Generate $\mathbf{g}$ of order $\phi(\mathbf{n})$ in $\mathbb{Z}_{\mathbf{n}^2}$
Generate another RSA-modulus $\mathfrak{n} = p_g \cdot q_g$, where p_g, q_g are big primes, $\phi(\mathfrak{n}) = (p_g - 1)(q_g - 1)$, and $|p_g| = |q_g|$.
Consider $\mathfrak{h} \xleftarrow{\$} \mathbb{Z}_{\mathfrak{n}}$ and $\mathfrak{g} \xleftarrow{\$} \langle \mathfrak{h} \rangle$.
return $par = (\mathbf{h}, \mathbf{n}, \mathbf{g}, \mathfrak{h}, \mathfrak{n}, \mathfrak{g}, q_{EC})$.

In Algorithm 2, the parties compute the signature $\sigma_i = g^{1/x}$ where $x \equiv (sk_m + sk_i) \cdot r_1 \ mod \ q_{EC}$. Therefore, this value is randomized and the value of $\sigma = g^{1/sk_m + sk_i}$ can only be restored by the user: $\sigma = \sigma_i^{r_1}$.

Algorithm 2 `2-party Signature` (par)

The issuer computes:
  $r \xleftarrow{\$} \mathbb{Z}_{\phi(\mathbf{n})}, r' \xleftarrow{\$} \mathbb{Z}_{\phi(\mathfrak{n})}$
  $e_1 = \mathbf{h}^{\mathbf{n}/2 + sk_m}\mathbf{g}^r \mod \mathbf{n}^2$
  $\mathfrak{c} = \mathfrak{g}^{sk_m}\mathfrak{h}^{r'} \mod \mathfrak{n}$
  $PK_m\{sk_m, r, r'\}$
The sender$_i$ checks PK_m and computes:
  $sk_i, r_1 \xleftarrow{\$} \mathbb{Z}_{q_{EC}}, r_2 \xleftarrow{\$} \{0...|2^{\kappa}q_{EC}|\}$, $\bar{r} \xleftarrow{\$} \{0...|2^{\kappa}\mathfrak{n}|\}$
  $e_2 = (e_1/\mathbf{h}^{\mathbf{n}/2})^{r_1}\mathbf{h}^{\mathbf{n}/2 + sk_i r_1 + r_2 q_{EC}}\mathbf{g}^{\bar{r}} \mod \mathbf{n}^2$
  $\mathfrak{c}' = \mathfrak{g}^{sk_i}\mathfrak{h}^{\bar{r}} \mod \mathfrak{n}$,
  $PK_i\{sk_i, r_1, r_2, \bar{r}\}$
The issuer checks PK_i and computes:
  $x = Dec(e_2) - \mathbf{n}/2$
  $\sigma_i = g^{1/x}$

B. Keyed-Verification Anonymous Credentials

KVAC [3] is an attribute-based authentication scheme that allows users to prove the ownership of their personal attributes. The scheme provides functions for issuing the credentials and for proving possession of them. In this scheme, the verifier must know the issuer's secret key in order to verify the user's attributes. The scheme consists of five main functions (`Setup`, `CredKeygen`, `Issue`, `Obtain`, `Show`, `ShowVerify`).

Algorithm 3 `Issue` $(sk = (x_0, ...x_n), (m_1, ..., m_n))$

The issuer receives attributes $(m_1, ...m_n)$.
The issuer computes:
  $\sigma = g_1^{\frac{1}{x_0 + m_1 x_1 + ... + m_n x_n}}$
  $\sigma_{x_1} = \sigma^{x_1}$, $\sigma_{x_2} = \sigma^{x_2}$, ..., $\sigma_{x_n} = \sigma^{x_n}$
The issuer sends $cred = (\sigma, \sigma_{x_1}, ..., \sigma_{x_n})$ to the user.

The `Setup` function outputs the system parameters par. The `CredKeygen` function generates the issuer private key sk and issuers parameters $ipar$. `Issue` takes as input the issuer's private key and attributes $(m_1, ...m_n)$ and outputs credential $cred$ as shown in Algorithm 3.

Algorithm 4 `Show` $(\langle m_i \rangle_{i=1}^n, \sigma, \langle \sigma_{x_i} \rangle_{i=1}^n, D))$

The verifier generates and sends $nonce \xleftarrow{\$} \mathbb{Z}_q$ to the user.
The user computes:
  $r, \rho_r, \rho_{m_{i \notin D}} \xleftarrow{\$} \mathbb{Z}_q$
  $\hat{\sigma} = \sigma^r$
  $t = \sum_{i \notin D} \sigma_{x_i}^{\rho_m \cdot r} \cdot g^{\rho_r}$
  $c = \mathcal{H}(D, \langle m_i \rangle_{i \in D}, t, \hat{\sigma}, par, ipar, nonce)$
  $s_r = \rho_r + cr$
  $\langle s_{m_i} = \rho_{m_i} - cm_i \rangle_{i \notin D}$
The user sends $proof = (\hat{\sigma}, t, s_r, \langle s_{m_i} \rangle_{i \notin D}, \langle m_i \rangle_{i \in D}, D)$ to the verifier.

The credential must then be sent to the user through a secure channel. `Obtain` lets the user check the validity of the credential and the attribute value. The `Show` Algorithm allows the user to prove ownership of his attributes with selective disclosure, meaning he can only disclose some of the attributes. With the `ShowVerify` Algorithm, the verifier can check the validity of the proof. The `Show` and `ShowVerify` algorithms are shown in Algorithm 4 and 5 respectively, note that D stands for the set of disclosed attributes.

Algorithm 5 `ShowVerify` $(\langle x_i \rangle_{i=0}^n, proof)$

The verifier checks that:
  $\hat{\sigma} \neq 1_{\mathbb{G}}$
  $c = \mathcal{H}(D, \langle m_i \rangle_{i \in D}, t, \hat{\sigma}, par, ipar, nonce)$
  $t \stackrel{?}{=} g^{s_r} \cdot \hat{\sigma}^{-c \cdot x_0 + \sum_{i \notin D}(x_i \cdot s_{m_i}) - \sum_{i \in D}(x_i \cdot m_i \cdot c)}$

III. Proposed Scheme

In this section, we propose an extension of the KVAC protocol by implementing blind issuance. This can be done with the NIZKPK 2-party computation and it makes the issued credential private to the user. Therefore, this credential cannot be misused by a malicious issuer to impersonate the user. To achieve these properties, we add the user's private key to the protocol, modifying the `Issue`, `Show`, `ShowVerify` algorithms, and utilizing the 2-party computation during issuing. Algorithms 6, 7 and 8 depict our proposed scheme with changes marked in red.

Algorithm 6 `Issue` $(sk = (x_0, ...x_n), (m_1, ...m_n), sk_i)$

The issuer receives attributes $(m_1, ...m_n)$.
The issuer computes: $d = x_0 + m_1 x_1 + ... + m_n x_n$.
Run 2-party computation to compute: $k = (d + sk_i) r_1$
The issuer computes:
  $\sigma^* = g_1^{1/k}$
  $\sigma^*_{x_1} = \sigma^{*x_1}$,$\sigma^*_{x_2} = \sigma^{*x_2}$,..., $\sigma^*_{x_n} = \sigma^{*x_n}$
The issuer sends $\sigma^*, \sigma^*_{x_1}, ..., \sigma^*_{x_n}$ to the user.
The user computes $\sigma = \sigma^{*r_1}, ..., \sigma_n = \sigma_n^{*r_1}$.

For our implementation, we chose the C/C++ programming language. The main reason for this is that C libraries tend to be faster for cryptography purposes. The existing KVAC code [6] and the partial NIZKPK implementation [5] were used as a starting point, that we then modified and extended to produce our scheme. The main libraries used are GMP for computation with big numbers in the NIZKPK and micro-ECC for the computation over an elliptic curve in KVAC.

Algorithm 7 `Show` $(\langle m_i \rangle_{i=1}^{n}, \sigma, \langle \sigma_{x_i} \rangle_{i=1}^{n}, D), sk_i)$

1: The verifier generates $nonce \xleftarrow{\$} \mathbb{Z}_q$ and sends it to the user.
2: The user computes:
3: $r, \rho_r, \rho_{m_{i \notin D}}, \rho_u \xleftarrow{\$} \mathbb{Z}_q$
4: $\hat{\sigma} = \sigma^r$
5: $t = \sum_{i \notin D} \sigma_{x_i}^{\rho_m \cdot r} \cdot g^{\rho_r} \cdot \sigma^{\rho_u \cdot r}$
6: $c = \mathcal{H}(D, \langle m_i \rangle_{i \in D}, t, \hat{\sigma}, par, ipar, nonce)$
7: $s_r = \rho_r + cr$
8: $s_u = \rho_u - csk_i$
9: $\langle s_{m_i} = \rho_{m_i} - cm_i \rangle_{i \notin D}$
10: The user sends $proof = (\hat{\sigma}, t, s_r, \langle s_{m_i} \rangle_{i \notin D}, s_u, \langle m_i \rangle_{i \in D}, D)$ to the verifier.

We had to integrate the NIZKPK algorithm to work with the KVAC implementation. Since GMP uses the `mpz_t` type and micro-ECC uses the `uECC_word` for storing numbers and points, we had to use conversion trough `uint8_t` to pass between these two types. Also, the previous implementation of NIZKPK did not include the calculations of the proofs of knowledge computed by the issuer and the user, so we had to implement those to make the NIZKPK complete.

Algorithm 8 `ShowVerify` $(\langle x_i \rangle_{i=0}^{n}, proof)$

1: The verifier checks that:
2: $\hat{\sigma} \neq 1_{\mathbb{G}}$
3: $c = \mathcal{H}(D, \langle m_i \rangle_{i \in D}, t, \hat{\sigma}, par, ipar, nonce)$
4: $t \stackrel{?}{=} g^{s_r} \cdot \hat{\sigma}^{-c \cdot x_0 + \sum_{i \notin D}(x_i \cdot s_{m_i}) - \sum_{i \in D}(x_i \cdot m_i \cdot c) + S_u}$

In the implementation, we used the *secp256_r1* curve, as a 256-bit elliptic curve provides a good level of security. We used the same security parameter κ as the previous NIZKPK implementation (1350) that follows the National Institute of Standards and Technology (NIST) security standard. In the KVAC we issued 10 attributes and shown 2 during the `Show` algorithm. The set of parameters used in the implementation are shown in Table I.

IV. Experimental Results

As mentioned before, we used C/C++ programming language. We tested the implementation on multiple devices and operating systems. Devices used were: 1) a PC: HP Pavilion 15 with OS: Windows 10 Home 20H2, RAM: 16 GB, CPU: Intel i5-9300H - 4 cores 2.4-4.1 Ghz; 2) a mobile phone: Xiaomi Redmi Note 8 Pro wit OS: Android 10, RAM: 6 GB, CPU: MediaTek Helio G90T 8 cores 2.05 GHz; 3) a smart watch: Galaxy Watch 4 Classic with OS: Wear OS 3, RAM: 1,5 GB, CPU: Samsung Exynos W920 - 2 cores 1.18 GHz; and 4) a Raspberry Pi 4 Model B with OS: Raspbian, RAM: 2 GB, CPU: ARM Cortex-A72 - 4 cores 1.5 GHz.

TABLE I
Parameters used in the implementation.

Parameter	Bitsize	Value
κ (security parameter)	32	1350
q_{EC}	256	0xffffffff00000000ffffffffffffffffbce6faada7179e84f3b9cac2fc632551
$\mathfrak{n}$,$\mathfrak{u}$ (RSA moduli)	4616	random
g	9232	random
h	4616	n+1
$\mathfrak{h}$,$\mathfrak{g}$	4616	random

Since the GMP library is built mainly for Unix systems, the library did not perform as expected on Windows PC, so we also tested the implementation on the same HP Pavilion 15 laptop in a virtualized environment running Ubuntu 64-bit Linux (VMWare, RAM: 8 GB, CPU: 4 cores). While the virtualized machine will not use the hardware of the laptop to its full potential, it still gives us some results for comparison.

Therefore, we conducted benchmarks on all the mentioned devices. The times shown in Table II are averages of 10 measures each. We grouped some of the parts of the protocol together to make our table easier to understand. In Table II, `Issuer 2-party` phase consists of computation of e_1, proof of knowledge PK_M, checking user's proof of knowledge, and decrypting of e_2, i.e., Lines 1-5,11, and 12 of Algorithm 2. `User 2-party` phase consists of times of checking issuer's proof of knowledge, computation of e_2 and creating the user's proof of knowledge PK_i, i.e., Lines 6-10 of Algorithm 2. We also compare the times of the modified KVAC with the original version to see if our changes in the `Show` and `ShowVerify` algorithms did affect the protocol's runtime. We ran the whole scheme on all devices but practically the issuer part will probably not be run on mobile devices.

Looking at the times in Table II, the `Setup` computation is the heaviest part of the scheme. Specifically, the `Setup` needs to generate the prime factors for the two RSA groups and this computation takes about 95% of the total `Setup` time. Since the GMP library only provides us with a function `mpz_nextprime` which finds the next prime after a number in parameter, we have to first generate a random number of desired bit length and then use this function. Because we are looking for prime numbers of around 2300 bits, this can take a long time. Note that the time of the `Setup` varies considerably depending on how long the search for prime numbers takes. Nevertheless, the `Setup` needs to be run only once, can be generated beforehand, and, therefore, it will not affect the time of the `Issue` algorithm in a real scenario.

Figure 1 depicts the consumption in percentage taken by the individual computations in the 2-party protocol on each side. Note that for the user the PK_M in the graph stands

TABLE II
BENCHMARKS OF THE SCHEME (TIMES IN MS).

	PC-Win 10	VM-Ubuntu	Raspbery Pi	Android Phone	Smartwatch
Our Proposal					
2-Party setup	8 512	3 063	37 089	10 697	47 359
Issuer 2-party	4 457	1 714	13 907	4 636	22 076
User 2-Party	4 414	1 833	13 787	4 625	23 076
Issue total	17 397	6 674	64 873	20 029	92 726
Issue without 2-party	14	64	90	71	215
Show	8	27	43	37	128
Verify	1,5	4	9	7	24
KVAC [3]					
Issue	7	30	39	35	132
Show	8	25	38	34	113
Verify	1,4	4	9	7	24

for checking the validity of the issuer's proof, same in PK_i for the issuer. The most time-consuming operation on both sides is computing and checking the PK_i, which is logical as it consists of the bigger number of modular exponentiations, which is a heavy operation in a big modulus.

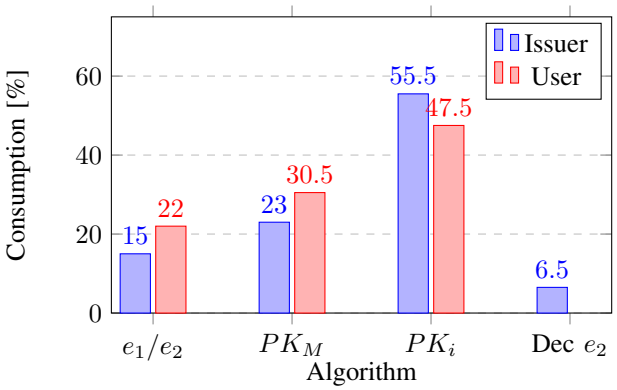


Fig. 1. Consumption used by the functions in the 2-party computation.

When we look at the total time of the `Issue` algorithm and exclude the 2-party computation we still get around twice the time than the original `Issue` algorithm, this is because there is no r_1 if we do not run the 2-party computation. This means that in our case for 10 issued attributes we have to do 10 multiplications on the elliptic curve. Note that our `Show` algorithm is only about 10% slower than the original while the `ShowVerify` takes almost the same amount of time as the original one.

As mentioned before the GMP library is built mainly for Unix systems and does not perform well on Windows. That is why the NIZKPK algorithm is so slow on the Windows PC while running almost three times faster on a Ubuntu virtual machine on the same PC, but the computations in KVAC on the elliptic curve are three times faster on the PC itself since we do not use GMP here and micro-ECC runs as good on Windows as on Unix. Therefore if we would run this algorithm on a similar PC running a Linux distribution (not virtualized), we could get times around 2.5 s for the issue algorithm. Raspberry Pi is really slow in the NIZKPK part of the scheme since this part is really memory intensive and the Raspberry Pi only has 2 GB of RAM, that is probably why the mobile phone is faster in this part while achieving similar times in the KVAC algorithms. But the Raspberry Pi could be easily used as a verifier device since the `ShowVerify` algorithm is very fast.

While true, the issue algorithm is slow, we have to consider that it only has to run once for each user, after that only the `Show` and `ShowVerify` algorithms will be run and those are almost as fast as the original, so they still should be suitable for real-life applications.

V. CONCLUSION

Anonymous credentials allow the selection of which personal data to show depending on the request of the authorized service provider. In this article, we extended the KVAC anonymous credentials scheme and show how blind issuance can be implemented into it. Moreover, to our best knowledge, we produce the first implementation of the NIZKPK scheme. While the setup algorithm of our scheme is rather slow, it only has to be run once for each user. The other algorithms of the protocol that will be used regularly are at most 10% slower than the original implementation.

REFERENCES

[1] EU General Data Protection Regulation GDPR, "Art. 5 GDPR - Principles relating to processing of personal data," GDPR.eu, November 14, 2018. Available at: https://gdpr.eu/article-5-how-to-process-personal-data/

[2] M. Belenkiy, J. Camenisch, M. Chase, M. Kohlweiss, A. Lysyanskaya and H. Shacham, "Randomizable proofs and delegatable anonymous credentials," In Annual International Cryptology Conference (pp. 108-125). Springer, Berlin, Heidelberg.

[3] J. Camenisch, M. Drijvers, P. Dzurenda, J. Hajny, "Fast keyed-verification anonymous credentials on standard smart cards," In ICT Systems Security and Privacy Protection: 34th IFIP TC 11 International Conference, SEC 2019, Lisbon, Portugal, June 25-27, 2019, Proceedings 34 2019 (pp. 286-298). Springer International Publishing.

[4] Camenisch, J., Stadler, M.: Efficient group signature schemes for large groups. In Kaliski, B., editor, Advances in Cryptology - CRYPTO '97, volume 1294 of Lecture Notes in Computer Science, pages 410–424. Springer Berlin / Heidelberg (1997).

[5] M. Sečkár, S. Ricci, "Secure Two-Party Computation for weak Boneh-Boyen Signature", Proceedings I of the 28th Conference STUDENT EEICT 2022 General Papers. 1. Brno: Brno University of Technology, 2022. s. 145-148. ISBN: 978-80-214-6029-4.

[6] T. Cvrček, "Cryptography on Arduino platform", Bachelor's thesis. Brno: Brno University of Technology, 2020.

How can Privacy enhancing technologies solve car sharing privacy issues?

1st Norbert LŐVINGER
Department of Telecommunications
FEEC, Brno University of Technology
Brno, Czechia
xlovin00@vut.cz

***Abstract*—This article explores the potential of privacy-enhancing technologies (PETs) for ensuring secure handling of private data in car sharing applications. We describe current car sharing applications, identify the main advantages of PETs, including protecting user privacy, reducing the risk of fraud and abuse, and addressing concerns of vehicle owners. We also discuss the technical challenges associated with implementing PETs in car sharing applications. Finally, we propose a proof of concept for implementing PETs in car sharing applications and suggest cryptographic tools that can be used to secure user's data. Our findings demonstrate that PETs have the potential to make car sharing a viable and attractive mode of transportation by addressing concerns regarding privacy and security.**

***Index Terms*—Smart transport services, Data and privacy protection, Privacy enhancing technologies, Pseudonym certificates**

I. Introduction

A car sharing is one of the most promising application of Smart Transport Services (STS), which has the potential to revolutionize the way we think about transportation. As urbanization continues to grow, more and more people are choosing to live in cities, leading to an increase in traffic congestion and pollution. Car sharing has emerged as a viable solution to these problems, offering a more sustainable and cost-effective alternative to traditional car ownership.

However, as car sharing becomes more popular, concerns about data privacy and secure handling of private data have become a significant issue. Current car sharing platforms collect a vast amount of personal data from users and these data are often stored and processed by third-party service providers. Due to these concerns, privacy-enhancing technologies (PETs) have been developed to ensure that the user's data privacy and security are maintained. PETs use cryptographic techniques to protect user data. By using PETs, car sharing platforms can offer a more secure and trustworthy service to their users.

This article explores the current situation of car sharing and why it is considered the future of transport, discussion about the advantages, main data privacy concerns and propose a part of a car sharing concept using PETs.

A. Related works

Recent research on STS has focused on various aspects, such as optimization of the traffic flow, where Fredianelli et al. (2022) have been working on a system of using low cost cameras to increase traffic flow efficiency in the Italian city of Piombino [1]. Equally important research area of STS is intelligent parking management where Dzurenda et al. (2021, 2022) have been exploring the possibilities of using blockchain technology and smart contracts for parking in city centre zones [2] and they also looked on user data protection [3].

Car and scooter sharing services have been increasingly used in the STS and became one of the most promising applications for research. Mounce et al. (2019) who have been looking into the idea of linking electric vehicles with a one-way car sharing system. Authors elaborated on current solutions, usability, legislation, benefits and autonomous driving [4]. The use of PETs in STS has been investigated in various studies. For instance, Pollicino et al. (2021) have proposed a new protocol for accountability in realistic car sharing scenarios [5]. In another study, Huang et al. (2020) proposed an architecture named DAPA to solve privacy concerns about users private data in car sharing service providers [6].

In this article, we explore the potential of PETs to ensure secure handling of private data in a car sharing application model.

II. Background

Data and privacy protection in STS can be achieved by implementing technical and organizational measures, such as obtaining informed consent from individuals, implementing privacy-by-design principles in system development, and providing clear and transparent privacy policies. Data protection and privacy considerations must be addressed throughout the entire system, from data collection to transmission, processing, and storage.

Privacy Enhancing Technologies (PETs) are a set of tools and techniques designed to protect the privacy of individuals in the digital world. They aim to enhance the privacy of users by **minimizing the collection, storage, and use** of personal information, while still allowing the intended functionality of the services being used.

In the context of STS, PETs could play a crucial role in protecting the privacy of individuals who use these services. STS generate a vast amount of data about users' travel patterns, behaviors, and preferences. These data can be sensitive, and could lead to privacy violations, such as unauthorized tracking or profiling of users.

The advantages of PETs in terms of STS are **anonymous identity**, better **control over personal data**, **privacy-by-design** software development with **minimizing users' data collection** with greater **transparency** about how personal data are collected, used, and shared.

There are also several challenges associated with the implementation of PETs such as **interoperability and compatibility problems** between different architectures, **impact on performance** due to additional data processing, **difficult and complex implementation** for providers with **balance between data privacy and service utility**, which comes with a **implementation cost**.

III. Reference scenario

Imagine a city where many people are looking for a *cost-effective and efficient way to travel* to work, run errands, or meet friends. Car sharing service providers offer a solution, where users can rent a car for a specific period of time, use it to travel to their desired destination, and return it to the designated parking spot. The service provider manages the fleet of cars and ensures that they are in good condition and ready for use.

In our proposed scenario, we consider a stationary car sharing business with multiple rental spots in the city, where service provider could rent their car fleet to other users under *predefined rules*, such as car availability, users' legal status, pick up and drop off area and length of the rental. Users with valid documents are authenticated by Trusted authority (TA) and allowed to enter the car sharing application under the created pseudonyms to prevent identity exposure. In the car sharing application, the user enters the parameters for renting a car and filters from the available database of vehicles from service providers. Service providers are also displayed in the form of pseudonyms, to protect their identity. After selecting a car from the database, the user pays the rental price and opens the car at the selected time and place using their smartphone. These steps are beyond the scope of this paper.

The aim of our proposal is the selection of a specific car sharing process that is adopted in commercial solutions and leverage PETs to improve data and privacy protection of the users. We have focused on the individual registration process of each entity into the car sharing application using TA and the pseudonyms certificates, which belongs to the PETs category of *anonymous credential systems*.

We assume that users and service providers have secured connection with TA and both parties are mutually communication using standardized secure cryptographic tools.

IV. Proposed solution

Popular car sharing applications present significant data privacy and security challenges, such as the risk of unauthorized access to personal data or tracking of user locations. PETs play a critical role in mitigating these risks and protecting user privacy in car sharing applications.

A. Model Proposal

In this proposed model of car sharing application using PET, we have five main components: **users, service providers, car sharing application, cars, and a trusted authority**. The users access the car sharing application through a mobile application or web portal. The TA oversees the users and service providers who manage the vehicles. The car sharing application is responsible for the secure exchange of users' and providers' data necessary for the fulfillment of the car rental act.

- **Users** – Individuals who wish to use the car sharing service. They access the service and provide personal information such as their name, contact information, and payment details.
- **Service providers / Renters** – Organizations that renting the car using car sharing application. They may collect data about their car fleet, such as location information, trip details and mandatory technical documents.
- **Car sharing application** – Web or mobile application which operates the car sharing service. It displays users and providers current offers and demands. Therefore may collect data about users habits or renters car fleet availability.
- **Cars** – Shared vehicles among users. Cars may be equipped with GPS or other sensors to enable tracking and monitoring of their location and usage.
- **Trusted Authority** – A trusted intermediary between users and service providers using PETs to protect user privacy and user control over their sensitive private data.

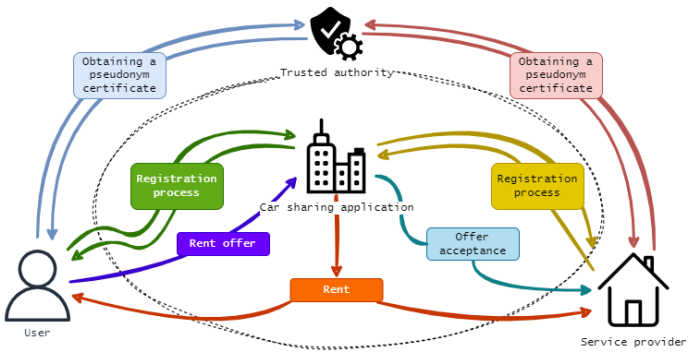


Fig. 1. Privacy-enhancing car sharing model proposal

B. Model Phases

The basic phases of a car sharing process can be divided into several steps with use of PETs:

- **User registration and authentication** During this phase, the user registers with the car sharing service by providing personal information such as name, email, and payment details. PET tools like *Zero-Knowledge Proofs (ZKP)* or *Pseudonyms identity using digital certificates* can be used for user authentication without revealing any

sensitive information, thereby ensuring confidentiality and privacy.

- **Vehicle selection and reservation** After registration, the user selects the preferred vehicle and makes a reservation. PET tools such as *Secure Multi-Party Computation (SMPC)* can be used to ensure data privacy and integrity during the vehicle selection and reservation phase. The User, Service Provider and Car sharing application can securely participate in the vehicle selection and reservation process without revealing their private information to each other.
- **Vehicle access and monitoring** During this phase, the user gains access to the reserved vehicle through a secure and private attribute-based authentication process. PET tools like *Homomorphic Encryption (HE)* can be used to ensure data confidentiality and privacy during car monitoring phase.
- **Secure payment and user revocation** Payment is a critical phase in the car sharing process, and PET tools such as *Private Information Retrieval (PIR)* can be used to ensure secure and private payment transactions. In case of any security concerns or misuse of the sharing service, user revocation is necessary. PET tools such as *Attribute-Based Encryption (ABE)* can be used to ensure conditional traceability and revocation of user access to the sharing service. In a car sharing application, each user could be assigned a set of attributes that describe their access rights to the car sharing service. If a user has not returned a rental vehicle or has violated the terms of service, their access to the car sharing service could be revoked by removing their attribute for vehicle rental.

These phases are interconnected and require secure communication and data handling using PETs to ensure user privacy and security throughout the car sharing process.

C. Model Requirements

Three main categories for the model requirements:

- **Security requirements** – *Data Integrity, Authentication, Confidentiality, Unlinkability, Revocation, Conditional Traceability* are essential for ensuring the data security of the car sharing system.
- **Privacy requirements** – *Data Minimization, Purpose Limitation, Transparency, User Control, Anonymity and Non-Discrimination* are critical for ensuring that the car sharing system respects user privacy and autonomy where users can trust the service to handle their personal data.
- **Functional requirements** – *User Registration, Vehicle Reservation, Access and Status monitoring, Payment Processing, Customer Support, Trustee Third Party Control* are mandatory for ensuring that the car sharing system meets the needs and expectations of users, and that the service is efficient, sustainable, and user-friendly way.

V. Implementation Verification

One of the first phases in car sharing application is *user registration and authentication*, which is crucial for ensuring secure and authorized access. We proposed **pseudonym certificate system** based on digital certificates and chose it due to its ubiquitous usage also outside of car sharing and to demonstrate the real-world use of PETs. Additionally, we included a **Trusted authority mechanism** to oversee the entire car sharing application and provide additional security measures.

The use of **Pseudonyms identity** ensures that the user's identity remains hidden behind a generated certificate, protecting their privacy while providing secure access to the car sharing application. To demonstrate the feasibility of our proposal, we implemented a proof of concept in **Python language** using the official **crypto library** [7]. The library can provide cryptographic primitives such as *public key cryptography and digital signature algorithms*. Full code is available on public GitHub repository [8].

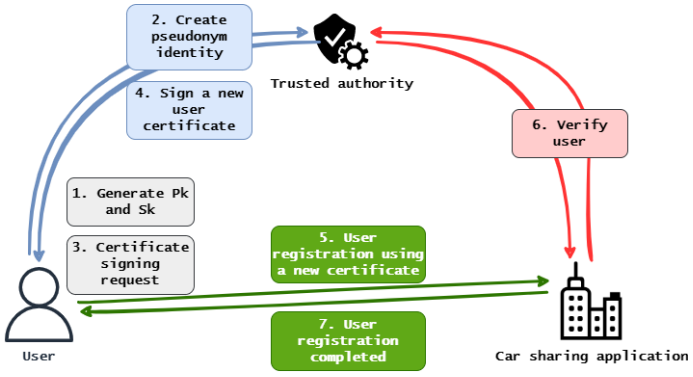


Fig. 2. Pseudonym certificates implementation proposal

A. Demonstration

The pseudonym certificate system implementation consists of four steps:

1) The user generates a public-private key pair using **RSA-3072** and creates a unique pseudonym identity which is based on the personal information using hash function **SHA-256**.
2) The TA receives the user's public key and created pseudonym identity from the user in a form of certificate signing request (CSR) using **X.509** standard. The CSR is then signed with a new certificate by the TA, verifying the user's data using again **X.509** standard.
3) The user approaches the registration process with their pseudonym identity and the signed certification received from the TA. The car sharing application verifies the user's signed pseudonym identity against the certificate provided by the TA.
4) If the user's certificate is correct, the user is allowed to enter the car sharing application under his pseudonym identity and user's registration can been completed.

This proof of concept demonstrates the feasibility of user registration and authentication phase using Pseudonyms identity and Trusted authority mechanism. This approach can provide a *secure and efficient way to manage user access* to car

sharing applications while *preserving user privacy and data confidentiality*.

B. Results

To evaluate the performance and functionality of our proof of concept implementation we used a Windows 10 Pro 64-bit machine equipped with an Intel i7-6500U CPU and 16GB of RAM. For the implementation, we used Python version 3.9.7 and compiled crypto library version 23.1.1. These versions were chosen for their stability and compatibility with the implementation requirements.

We conducted 100 runs to measure the system's performance, including the speed of the cryptographic primitives of generating *RSA-3072* key pairs. The average time for key generation was **841.09ms**. It is important to ensure that sufficient computational resources are available when generating RSA keys, especially for large key sizes such as 3072 bits. We also conducted another 100 runs to measure the efficiency of the generating CSR request and user's certification where the average time was **7.06ms**. Additionally, it may be useful to consider alternative cryptographic algorithms such as **PQ Dilithium algorithm** based on lattice-based cryptography that may be faster and more secure. This method was also tested in comparison to the RSA-3072 key pairs generation.

The testbed provided a stable and reliable platform to evaluate the efficiency and effectiveness of our proposed approach under real-world conditions. The use of PETs has become more important than ever before and with the availability of powerful cryptographic tools, implementing PETs in these applications can be fast and more secure than ever before as long as they are implemented in the correct way.

Conclusions

In this article, we presented the use of the PETs in car sharing applications. Car sharing has gained popularity in recent years as a cost-effective and eco-friendly alternative, it has also raised concerns about the user privacy and security of personal data collected through these platforms. PETs can protect personal data while still allowing car sharing platforms to function effectively. Implementing PETs in car sharing applications requires careful consideration of technical requirements, user needs, and regulatory compliance as potential trade-off between privacy and functionality. We proposed a simple Python implementation of an individual user registration process for a car sharing application using pseudonym certificates as one of the PET tools. Future work in this area should focus on the development of standardized and interoperable PETs that can be seamlessly integrated into different scenarios in car sharing applications. Also the adoption of Post Quantum algorithms in PETs should be explored to improve the future-proof of car sharing applications.

Acknowledgement

This work is supervised by: Lukas Malina (Department of Telecommunications, FEEC, Brno University of Technology, e-mail: malina@vut.cz).

References

[1] FREDIANELLI, Luca, Stefano CARPITA, Marco BERNARDINI, Lara Ginevra DEL PIZZO, Fabio BROCCHI, Francesco BIANCO a Gaetano LICITRA. Traffic Flow Detection Using Camera Images and Machine Learning Methods in ITS for Noise Map and Action Plan Optimization. Sensors [online]. 2022, 22(5) [cit. 2023-03-09]. ISSN 1424-8220. Available from: doi:10.3390/s22051929

[2] DZURENDA, Petr, Carles Anglès TAFALLA, Sara RICCI a Lukas MALINA. Privacy-Preserving Online Parking Based on Smart Contracts. The 16th International Conference on Availability, Reliability and Security [online]. New York, NY, USA: ACM, 2021, 2021-08-17, 1-10 [cit. 2023-03-09]. ISBN 9781450390514. Available from: doi:10.1145/3465481.3470058

[3] DZURENDA, Petr, Florian JACQUES, Manon KNOCKAERT, Maryline LAURENT, Lukas MALINA, Raimundas MATULEVICIUS, Qiang TANG a Aimilia TASIDOU. Privacy-preserving solution for vehicle parking services complying with EU legislation. PeerJ Computer Science [online]. 2022, 8 [cit. 2023-03-09]. ISSN 2376-5992. Available from: doi:10.7717/peerj-cs.1165

[4] MOUNCE, Richard a John D. NELSON. On the potential for one-way electric vehicle car-sharing in future mobility systems. Transportation Research Part A: Policy and Practice [online]. 2019, 120, 17-30 [cit. 2023-03-09]. ISSN 09658564. Available from: doi:10.1016/j.tra.2018.12.003

[5] POLLICINO, Francesco, Luca FERRETTI, Dario STABILI a Mirco MARCHETTI. Accountable and privacy-aware flexible car sharing and rental services. 2021 IEEE 20th International Symposium on Network Computing and Applications (NCA) [online]. IEEE, 2021, 2021-11-23, 1-7 [cit. 2023-03-09]. ISBN 978-1-6654-9550-9. Available from: doi:10.1109/NCA53618.2021.9685942

[6] HUANG, Cheng, Rongxing LU, Jianbing NI a Xuemin SHEN. DAPA: A Decentralized, Accountable, and Privacy-Preserving Architecture for Car Sharing Services. IEEE Transactions on Vehicular Technology [online]. 2020, 69(5), 4869-4882 [cit. 2023-03-09]. ISSN 0018-9545. Available from: doi:10.1109/TVT.2020.2980777

[7] CRYPTO — Generic cryptographic module: OpenSSL — Python interface to OpenSSL [online]. 28.03.2023 [cit. 2023-04-02]. Available from: https://www.pyopenssl.org/en/latest/api/crypto.html

[8] GitHub: Practical part of the article "How can Privacy enhancing technologies solve car sharing privacy issues" [online]. 30.3.2023 [cit. 2023-04-02]. Available from: https://github.com/xlovin00/EEICT2023

Cyber Arena Automation Component Modernization including Backup and Recovery Processes

1st Veronika Fisarova
Department of Telecommunications
FEEC, Brno University of Technology
Brno, Czechia
xfisar01@vut.cz

2nd Tomas Stodulka
Department of Telecommunications
FEEC, Brno University of Technology
Brno, Czechia
xstodu07@vut.cz

***Abstract*—This paper describes the modernization of the BUTCA (Brno University of Technology Cyber Arena) component responsible for automating processes, as well as the implementation of backup and recovery procedures for critical parts of the OpenStack cloud computing platform used by Cyber Arena. Specifically the paper focuses on modernizing the AWX (Ansible Web eXecutable) component and creating automated backups for OpenStack critical data. It was researched that already implemented tools for the administration of OpenStack, such as existing OpenStack CLI (Command Line Interface) and Ansible OpenStack collections, lacked the necessary capabilities for creating complex backup and recovery processes. Therefore the focus is on obtaining data for backups from the MariaDB database that OpenStack uses to store data. The paper concludes with an overview of proposed approach to modernizing and streamlining Cyber Arena's backup and recovery processes on OpenStack.**

***Index Terms*—Ansible, automation, AWX, backup and recovery, BUTCA, cloud computing, Cyber Arena, cyber range, MariaDB, OpenStack**

I. Introduction

Information and communication systems have become an indispensable part of our daily lives and are crucial to the functioning of many organizations. However, these systems are susceptible to cyber attacks, which can compromise the security and integrity of the information they store and transmit. Recent events, such as the attack on Czech hospitals during the COVID-19 pandemic and the conflict in Ukraine, have highlighted the critical nature of cybersecurity and the urgent need to be prepared for such attacks. According to cybersecurity expert Dr. Kenneth Geers, future conflicts are likely to test the security of cyberspace more than military strength [1]. To meet this challenge, cyber ranges are being developed to provide a realistic environment for testing ICT network defenses. The Cyber Arena at Brno University of Technology (BUTCA) is one such cyber range that was built [2].

To build a cyber range that can withstand the onslaught of participants at any given moment, cloud computing technology is essential. The Cyber Arena utilizes the OpenStack cloud to provide computing and storage resources, and to allocate these resources, virtualize the infrastructure, network, and orchestrate the training scenarios. However, administering OpenStack to provide these features to the Cyber Arena is necessary. This administration is handled by AWX (Ansible Web eXecutable) tool, which serves as the orchestration layer throughout the entire Cyber Arena complex [3]. While the AWX platform is still in development with new features being added, the present Cyber Arena's AWX is outdated.

This paper focuses on modernizing the AWX automation tool and designing the backup and recovery processes in AWX for OpenStack administration to contribute to improved ICT network security. The modernization of the AWX platform is expected to enhance the Cyber Arena's capabilities and improve its ability to simulate real-world cyber attacks by improving the administration of OpenStack.

II. Ansible Web Executable

AWX is an extension of Ansible that provides a web-based user interface, REST API (Representational State Transfer Application Programming Interface), and a task engine built on top of Ansible. It offers features such as job scheduling, run history, multi-tenancy, and job templating.

A. The Essence of AWX in the Cyber Arena Environment

In the Cyber Arena environment, AWX serves as an intermediary between OpenStack and the Cyber Arena web UI. It allows for easy integration with other applications through simple web-based calls. AWX facilitates specific actions, including generation or destroying of game training scenarios based on game status in Cyber Arena, generating participant consoles, and administration tasks [3].

B. Key Differences between AWX Versions

There are several key differences between AWX version 17 and version 21. One major addition in AWX 21 is the job slicing feature, which allows large jobs to be split into smaller pieces and run concurrently across multiple machines. Another important addition is the ability to create custom credential types, which can be used to store and manage any type of secret data needed by the automation workflows. The inventory management functionality has also been improved in AWX 21, making it easier to manage hosts and groups, and enabling more flexible grouping and filtering of inventory data. Additionally, AWX 21 includes a number of security enhancements, such as improved password hashing algorithms

and support for two-factor authentication. Overall, these new features make AWX 21 more robust and secure for managing automation workflows than its predecessor. Version 21 also uses a new version of ansible-runner, specifically version 2.10, which includes many major changes. The main change for the purpose of backup and restore processes is the way Ansible collections and plugins are used [4]. The importance of the Ansible collections is explained later in the text.

Starting at version 18, deployment methods have also changed, AWX 17 can be deployed using a Docker Compose file, while AWX 21 is recommended to be deployed using the awx–operator tool, which introduces better support for Kubernetes clusters. Deployment into a Kubernetes clusters allows more flexible customization options of the future AWX environment and can simplify management of the AWX infrastructure [5], [6]. The following table I summarizes the requirements for installing AWX with and without awx-operator. All information listed is available in the references [5], [6].

TABLE I
SUMMARY OF AWX DEPLOYMENT ANALYSIS

Container manager	OpenShift	Kubernetes	Docker
Container manager hardware requirements	3 CPU 6 GB RAM	3 CPU 6 GB RAM	2 CPU 4 GB RAM
Deployment without awx-operator			
Supported versions	<18	All	<18
AWX hardware requirements	2 CPU and 4 GB RAM 20 GB disk space		
Upgrade possible	Yes		
Deployment with awx-operator			
Supported versions	>18	>18	None
AWX hardware requirements[a]	4 CPU and 8 GB RAM 20 GB disk space		None
Upgrade possible	Yes		None

[a]It depends on the specific needs of the container manager distribution.

III. DEVELOPMENT ENVIRONMENT

As part of the modernization of the AWX component and the design of a backup and recovery processes, it was necessary to create an appropriate development environment that mimics the key components and parameters of the Cyber Arena infrastructure. Based on the Cyber Arena infrastructure described in the literature [3], a development environment was created as shown in the Figure 1.

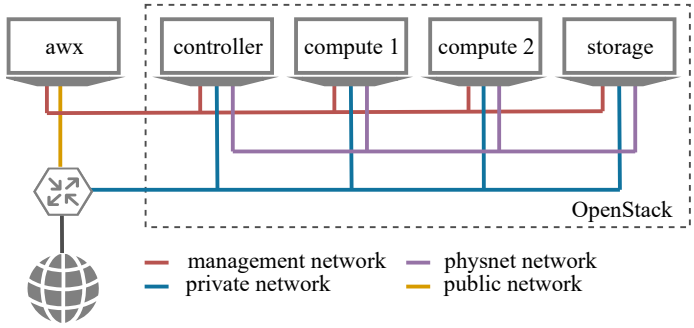


Fig. 1. Created development environment

The OpenStack infrastructure shown in Figure 1 includes a controller node, two compute nodes, and a storage node. The Victoria version of OpenStack was deployed using the Kolla-Ansible deploying tool.

The figure 1 also includes the AWX component, which was deployed using the axw-operator tool into the Kubernetes cluster. As a result of the successful deployment of the AWX component, a guide was created so the deployment can be replicated in the modernization of the Cyber Arena.

IV. THE ESSENCE OF OPENSTACK BACKUP AND RECOVERY

It's important to clarify that the purpose of backing up OpenStack is not to create backups of the parts related to the training scenarios. All the data OpenStack needs to successfully create the training infrastructure are in the Heat templates, so there's no need to back them up. The essence of backing up and recovering OpenStack is to ensure that critical parts that affect the functioning of the Cyber Arena can be recovered in case there are any issues during OpenStack modification or update.

A. OpenStack Migration and Backup Tools

In the backup and recover processes of OpenStack selected parts, it is necessary to distinguish two different approaches. The first approach, called OpenStack migration, is suitable if it is necessary to transfer most of the data from one cloud to another, either within the same version or a different version. The migration process involves copying the required data and configurations, as well as ensuring the preservation of all dependencies. It is very important to consider that the possibility of migration is strongly influenced by the method of deploying OpenStack. For example, OpenStack deployed using TripleO method mainly uses Heat templates for deployment, while OpenStack deployed using the Kolla-Ansible method uses Docker containers [7], [8]. Therefore the major differences will be present, configuration files and OpenStack directory structure is affected the most, which implies that there is no uniform approach to a cloud migration.

Several tools have been developed for cloud migration, with a significant portion of them being proprietary and allowing the data to be migrated between two different cloud platforms. The native OpenStack tool for migration, backup, and recovery is the Freezer project, which allows for taking snapshots of various parts of the cloud, and also supports incremental backups [9]. Although, the implementation of the Freezer into the Cyber Arena's OpenStack is one possible approach for the back up and recovery processes, it would need proper understanding of a quite large tool that would bring another complex layer into the Cyber Arena's environment.

Another tool native to Kolla-Ansible deployment is named Mariabackup, which takes a snapshot of the database files at a particular point in the time. This backup includes all the database's tables, indexes, and data, and the backup files are stored on a separate volume or server outside of the MariaDB container. However, Mariabackup provided by Kolla-Ansible

has a few disadvantages and limitations. First, the snapshot files created by Mariabackup can be quite large, presenting challenges for storing and transferring backup files, which must be considered while creating automated processes using Mariabackup. Second, the snapshots are stored in a dedicated backup volume, and therefore, the recovery process must be able to handle the fact that the original MariaDB container may be destroyed and recreated, which changes the metadata of the container. Lastly, configuring Mariabackup is quite complex, requiring a deep understanding of containerization, thorough knowledge of MariaDB and its architecture, as well as the underlying system configuration and its customization [10].

Furthermore, while these tools are being integrated into the OpenStack cloud, they lack the advantages of the AWX automation layer. Therefore, they are not able to leverage API calls or it would be needed to implement them above the standard capabilities offered by these tools.

B. Different Approaches of Designing the Backup and Recovery Processes

In the previous sections, various methods of creating backups of OpenStack data were described. However, it follows from the above that implementing these methods and tools into the life cycle of the Cyber Arena would add a complex layer which would require more complicated configuration and preparation. The purpose of designing our own automated process for obtaining backups and their subsequent recovery is to minimize complexity, even in the case of reduced customization options.

To ensure successful and recoverable backups, it is essential to collect all necessary data and ensure its integrity in the backup process. To design the backup process, three different approaches can be taken in general:

1) Retrieving the necessary data using the OpenStack CLI (Command Line Interface) and processing it with Ansible;
2) Using convenient Ansible collections, such as the `openstack.cloud` collection;
3) Accessing the MariaDB OpenStack database to retrieve the necessary data.

OpenStack CLI is a native tool for managing OpenStack clouds, but using it for backup and recovery may not provide complete data or important metadata, requiring additional API calls. Also this approach is not scalable for larger and highly customized OpenStack deployments, even with automation.

Ansible provides the `openstack.cloud` collection, which allows for obtaining and managing certain OpenStack services using a pre-implemented Python module. Using automation can simplify and speed up the process of collecting data for backups. However, Ansible collections are incomplete for all components needed for backup.

The MariaDB database is used as the primary database backend for storing and managing information about OpenStack's configuration and state and user data, which is based on MySQL. Therefore, it is possible to export the database content using the `mysqldump` command. It is also possible to modify data directly in the database, but this should be done with caution to avoid inconsistencies or damaging the data.

V. Design of the Backup and Recovery Process

When designing an automated backup process to support the Cyber Arena life cycle, the aim is to create a straightforward, easy-to-use, and simple-to-debug implementation. The process should require as few additional system packages as possible and be as independent as possible from the OpenStack version and customization. To achieve this, it is necessary to combine all the approaches described in section IV-B. Specifically, the usage of the OpenStack CLI for cloud authentication, Ansible collections to access the MariaDB container, and the MariaDB databases and tables for data retrieval.

A. The General Workflow of Automated Processes

The entire backup and recovery process is initiated through the AWX web interface. To enable this, access to the OpenStack instances (specifically the controller node where the MariaDB container is located) must be ensured. Ansible generally manages stations through an SSH (Secure Shell) connection using a certificate, username and password, or other authentication methods. In the AWX environment, access is set up using an inventory and a credentials variable. Authentication for using the OpenStack CLI is then provided on the station that AWX will manage, using a file with authentication details provided to the Ansible engine.

Finally, AWX must have access to the Ansible playbooks themselves, which outline the individual backup and recovery steps. AWX has a feature to access GitHub repositories (and similar) from which it can retrieve and potentially update the playbooks. This is done using a personal Git token, hence the access is possible even for private repositories. The described workflow is shown in the Figure 2.

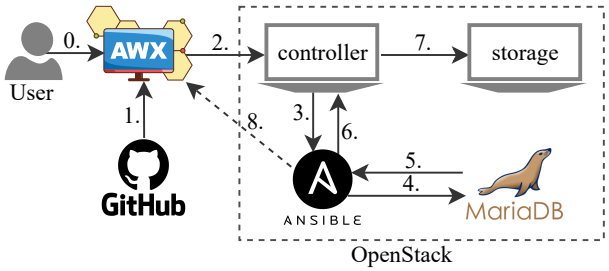


Fig. 2. Workflow diagram

The user interacts with the AWX and runs the required playbooks from there (process 0). GitHub provides source files for running playbooks (process 1). AWX executes playbooks and manages access to the OpenStack instances (process 2). The specific OpenStack instance provides a system environment for the Ansible engine (process 3). The Ansible engine is responsible for executing Ansible playbooks, which are a set of instructions written in YAML (YAML Ain't Markup Language) format that defines the desired state of a MariaDB component (process 4). The result of process 4 is sent back

to the Ansible engine for processing, including the return code and outputted backup files (process 5). The Ansible engine then saves the output files into a pre-defined directory, volume, or external server (process 6 and 7). Simultaneously, the Ansible engine reports back to AWX (process 8). The same workflow diagram is used for the recovery process except the process 6.

B. Ansible Project Structure

When creating an Ansible project to manage MariaDB databases in OpenStack, it's important to minimize code redundancies while maintaining understandability. This approach is important to minimize the possibility of errors as much as possible. MariaDB is used by the following OpenStack services for database storage: Cinder, Glance, Heat, Keystone, Neutron, Nova, and Placement. The Ansible project's design adheres to the best practices and conventions detailed in the Ansible documentation [11]. As a result, the project features two primary playbooks – `backup.yaml` and `restore.yaml`, which call their respective tasks and roles where the specific steps for backup and recovery of the MariaDB are being implemented.

Ansible tasks use the `community.docker` Ansible collection, which allows easy access to the Docker container where the MariaDB database is running, and obtain a backup using `mysqldump` command. This way, a schema of the entire database is created, which contains SQL commands that allow its recreation in the recovery process. The output of this process is a file called *database-name_dump.sql*, along with the hash of the dump file data. Prior to beginning the restore process, a new hash is calculated and compared with the existing hash.

To facilitate repeatable testing of both the backup and restore processes, additional Ansible playbooks were created to automate the creation of the mentioned resources and verify their presence. In this manner, before each backup process, the specified OpenStack resources are deployed, and after the creation of the database dump files, these resources can be edited or deleted. Once the restore process is completed, the playbook that checks for the presence of the previously created OpenStack resources can be run to verify a successful restoration.

C. Current State of Development and Results

Currently, both backup and recovery processes have been implemented for each of the MariaDB databases. The entire process was executed in the development environment using the AWX platform, and its functionality was tested through the automated creation of various OpenStack resources, such as projects, users, networks, subnets, routers, access rights, instances, and their combinations, to cover as much of the MariaDB database as possible. Based on the results of the testing, the dump files were successfully created and then the dump files were used to restore MariaDB to previous state.

The latest implementation can be found in the GitHub repository referenced as [12]. In the close future, the goal is to refine the tests to create more complex resources, increasing the probability of detecting any uncaught errors and problems. Since MariaDB does not contain any images and snapshots themselves, which OpenStack is using, a new role is going to be created to directly backs up these images and snapshots from the Glance service.

VI. Conclusion

This paper highlights the importance of the automated component AWX, which is utilized by Cyber Arena. The analysis in this paper explores deployment options that will be used to modernize AWX to a newer version. Additionally, the creation of a development environment that replicates the OpenStack cloud on which Cyber Arena is built has been described. The new version of AWX was deployed in this environment to test the updated features.

AWX is also used for creating backups of important user data, which need to be recoverable in the event of OpenStack failure or data corruption. The process for backup and recovery was designed, implemented, and tested. However, it is still under development, so the completed implementation may undergo changes. The overall goal of this paper and the Master's thesis on which it is based is to improve the life cycle of Cyber Arena and simplify the administration of its OpenStack infrastructure, which will lead to an increase in the quality of Cyber Arena as a whole.

Acknowledgment

I would like to thank the Department of Telecommunications for the provided computing environment.

References

[1] Antoniuk D., "Q&A: Kenneth Geers on the cyber war between Ukraine and Russia," [online], October 22, 2022 [cit.2023–03–01]. https://therecord.media/qa-kenneth-geers-on-the-cyber-war-between-ukraine-and-russia/.

[2] Technology agency of the Czech Republic, "Cyber-arena for research, testing and education in cybersecurity", [online], 2019 [cit.2023–03–01]. https://starfos.tacr.cz/en/project/VI20192022132#project-main.

[3] T. Stodulka and R. Fujdiak, "Building the Cyber Range platform with cloud computing technology," Brno, Czechia: Brno University of Technology, 2022, p. 153, ISBN: 978-80-214-6064-5.

[4] Ansible, "Ansible project 2.10", [online], 2023 [cit.2023–03–01], https://docs.ansible.com/ansible/latest/roadmap/COLLECTIONS_2_10.html.

[5] Ansible, "AWX Project", [online], 2023 [cit.2023–03–01], https://github.com/ansible/awx/.

[6] Ansible, "AWX Operator', [online], 2023 [cit.2023–03–01], https://github.com/ansible/awx-operator.

[7] OpenStack, "TripleO Deployment Guide", [online], 2023 [cit.2023–03–01], https://docs.openstack.org/project-deploy-guide/tripleo-docs/latest/.

[8] OpenStack, "Kolla Ansible Deployment Guide", [online], 2023 [cit.2023–03–01], https://docs.openstack.org/project-deploy-guide/kolla-ansible/latest/.

[9] OpenStack, "OpenStack Freezer's documentation", [online], 2023 [cit.2023–03–01], https://docs.openstack.org/freezer/latest/.

[10] OpenStack, "Kolla-Ansible MariaDB database backup and restore", [online], 2023 [cit.2023–03–01], https://docs.openstack.org/kolla-ansible/latest/admin/mariadb-backup-and-restore.html.

[11] Ansible, "Ansible Documentation", [online], 2023 [cit.2023–03–29], https://docs.ansible.com/ansible/latest/index.html.

[12] Deydra71, 'OpenStack-MariaDB-backup-and-restore", [source code], GitHub 2023 [cit.2023–03–29], https://github.com/Deydra71/OpenStack-MariaDB-backup-and-restore.

Enhancing the efficiency of fiber optic eavesdropping

1st Vladimir Spurny
Brno University of Technology,
FEEC, Dept. of Telecommunications,
Technicka 12, 616 00 Brno, Czech Republic
orcid.org/0000-0003-4132-1444

2nd Petr Munster
Brno University of Technology,
FEEC, Dept. of Telecommunications,
Technicka 12, 616 00 Brno, Czech Republic
orcid.org/0000-0002-4651-8353

***Abstract*—It is generally known that optical fibers is difficult to eavesdrop and are therefore considered secure. Macrobending-based eavesdropping methods are relatively easy to implement without link disrupting, and that's why the risk of these attacks is higher. However, it is necessary to use special equipment that will allow the emitted light to be coupled into another fiber. The paper evaluates a simple commercially available taping clip on coupler. The measurement is carried out on standard telecommunication fibers G.652, while not only the power level of emitted light is evaluated, but also the attenuation of the intercepted link. Moreover, we present a modification of the equipment to increase its efficiency.**

***Index Terms*—fiberoptics, security, eavesdropping, attack, transmission, fiber**

I. Introduction

Optical networks are considered secure and data transmission is relatively eavesdropping-proof due to the principle of signal transmission in optical fibers. However, there are several ways to intercept part of the signal in an optical fiber without the user realising that he is being eavesdropped on. Although eavesdropping on the physical layer seems impossible and complicated, the fact that it occurs in practice is obvious. The basic methods are known to all, but very few measurements have been made with practical verification to confirm or refute the possibility of eavesdropping. Although cryptography is used to secure data today, it is expected that in the future current encryption methods will not be sufficient due to quantum computing. Physical layer and data security will need to be protected, for example, by quantum key distribution [1]–[3].

Attacks on fiber optic networks are very diverse in their approach. These range from simple non-interference with the infrastructure, to individual eavesdropping methods that exploit the physical properties of the optical fiber transmission, or methods that involve adding passive components to the route to enable eavesdropping. These methods are generally difficult to detect, can serve undetected for a very long time and can be very effective. Other eavesdropping methods include handling crosstalk that occurs with passive components, or using unsecured unused ports of passive optical splitters. When designing the infrastructure, it is necessary to take into account the possibility of individual attacks.

Risks are not only local, but also affect global backbone networks to a much greater extent, particulary in areas affected by war or political regimes [4]–[6].

II. Related Studies

In most of the literature, one can find a theoretical summary of the security risks of fiber optic transmissions. An interesting evaluation is offered by the work of Skorin-Kapov and Furdek [7] who describe common methods of attacks on the physical layer. They describe attacks exploiting jamming by tying an interfering beam to the fiber, attacks consisting in disrupting the function of erbium doped fiber amplifiers (EDFAs), common tapping attacks or attacks targeting denial of service. Unfortunately, these methods are only theoretically described and there is no experimental verification and risk evaluation. The authors in [8] mention the possibility of simple eavesdropping using a clip-on coupler, but they do not describe the effect on ongoing communication, the possibility of detecting attacks using optical time domain reflectometry (OTDR) or measure the increase of optical path attenuation or the power of the eavesdropped signal. The authors in [9] performed a similar experiment with the possibility of eavesdropping on video transmissions, but the results are reported in an analogous way to the previous work. They also performed a verification of the macro bend radius on the amount of attenuation. They subsequently describe the different methods of information leakage rather theoretically.

The authors in [10] refer to a previous paper regarding the attenuation caused by macrobends. They perform an experiment on the sensing of radiation emanating from an optical fiber in the case of bends with very small radius. They then process the captured radiation in terms of signal characteristics and compare the individual output.

Author in the paper [11] discuss ways to improve the security of passive optical networks (PONs) and perform spectrometric measurements. Information on the use of crosstalk can also be found in the literature [12], but unfortunately without any verification.

III. Methodology

The phenomenon that enables transmissions in optical fibers is called total internal reflection (TIR). This is the phenomenon where a wave incident on an interface with two different refractive indices at more than a critical angle is not radiated into the second medium but reflected back into the first medium, thus propagating the radiation along the length of the fiber as seen in the Figure 1.

The calculation of the critical angle can be expressed by the equation 1,

$$\theta_c = \arcsin(n_2/n_1) \tag{1}$$

where n_2 denotes the refractive index of the core and is also greater than n_1 which denotes the refractive index of the cladding.

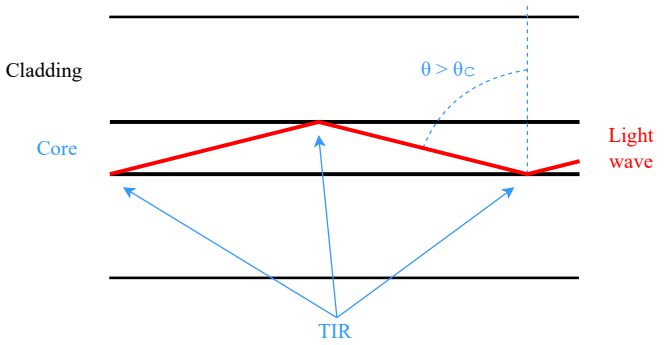


Fig. 1. Total internal reflection.

However, if the beam is incident at an angle less than the critical angle, part of the light is emitted into the cladding and further out of the fiber, increasing the path attenuation. We refer to this phenomenon as macro-bend loss, as is shown in Figure 2. This loss can occur naturally when installing fiber with small radius bends or due to the previously mentioned eavesdropping when macro-bend is produced by an attacker in order to attack the infrastructure. The assumption is that the fewer layers of protection a fiber contains, the more light is emitted from the fiber.

IV. Experimental measurement

The OPT130 optical coupler is used to create a temporary connection between an existing single-mode link and another single-mode SC/APC pigtail. According to the manufacturer, the device is designed for 250 μm fibers with transparent protection. The method of coupling is non-destructive and is also not dependent on the location of the application, apart from the need to use the fiber. Primarily it serves as a simple maintenance and testing tool for optical paths. The experiment was devoted to verifying the method of simple eavesdropping of optical transmission with with G.652 fiber type which is commonly used. The paper also describes method for increasing efficiency of coupling.

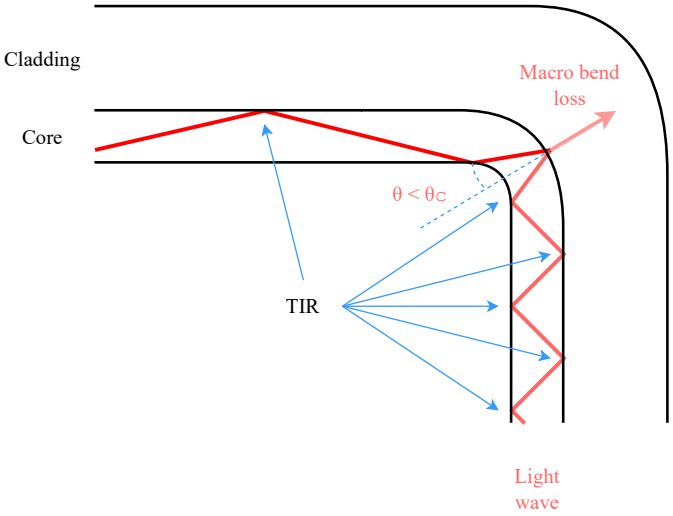


Fig. 2. Total internal reflection.

A. Clip-on coupler evavesdropping

From the datasheet, the attenuation at 1550 nm caused by the coupler should be less than 7 dB and the coupling efficiency between 13 and 22 dB. An integrable tunable laser assembly (ITLA) from Purephotonics was used as lightsource capable of output power up to 13.5 dBm. A 10 km spools of 250 μm transparent fiber type G.652 was used as a tested fiber. For power output measurement the EXFO FPM-600, which is sensitible up to -70 dBm.

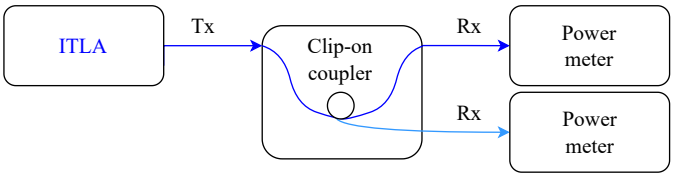


Fig. 3. Schematic of the experimental setup - evavesdropping.

The measurement results showed that the values quoted by the manufacturer are completely erroneous and unattainable. The insertion loss was in the range of 4 dB. The results are shown in Table I. The eavesdropping optical power values cannot be determined accurately as they oscillate in the range of 3 dB.

TABLE I
Evavesdropping values

Fiber type	Power output - data channel [dBm]	Insertion loss [dB]	Evavesdropping output [dBm]
G.652	10.57	4.26	-55

B. Clip-on coupler improvements

In the middle of the measurements there were often large differences between the measured values. It turned out that it is not only important to place the fiber precisely in the coupler,

but also the results are very influenced by the position of the fiber in close proximity before and after the bend created by the coupler. When the tension was increased and the fiber was placed in the correct position, a response in the order of 30 dB higher than the previous measurement was achieved. At the same time, there was the expected increase in insertion loss of the data channel

For this purpose, a special 3D printed holder was designed and developed to hold the fiber in an ideal position to maximize the effectiveness of the eavesdropping. During installation, it is important to pay attention to the orientation with respect to the direction of the eavesdropped traffic. The 3D printed part with attached fiber is shown in Figure 4. At the same time, when installed for a longer period of time, the product allows the fiber to be fixed in order to prevent gradual loosening and thus prevent degradation of the eavesdropping performance.

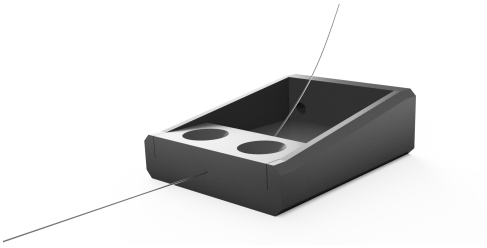


Fig. 4. 3D printed part with fiber.

The Table II shows improvements of the eavesdropping results is shown compared to the values from the previous measurement without the 3D printed holder. It is important to note that each fiber type requires a different placement after exiting the coupler for the best results.

TABLE II
IMPROVED EVAVESDROPPING VALUES

Fiber type	Power output - data channel [dBm]	Insertion loss [dB]	Evavesdropping output [dBm]
G.652	11.01	8.85	-21

V. RESULTS

The measurements showed that the device without modification is capable of eavesdropping with an isolation of around 60 dB for G.652 fiber. After designing and testing an efficiency-improving device, there was a dramatic improvement of up to 30 dB. At the same time, the possibility of transmission distortion through the component was verified and the power transfer characteristics were also verified. The measurements showed that the dependence of the source power and the coupling power to the fiber is constant and the isolation corresponds to 35 dB.

At the same time, it can be said that the repeated use of the clip-on coupler is likely to negatively affect the transmission characteristics of the coupled fiber, given that after a large number of measurements, the performance of the eavesdropped part gradually decreased.

VI. CONCLUSION

This paper describes a method of simple fiber optic eavesdropping and jamming was described using a clip-on coupler that creates a macro bend on a 250 μm diameter transparent fiber. Consequently, it describes possible methods of using the device and also suggests the possibility of improving the efficiency of the device by using a 3D printed part that increases the eavesdropping and efficiency by as much as 30 dB. Therefore, it can be concluded that the existence of attacks using these devices is possible and from a fundamental physics point of view they work correctly. For an effective eavesdropping or infrastructure attack, the attacker will have to be equipped with at least partially more technically advanced equipment. It can therefore be stated that the risk of an attack depends on the equipment of the attacker and is therefore low in most cases.

REFERENCES

[1] B. Fröhlich, M. Lucamarini, J. F. Dynes, L. C. Comandar, W. W.-S. Tam, A. Plews, A. W. Sharpe, Z. Yuan, and A. J. Shields, "Long-distance quantum key distribution secure against coherent attacks," vol. 4, pp. 163–167, Jan 2017.

[2] V. Zapatero and M. Curty, "Secure quantum key distribution with a subset of malicious devices," *npj Quantum Inf*, no. 26, pp. 1–8, 2021.

[3] M. Furdek, N. Skorin-Kapov, S. Zsigmond, and L. Wosinska, "Vulnerabilities and security issues in optical networks," in *2014 16th International Conference on Transparent Optical Networks (ICTON)*, (Graz, Austria), pp. 1–4, IEEE, 2014.

[4] M. P. Fok, Z. Wang, Y. Deng, and P. R. Prucnal, "Optical layer security in fiber-optic networks," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 725–736, 2011.

[5] M. Shah, S. Gala, P. Doshi, V. Venkataramanan, and R. Shah, "Security management in optical fiber," in *2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, pp. 1–4, 2022.

[6] D. Dahan and U. Mahlab, "Security threats and protection procedures for optical networks," *IET Optoelectronics*, vol. 11, 05 2017.

[7] N. Skorin-Kapov, M. Furdek, S. Zsigmond, and L. Wosinska, "Physical-layer security in evolving optical networks," *IEEE Communications Magazine*, vol. 54, no. 8, pp. 110–117, 2016.

[8] A. Akpan, "The vulnerability of fiber- optics communication systems: The role of optical tapping.," *Journal of Information Engineering and Applications (JIEA)*, vol. 4, pp. 145–153, 01 2014.

[9] M. Zafar Iqbal, H. Fathallah, and N. Belhadj, "Optical fiber tapping: Methods and precautions," in *8th International Conference on High-capacity Optical Networks and Emerging Technologies*, pp. 164–168, 2011.

[10] Yilmaz, Kaan, A. Deniz, and H. Yuksel, "Experimental optical setup to measure power loss versus fiber bent radius for tapping into optical fiber communication links," in *2021 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, pp. 1–6, 2021.

[11] A. Harris, A. Sierra, S. V. Kartalopoulos, and J. J. Sluss Jr., "Security enhancements in novel passive optical networks," in *2007 IEEE International Conference on Communications*, pp. 1399–1403, 2007.

[12] R. Shuai, Z. Yinfa, X. Guijin, R. Xuejun, and W. Xiaoshuang, "Research on security protection of key information infrastructure in optical networks," in *2019 IEEE 11th International Conference on Communication Software and Networks (ICCSN)*, pp. 442–445, 2019.

Multifunction Speaker with Wireless Connectivity

1st Michal Baránek
Department of Radioelectronics
Brno University of Technology
Brno, Czech Republic
xbaran17@vutbr.cz

2nd Jan Kufa
Department of Radioelectronics
Brno University of Technology
Brno, Czech Republic
kufa@vut.cz

***Abstract*—The topic describes a multifunction system with multiple choices to transmit the audio signal to the speaker. Emphasis is placed on the competitiveness and higher quality of the final product compared to today's highly popular commercial portable speakers. The product uses wireless technologies such as Bluetooth using A2DP and AVRCP for audio transmission and convenient control or Wi-Fi for audio streams from an online HTTP. An SD card is used for local playback, from which the data are decoded and streamed to the speaker. The audio processing is used for compensating the size of the enclosure, balancing the characteristics of the loudspeakers, and human hearing. Since the product is portable thanks to battery power with all protections such as BMS (Battery Management System) and safe charging, it is equipped with energy-saving mechanisms such as sleep control of the power amplifier whenever the speaker is muted, as well as efficient components such as a class D amplifier for maximum utilization of accumulated energy. All controls are simple and intuitive, and the speaker has enough status indicators to keep the user informed.**

***Index Terms*—ESP32, multifunction speaker, bluetooth, audio, internet radio, sd card**

I. Introduction

Currently, several manufacturers offer high-quality multifunction portable speaker systems with wireless connectivity. These speakers must meet a range of features, from long battery life to great audio output. With growing competition and new technology, these speakers are constantly being improved and are at the forefront of the best-selling speaker systems.

In this work, an illustrative procedure for processing such a portable speaker is captured. The multifunctional speaker has features such as wireless communication via Bluetooth, a slot for a microSD card, which is able to play music locally on the speaker, also the radio function via Wi-Fi with multiple accessible radio stations, and its unique design [1].

The paper is organized as follows: choosing suitable building blocks for the project in Section II. Software part development procedures are outlined in Section III. The development, selection, and subsequent measurement of hardware components are described in Section IV. The final implementation of the project together with a comparison of the multifunction speaker with the competitors are in Section V, and finally, concluding remarks are in Section VI.

Research described in this paper was supported by the Internal Grant Agency of the Brno University of Technology under project no. FEKT-S-23-8191.

II. Building blocks

A. Selection of a hardware

When designing a project with the ability to process audio signals or audio data, a subset of the following components are considered, such as inputs like storage media, e.g. microSD card with audio files to read, a Wi-Fi interface to obtain an audio data stream from the internet, Bluetooth interface to obtain an audio data stream from a Bluetooth headset, user interface e.g. buttons or some other means to provide user input, and the outputs, such as analog signal output for power amplifier input or speakers. Then more deeply I2S interface to stream audio data to a codec chip, and the main processing unit, which is a microcontroller with processing power to read the data from the input, processes, and send them to the output.

On this basis, the low-power system on a chip microcontroller ESP32 with integrated Wi-Fi and dual-mode Bluetooth was chosen for this work which ideally meets requirements for Bluetooth, Wi-Fi, and processing power. A class D amplifier was selected for the appropriate amplification method, namely an amplifier from Texas Instruments the TPA3110D2. This amplifier has a performance exceeding 90 % and a sufficiently high power, which makes it an ideal choice for amplifying the audio output from the codec chip. The main power supply is the battery pack consisting of five 18650 Li-ion batteries that supply all the individual hardware components. From this pack, the power output is directly connected to the power amplifier and via the synchronous buck DC/DC converter with built-in low on-resistance power MOSFETs and low quiescent current (BD9E104FJ) that supplies the power branch for the ESP32 microcontroller.

These parameters are shown in the block diagram that is in Fig. 1. for a better idea of the design of the portable multifunction speaker itself.

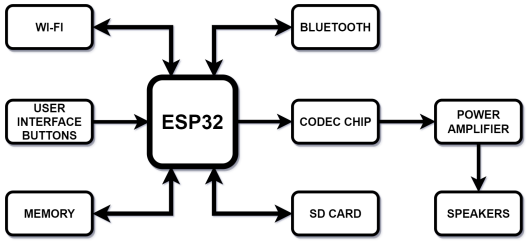


Fig. 1. The block diagram of multifunction speaker

Based on the block diagram, a development board ESP32-LyraT-Mini from Espressif was selected, which perfectly meets all requirements. It has all the peripherals like control buttons, Wi-Fi, Bluetooth, an SD card slot, and other features.

B. Selection of a development framework

Inventing and writing the whole software from scratch would be very demanding and difficult to accomplish. Therefore, a development framework from the manufacturer of the development board is used, namely the Espressif IoT Development Framework, and on top of it the Audio Development Framework. These frameworks provide a way to develop audio applications using Elements like Codecs, Streams, or Audio Processing functions and connecting them to the Audio Pipelines.

III. Software part

A. State and flowchart diagrams

The whole concept stands on FSM (Finite State Machine) which is shown in Fig. 2.

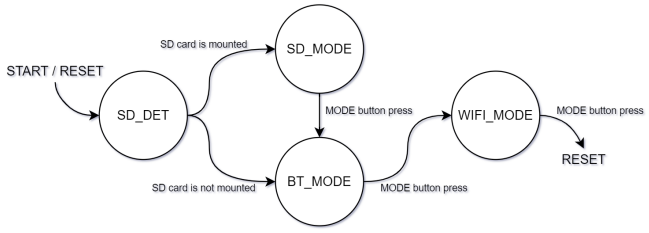


Fig. 2. State diagram of the interaction between individual modes

It is based on the user's decision to move between the various modes such as SD card mode, Bluetooth mode, and Wi-Fi radio mode. Each mode is composed of an Audio Pipeline that consists of elements from the Espressif ADF (Audio Development Framework). The control of the basic playback parameters (play/pause, volume settings, switching between tracks, or directly switching between modes) depends on the pressed button by the user.

Each mode is announced by their specific system prompt tone file, which is stored in the ESP32 flash memory. The audio announcements were created via a text-to-speech API from Google Cloud, and it was necessary to convert these announcements from MP3 to a binary file using a Python script provided by Espressif.

B. Audio Pipeline

As discussed before, each mode consists of an audio pipeline. This audio pipeline is constructed of the elements from ADF. To better understand the operation of the modes, the block structure of the audio pipelines is shown in Fig. 3.

C. Amplifier shutdown control

The selected amplifier has a shutdown mode function in which the amplifier is turned off and the quiescent current is kept at approx. 250 μA. The application also includes this SHUTDOWN control of the power amplifier using pin IO22 of ESP32.

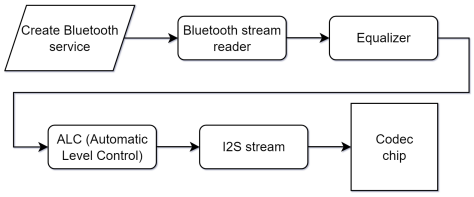


Fig. 3. Audio pipeline example of Bluetooth mode

D. Audio processing

The processing of the audio signal is done by equalization and ALC (Automatic Level Control). It was necessary to add equalization to the system to compensate for the size of the enclosure, the properties of the loudspeakers, and the frequency response of the power amplifier. The equalization levels are adaptive based on player volume and considering properties of human hearing. The levels are set as shown in Fig. 4.

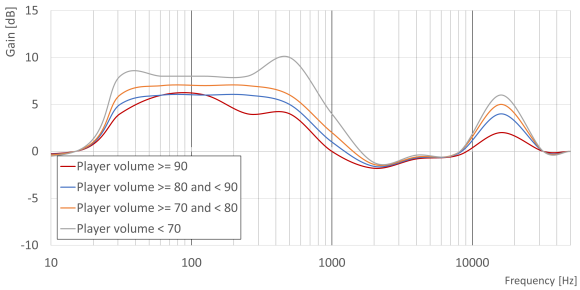


Fig. 4. Equalization frequency levels

The ALC detects the amplitude of the audio signal and determines whether it exceeds the preset maximum value. If it exceeds the range, the amplitude of the audio signal will be adjusted according to the set gain to make it within a reasonable range [2]. This feature is added to protect the amplifier and loudspeakers from overdriving and damage.

IV. Hardware part

A. PCB layout

The final PCB layout consists of the power amplifier TPA3110D2, synchronous buck DC/DC converter BD9E104FJ, Jack-Sensing circuit, and battery indicator circuit as can be seen in Fig. 5.

B. Power amplifier measurements

The measurement of the TPA3110D2 amplifier was done by the following setup shown in Fig. 6.

The first measurement was to measure the maximum output power for limitation (all parameters are measured in RMS values). The maximum power, where the waveform is still without visible distortion (at THD+N = 1 %, V_{IN} = 21 V, and amplifier gain set at 20 dB) was 11 W and with the value of THD+N = 10 % that manufacturers like to quote was power 16 W. This is a case when is the battery pack fully charged.

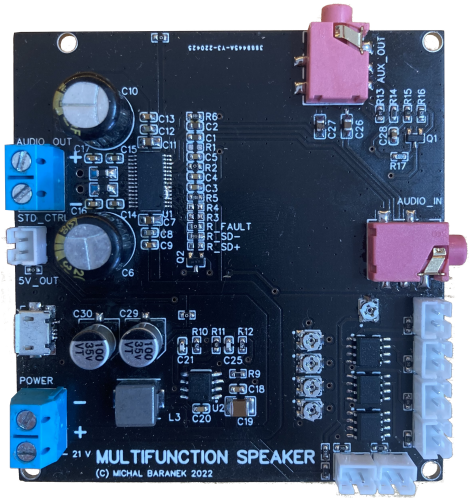


Fig. 5. Final PCB layout

The maximum power with a bottom threshold of the battery pack, where the waveform is still without visible distortion (at THD+N = 1.67 %, V_{IN} = 16 V, and amplifier gain set at 20 dB) was 5.6 W and with the value of THD+N = 10 % was power 8 W. Subsequently, efficiency measurements were performed. The efficiency of this amplifier is 81 % (at V_{IN} = 21 V and P_{OUT} = 11 W). This efficiency corresponds to information in the datasheet of the component.

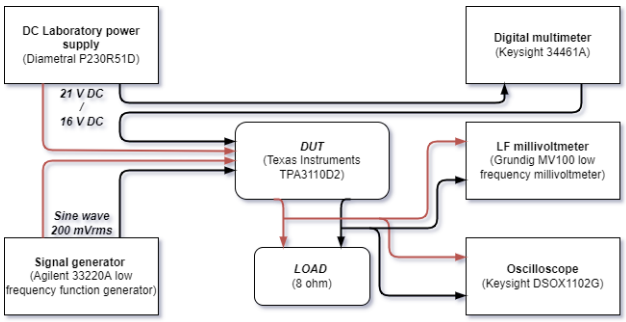


Fig. 6. Block diagram and connection of measurement conditions

The last measurement was the frequency response of the TPA3110D2 amplifier which is illustrated in Fig. 7.

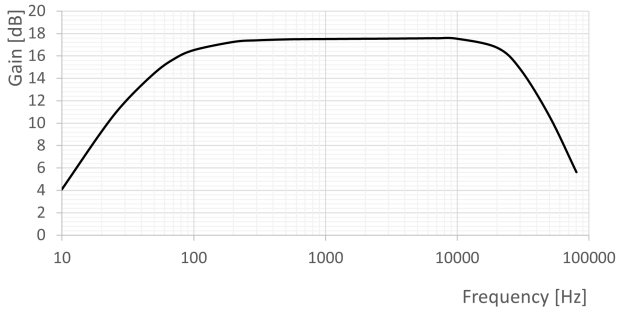


Fig. 7. Measured frequency response of TPA3110D2

The frequency response should be constantly at 20 dB, but in this case, it is at 17.5 dB. This decrease and other problems related to measured data may be due to the chip not being original. The lower limit of the 3 dB drop is at 70 Hz and the upper is about 27 kHz where the output LC low pass filter is applied.

C. Battery and power consumption

The entire hardware had very similar current consumption across the entire 16 to 21 V supply range of the battery pack. In an idle mode without the amplifier shutdown mode enabled are consumption 50 mA (otherwise 25 mA), 60 mA at low volume 100 mA at medium volume, and 250 mA at full volume with up to 500 mA peaks. These values are very dependent on the mode used. In SD mode is the lowest and in Bluetooth or Wi-Fi mode, the consumption is higher by about 15 mA due to the use of the radio.

On this basis, Li-Ion 18650 3.7 V/2000 mAh 3 C batteries from MOTOMA were selected. Five of these batteries should keep the multifunction speaker running at the highest volume of approx. 7 hours, at the medium volume about 20 hours, and at the lowest volume up to 30 hours. These five batteries are connected in series to a final voltage of 16 to 21 V and protected by a BMS for safe use.

For safe and smooth charging, a Buck/Boost converter with a function to limit the output current was added to the multifunction speaker, which is a highly sought-after function for charging Li-Ion batteries. This module is set to a maximum output current of 500 mA and an output voltage of 21 V. Thanks to this module, it is also possible to charge the multifunction speaker with a large voltage range, from 4.8 to 30 V (the output power of the supply source must be at least 11 W). Charging from a fully discharged battery pack to a fully charged takes about 4 and a half hours (tested).

D. Loudspeakers and crossover selection

The multifunction speaker's enclosure is taken from TVM ARS-204-50/8 which was originally a central component in the 6.1 audio system. This center speaker is equipped with two 8-ohm TVM ARN-100-10/8 midrange speakers and one 8-ohm TVM ARV-104-61/8 tweeter. Since the audio from the midrange speakers wasn't acceptable, these speakers were replaced by the 4-ohm AIYIMA mid-range speakers with significantly better frequency response in the lower band of the audible spectrum. This frequency response can be seen in Fig. 8.

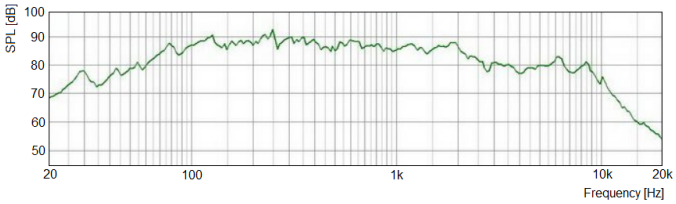


Fig. 8. Frequency response of the 4-ohm AIYIMA mid-range speaker [3]

The 8-ohm TVM ARV-104-61/8 tweeter was kept. The frequency response is in Fig. 9.

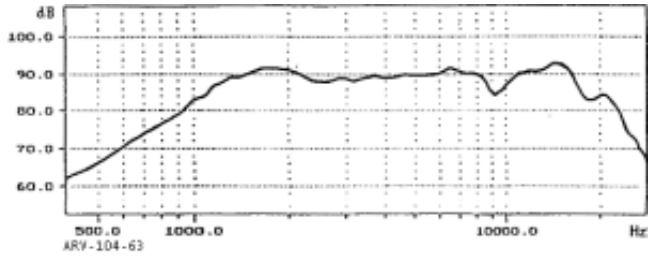


Fig. 9. Frequency response of the 8-ohm TVM ARV-104-61/8 tweeter [4]

The crossover, which consists of two LC filters, remained the same as it was in TVM ARS-204-50/8. An LC high pass filter with a cut-off frequency of 2.3 kHz for the tweeter. And for mid-range speakers used LC low pass filter with a cut-off frequency of 2.8 kHz.

V. Final result

Fig. 10. shows the inner part of the multifunction speaker with all the parts already discussed and installed. The advantage is that all parts are replaceable. Thus, for example, replacing the battery pack is not a problem (or any other part). The separability of the functional parts has been maintained for easy future improvements. For better vibration resistance, all parts were installed on rubber or foam pads and glued.

A particleboard enclosure is used for the safe storage of all hardware parts. This enclosure is recycled from the old enclosure of TVM ARS-204-50/8 with bass reflex.

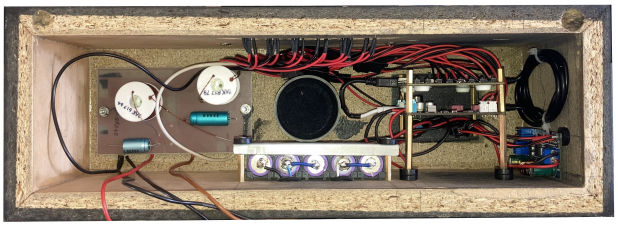


Fig. 10. Inside of multifunction speaker

The rear panel has been designed for easy access to features like an SD card, headphone output, and power input for charging and switching on/off the multifunction speaker.

Another integral part is the top control panel with buttons and their functions marked with simple symbols such as play/pause, next, previous, etc. In this panel are buttons that have integrated LEDs. These LEDs are connected to the battery indicator circuit and show the user an overview of how the battery pack is charged.

To protect the entire multifunction speaker, protective feet have been designed. These feet are suitable for any placement of the multifunction speaker (it can stand both vertically and horizontally). For convenient transmission of the multifunction speaker, a holder is installed on the top of the enclosure, which fits perfectly into the entire design of the multifunction speaker, as can be seen in Fig. 11.



Fig. 11. The final multifunction speaker

The comparison of the multifunction speaker with the best-selling speakers is in Table 1.

TABLE I
Parameters of multifunction speaker and best-selling speakers from the e-shop alza.cz

Name	Features	Power	Batt. life	Weight	Price
Multifunction speaker	Wi-Fi,BT, SDcard, Radio	11W	20hrs	5.7kgs	140€
JBL Extreme 3	BT	50W	15hrs	1.8kgs	300€
Bang & Olufsen Beolit 20	BT	70W	20hrs	2.7kgs	490€
Gogen SMILEE BPS 636	BT,Radio	40W	6.5hrs	5.2kgs	130€

In terms of multifunctionality, battery life, and price, this project is surely a better choice. On the contrary, it stands behind in weight and power. However, the manufacturers do not state under what conditions the speakers are tested. Therefore, it is difficult to compare these parameters unless it is determined under what conditions the manufacturers state them.

VI. Conclusion

This paper does not discuss all the procedures and problems that had to be overcome to complete this project (more on them in [1]). An example of functionality can be found on YouTube[1]. The main aim was to make a modern portable speaker and it was successfully achieved. The final result has all the required features and the final price of all the necessary components was 140 €. This project is still under development and new features and improvements are planned.

References

[1] BARÁNEK, Michal. Multifunction speaker with wireless connectivity. Brno, 2022. Available from: https://www.vutbr.cz/studenti/zav-prace/detail/141495. Bachelor's thesis Brno University of Technology, Faculty of electrical engineering and communications, Dept. of telecommunications, 2022. 84 p. 12 p. of attachments.

[2] ESPRESSIF, Adjust wav volume with alc [online]. 2023 [cit. 2023-03-08] Availible from: https://github.com/espressif/esp-adf/tree/master/examples/audio_processing/pipeline_alc

[3] ALIEXPRESS, AIYIMA 4 inch 4ohm Subwoofer Speaker [online]. 2023 [cit. 2023-03-08] Available from: https://ae01.alicdn.com/kf/H7e91cda414174ac38086a238e432b0cfk.jpg

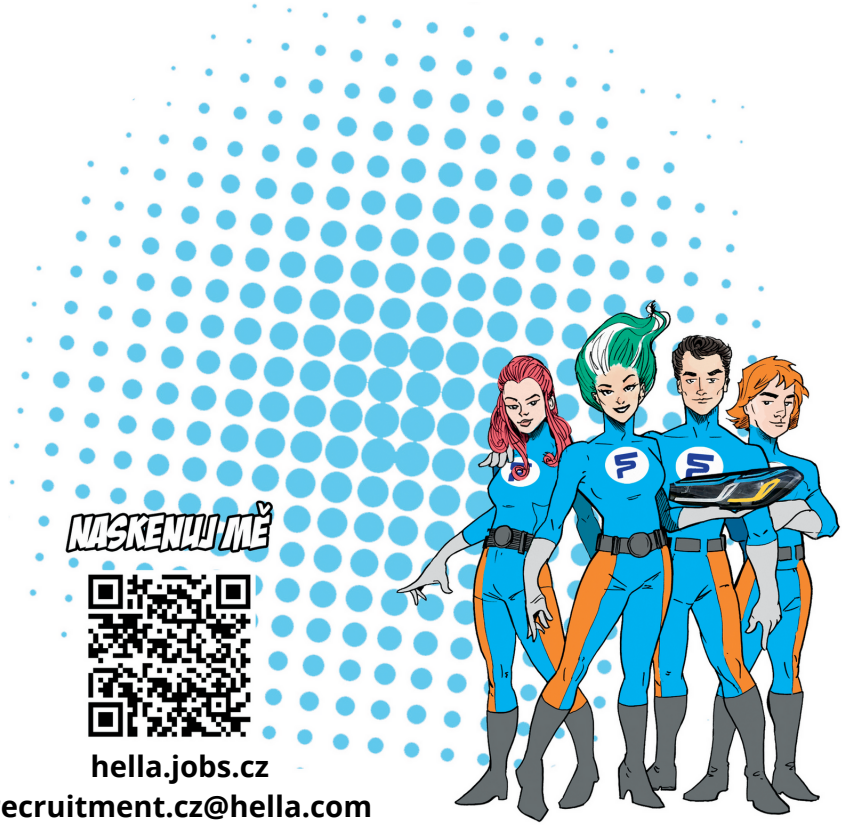
[4] ANGELICAAUDIO, TVM ARV 104-63/4 [online]. 2023 [cit. 2023-03-08]. Available from: http://www.angelicaaudio.cz/tesla/reproduktory-tvm/tvm-arv-104-634/

[1] https://www.youtube.com/watch?v=J5AC4qzRAgA

STAŇ SE SOUČÁSTÍ NAŠEHO

TRAINEE PROGRAMU

Pojď s námi vyvíjet, testovat a vyrábět světlomety
a elektroniku pro automobilky prémiových značek.
V Mohelnici, Olomouci nebo Ostravě.



hella.jobs.cz
recruitment.cz@hella.com

FORVIA
HELLA

KDE SE MŮŽEŠ U NÁS V RÁMCI **TRAINEE PROGRAMU** UPLATNIT?

IT
MES systémy, JAVA vývojáři, Front-end vývojáři analytici s SQL, aj.

elektronika
embedded software, hardware a PCB layout

optika
optické simulace a výpočty, měření světelných vlastností

výroba
lisování a povrchová úprava plastů, procesní inženýrství v rámci montáže, technický servis a automatizace + robotizace

testování
životnostní a environmentální testy elektroniky a celých světlometů, testy elektromagnetické kompatibility atd.

konstrukce
konstrukce plastových dílů nebo strojního zařízení, simulace (CFD, Moldflow, FEM)

projektové řízení a další podpůrné role
HR, logistika aj.



V případě jakýchkoliv dotazů k trainee programu jsem k dispozici.

Hana Zámečníčková

hana.zamecnickova@forvia.com

+420 770 157 490

hella.jobs.cz

PPG QUALITY EVALUATION

1st Pavlína Sikorová
Brno University of Technology
Department of Biomedical Engineering
Brno, Czech Republic
xsikor19@vutbr.cz

2nd Lukáš Smital
Brno University of Technology
Department of Biomedical Engineering
Brno, Czech Republic
smital@vut.cz

***Abstract*—This paper addresses the issue of the difficulty of assessing the quality of photoplethysmographic (PPG) signals. Three different approaches for PPG signal quality evaluation were implemented in Python to select the best field for the development of a new classifier: a trend-based approach that assesses signal quality based on skewness and maximum autocorrelation (accuracy 68%), a random forest classifier assigning quality labels according to seven different features (accuracy 77%) and a 1D deep convolutional neural network which evaluates signal quality directly from the signal (accuracy 85%).**

***Index Terms*—PPG, photoplethysmography, signal quality assessment, deep learning, DCNN, deep convolutional neural network, random forest classifier**

I. Introduction

For many years, electrocardiography (ECG) has been considered the dominant technique and the gold standard in cardiac monitoring [1]. However, despite numerous improvements and upgrades over the years, this technology has not yet reached a state where it offers users the required flexibility, portability and convenience [2].

An alternative to the ECG method may be PPG [1], [2], which is commonly used to measure blood oxygen saturation levels [3], but the signals also carry cardiac information that can be derived and which is more than comparable to the cardiac information contained in ECG signals. Moreover, PPG has other advantages, such as simplicity in operation, convenience for the user and also its cheapness [2].

The disadvantage of this method is that it is highly susceptible to artifacts, which are unfortunately very common in these measurements. These artifacts then affect the accuracy of heart rate (HR) estimation [2] and other possible diagnostic or health assessment information.

A solution to the problems outlined above, which prevent PPG from becoming more widely used in medicine and replacing ECG, could be a method that somehow evaluates the recorded PPG signal and decide whether its quality is sufficient for the given task. This would avoid potential errors (caused by using low quality signals corrupted by noise and other common artifacts instead of omitting them from analysis) [4].

II. Methods for quality assessment

A. Trend-based approaches

The simplest methods are based on classical signal processing techniques. They search for regularity and irregularity in shorter segments of the signal. They are based on the expectation that the clean signal consists of morphologically similar pulse waves. Any kind of noise or interruptions causes distortions in the regularity and similarity [4], [5]. This can be measured and described via skewness (1) (symmetry of the distribution of PPG values [4], [6]), kurtosis (2) (sharpness of the peak of the distribution) and Shannon entropy (3) (measure of the information and uncertainty of the system [4], [6]).

$$\hat{S} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x})^3}{\hat{\sigma}^3}, \tag{1}$$

$$\hat{K} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x})^4}{\hat{\sigma}^4}, \tag{2}$$

$$SE = \sum_{i=1}^{N}(x_i)^2 \log_e(x_i)^2, \tag{3}$$

where x_i is the PPG signal, $\hat{\sigma}$ is the standard deviation of the distribution of the signal, $\hat{x}$ is the mean value of the PPG signal, and N is the number of samples [4].

Low-quality signals have lower values of skewness and kurtosis and Shannon entropy is higher – for high-quality signals it is the contrary [4], [6].

B. Machine learning approaches

Machine learning (ML) is useful in quality assessment because the relationships between available features and corresponding labels are often unknown or not fully understood. These methods only need to learn how to replicate the labeling pattern suggested by the training dataset. Moreover, the process of assigning the correct class need not to be linear (as in simpler approaches, e.g. trend-based). ML approaches use a combination of multiple features to solve problems. In the case of signal quality classification, the output is the quality class to which the signal belongs. Compared to the trend-based approach, which works on much simpler principles, ML-based methods combine different features during the learning process in a more sophisticated way and the final decision is more robust and reliable [4].

Commonly used ML methods are: *k-Nearest Neighbour, Multi-class SVM, Naïve Bayes, Decision tree, Random forest*, of which the best results are usually achieved by the random forest classifier [7].

C. Neural networks and deep learning approaches

Neural networks (NNs) are one branch of machine learning algorithms. Simplistically, artificial NNs simulate the principle of a biological NN. They are composed of multiple layers of artificial neurons called perceptrons. Such a network needs to be trained on a training dataset to learn dependencies and relationships in the data, so that it can adapt in order to correctly classify newly introduced data [8], [9].

The complexity of the network increases as more hidden layers are added to the network. If there are more than two hidden layers, the model is referred to as a deep neural network (DNN) and the training process becomes deep learning (DL) [8].

The range of possible NN architectures is very wide, the use of a particular one depends on the problem to be solved. Commonly used architectures include *Feedforward, ResNet, Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), Encoder-Decoder and Autoencoders* [8].

III. PPG DATASET

The PPG dataset was collected at the Department of Biomedical Engineering (UBMI) of Brno University of Technology (BUT) by students and subsequently processed by the UBMI cardiology team, which created the final database.

The database consists of 3894 PPG signal segments (10 seconds, sampling frequency 30 Hz) obtained accordingly to a special experimental protocol, using a smartphone to acquire the data. Each PPG recording includes a quality annotation created by biomedical experts using a reference HR from an ECG that was recorded simultaneously. The quality index is binary:

- Label 0: low-quality signals from which the HR estimation is not reliable (not suitable for further analysis)
- Label 1: good-quality signals from which the HR estimation is reliable (suitable for further analysis)

For each implemented model, the dataset was divided into two subsets – 80% for training and 20% for testing. Since the dataset was highly imbalanced (only 24% high-quality signals), the dataset division was done with respect to this problem, while maintaining the imbalance. For each classifier, it was treated accordingly to the method.

IV. IMPLEMENTED CLASSIFIERS

From all fields of possible signal classification, one selected method (the one achieving the best results accordingly to works of other authors [4], [7], [10]) was implemented in order to compare different approaches and decide which area is the most promising for the development of a new method.

A. Trend-based classifier

For each database record, several characteristics describing the regularity and irregularity of the signal were calculated: *skewness, kurtosis in frequency and time domain, Shannon entropy, maximum autocorrelation* [4]–[6].

For all features, the optimal threshold for class separation was calculated via the OTSU method. However, only *skewness* and *maximum autocorrelation* showed some ability in signal classification. Each descriptor was used alone for classification via thresholding and only these two did not label all signals as low-quality ones.

The final quality estimation was done via majority vote – if both selected descriptors marked the signal as high-quality, it was labeled with 1, otherwise it was labeled with 0.

B. Random forest classifier

The first step was to extract multiple features from each signal to represent it during the classification. Features were extracted from both the raw PPG and detrended PPG (baseline wander removed) – there were 16 different features in total (number of peaks, mean peak distance, standard deviation, mean, signal range, kurtosis in time and frequency domain, skewness, perfusion, number of zero-crossings of the original/once derived signal, SNR, entropy, signal energy, energy of a frequency band 0.25-4.00 Hz, relative energy of the frequency band).

Using a wrapper method *backward elimination*, a subset of features was selected: number of zero-crossings, mean peak distance, number of peaks, Shannon entropy, SNR, signal energy, relative energy of the frequency band 0.25-4.00 Hz.

To define the hyperparameters of a random forest classifier (RFC), both *grid* (GS) and *random search* (RS) were used to find the optimal values for model building. The *GS* limited the search space of the other method. The model was then trained on the training dataset represented by pre-selected features.

To address the imbalance problem, the majority class of the dataset was randomly undersampled (randomly picking samples of the majority class without replacement).

C. 1D deep convolutional neural network

Using the architecture proposed by [10], a 1D deep convolutional neural network (1DDCNN) was built. The properties of the network were specified via Bayesian optimization. The 1DDCNN consisted of 6 convolutional (CNN) layers (32 kernels of size 10×1, 64 kernels of size 6×1, 128 kernels of size 4×1, and the residual 3 layers had all 256 kernels of size 2×1) followed by 2 fully-connected (FC) layers (1546 and 2 nodes). CNN layers were all using He initialization and ReLu activation, the mutual stride was 2. Each layer's output was normalized via batch normalization layer and the 6th layer was followed by a max-pooling (maxPool) layer. The first FC layer had ReLu activation and the output FC layer had Softmax activation. The architecture is pictured in Fig. 1.

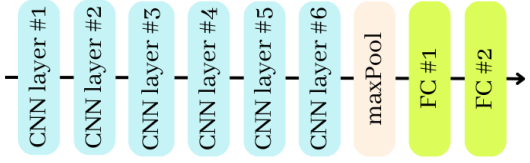


Fig. 1. 1DDCNN architecture proposed by [10]

The problem of the imbalanced dataset was solved by applying weights to each class during network training. The weights for both classes were calculated as follows, using a *sklearn* library in Python:

$$w_{class0} = \frac{N_{samples}}{N_{classes} \cdot N_{class0}} = 0.6645, \quad (4)$$

$$w_{class1} = \frac{N_{samples}}{N_{classes} \cdot N_{class1}} = 2.0195, \quad (5)$$

where $N_{samples}$ is number of samples in the training dataset, $N_{classes}$ is number of classes, and N_{class0} and N_{class1} is number of samples from class 0 and class 1, respectively.

For the model generation, stratified 5-fold cross validation was used. The training data was divided into five subsets of the same size, ensuring equal class-distribution in all subsets. In each run, a model was created and the average accuracy and validation loss per fold was assessed – the best model from each fold was saved and used for model ensemble for signal classification. The majority vote of the sub-model ensemble decided whether the signal was of high or low quality.

V. Results and Discussion

The trend-based approach was based on the calculation of selected features, their thresholding and subsequent fusion of the results to determine the signal quality. As expected, the results of this most naive approach were really poor, achieving an overall accuracy of 71% on the training dataset and 68% on the testing dataset – see confusion matrices in Fig. 2 and scores in Tables I and II. However, the significant imbalance of the dataset must be also considered when assessing performance of the model.

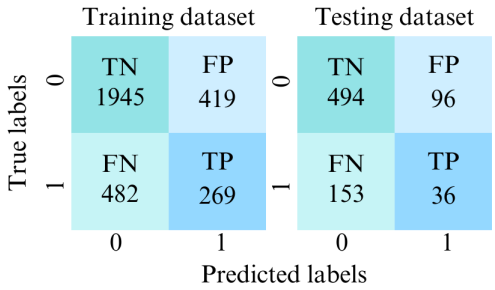


Fig. 2. Results of the trend-based approach via a confusion matrix.

A more sophisticated approach, the RFC, provides more reliable results (Fig 3). Although the model accuracy was 81% for the training dataset and 77% for the testing dataset, the precision and sensitivity showed considerably lower values than desired. However, it is still more reliable than the previous approach – Tables I and II.

The last implemented model was the 1DDCNN. The final model was an ensemble of five models obtained from a five fold cross-validation during training. After the ensemble, the overall accuracy was 88% for the training dataset and 85% for the testing dataset. Based on the scores in Tables I and II, 1DDCNN shows better reliability and performance on both

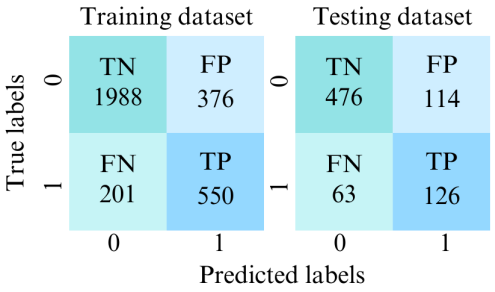


Fig. 3. Results of the ML approach (RFC) via a confusion matrix.

datasets than the two previous approaches. Nevertheless, there is still a big room for improvement, since any model working with medical data should have performance as high as possible to be usable in a health assessment [11].

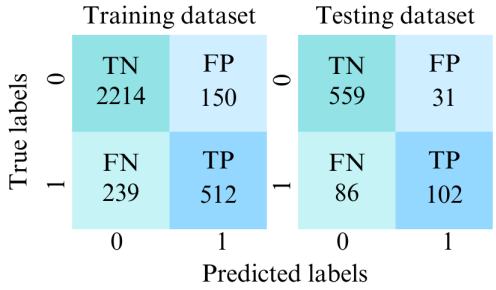


Fig. 4. Results of the DL approach (1DDCNN) via a confusion matrix.

TABLE I
Performance of methods on the same training dataset

Approach	Accuracy	Precision	Sensitivity	Specificity
trend-based	71%	39%	36%	82%
ML	81%	59%	73%	84%
DL	88%	77%	68%	94%

TABLE II
Performance of methods on the same testing dataset

Approach	Accuracy	Precision	Sensitivity	Specificity
trend-based	68%	28%	20%	83%
ML	77%	53%	67%	81%
DL	85%	77%	54%	95%

To compare the performance of the models, their execution time was calculated (dataset preparation included) – see Table III. The computer (Dell Latitude 5410) properties are as follows: *Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz, GPU UHD Graphics, RAM(DDR4) 8GB, SSD 256GB, OS Windows 11 Pro 64-bit.*

The fastest and least computationally expensive model was the trend-based model. However, its performance was poor in comparison to the others, so computational inexpensiveness is not a defensible reason to prefer this method over the others.

The execution time of the RFC was the highest, but it can be traded for its classification performance. The long execution

time is due to the computation of features that represent the input signal.

The DL approach is slower than the trend-based approach, but it outperforms it immensely in terms of classification accuracy. It is faster than the RFC – therefore, the 1DDCNN is the best model even based on the comparison of their execution time.

TABLE III
EXECUTION TIME OF IMPLEMENTED METHODS ON A SET OF 779 SIGNALS

Approach	Total elapsed time [s]	Average per signal [s]
trend-based	0.2337	0.0003
ML	2.0254	0.0026
DL	2.2591	0.0029

The models also differ in their requirements for dataset pre-processing. The trend-based approach calculates two higher-order statistics which then represent the signal. No prior knowledge about the signal nor further signal pre-processing is needed.

The RFC extracts seven different features and the sampling frequency of the signal must be known. Similar to the simplest approach, the signal can be arbitrarily long (however, both methods have been implemented when working with signals of 300 samples (10 seconds)).

The 1DDCNN works only with signals of a given length of 300 samples, as this has been set as a hyperparameter of the input layer of the network. Newly introduced signals must therefore be reshaped, resampled or split into shorter segments. However, this is the only mandatory pre-processing step. It would be advisable to test the 1DDCNN on signals that had to be resampled in terms of length or number of samples before entering the network to determine its ability to classify any PPG signal.

All methods, especially the 1DDCNN, would benefit from working with a larger training dataset than the available dataset, which consisted of only 3894 signals, which in most applications is not sufficient to train a well-performing and more sophisticated model with a good generalization on the newly introduced data. In addition, more high-quality signals need to be recorded to avoid using undersampling methods that further reduce the already small dataset, or data augmentation methods to address the problem of class imbalance.

Another thing that would increase the reliability of the results and the ultimate performance of the models would be the use of independent data in the testing dataset – e.g. recorded using a different phone, a different acquisition protocol and measured on multiple body locations, collected by different experimenters, etc.

VI. CONCLUSIONS

After implementing three different methods from three different fields of signal quality assessment, it was discovered that the newest and also the most complex model (1DDCNN) is able to perform well in the difficult task of PPG signal quality evaluation - with accuracy of 85%, precision of 77%, sensitivity of 54% and specificity of 95%. DL has emerged as the most promising field to solve this issue. Therefore, it is recommended to develop new methods using DL.

REFERENCES

[1] H W Loh, et al. Application of photoplethysmography signals for healthcare systems: An indepth review. Computer Methods and Programs in Biomedicine, Volume 216, 2022, 106677, ISSN 0169-2607, https://doi.org/10.1016/j.cmpb.2022.106677.

[2] M Ghamari, K Singh, A Shankar. A review on wearable photoplethysmography sensors and their potential future applications in health care: A survey. 2017 International Conference on Recent Innovations in Signal processing and Embedded Systems (RISE). IEEE, 2018, 2017, **4**(4), 139-141, ISBN 978-1-5090-4760-4, ISSN 25732838, https://doi.org/10.15406/ijbsbe.2018.04.00125

[3] M A Almarshad, , M S Islam, S Al-Ahmadi and A S Bahammam. Diagnostic Features and Potential Applications of PPG Signal in Healthcare: A Systematic Review. Healthcare. 2022, **10**(3), R1-R39. ISSN 2227-9032, urlhttps://doi.org/10.3390/healthcare10030547

[4] C Orphanidou. Signal Quality Assessment in Physiological Monitoring: State of the Art and Practical Considerations. 2018. Springer Cham [cit. 2022-10-07]. ISBN 978-3-319-68415-4. https://link.springer.com/book/10.1007/978-3-319-68415-4

[5] Q Li, R G Mark and G D Clifford. Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. Physiological Measurement. 2008, **29**(1), 15-32 [cit. 2022-11-17]. ISSN 0967-3334, https://doi.org/10.1088/0967-3334/29/1/002

[6] A Mahmoudzadeh, I Azimi, A M Rahmani and P Liljeberg. Lightweight Photoplethysmography Quality Assessment for Real-time IoT-based Health Monitoring using Unsupervised Anomaly Detection. Procedia Computer Science. 2021, **184**, 140-147 [cit. 2022-10-22]. ISSN 18770509, https://doi.org/10.1016/j.procs.2021.03.025

[7] N Pradhan, S Rajan, A Adler. Evaluation of the signal quality of wrist-based photoplethysmography. Physiological Measurement. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2019, 2004-04-22, **40**(6), 1491-1501 [cit. 2022-10-22]. ISBN 978-0-89871-568-2. ISSN 1361-6579. https://doi.org/10.1088/1361-6579/ab225a

[8] J Allen, P Kyriacou. Photoplethysmography: Technology, Signal Analysis and Applications. Academic Press, 2021. ISBN 9780128235255.

[9] A Panesar, M P Barde. Machine Learning and AI for Healthcare: Standard deviation or standard error of mean?. Perspectives in Clinical Research. 2012, 2021, **3**(3). ISSN 2229-3485, https://doi.org/10.1007/978-1-4842-6537-6

[10] H Shin, X Wu, C Shi, Q Gao and J Geng. Deep convolutional neural network-based signal quality assessment for photoplethysmogram: Implications for Training and Classification Users. Computers in Biology and Medicine. 2022, 2021, **145**(6), 3826-3835. ISBN 978-1-6654-4899-4. ISSN 00104825. https://doi.org/10.1016/j.compbiomed.2022.105430

[11] N Lavrac, B Zupan, I Kononenko, K Matjaz, K Elpida. Intelligent Data Analysis for Medical Diagnosis: Using Machine Learning and Temporal Abstraction. AI Commun., 1998, **11**(3), 191-218

An insight into horizontal gene transfer in *Bacteroidetes spp.* using *in-silico* analysisof operon structures

Julie Nejezchlebová
Department of Biomedical Engineering
Faculty of Electrical Engineering and Communication,
Brno University of Technology
Brno, Czech Republic
xnejez10@vutbr.cz

Jana Schwarzerová
Department of Biomedical Engineering
Faculty of Electrical Engineering and Communication,
Brno University of Technology
Brno, Czechn Republic
Molecular System Biology (MOSYS),
University of Vienna,
Vienna, Austria
Jana.Schwarzerova@vut.cz

Darina Čejková
Department of Biomedical Engineering
Faculty of Electrical Engineering and Communication,
Brno University of Technology
Brno, Czech Republic
cejkovad@vut.cz

***Abstract*— Overuse of antibiotics in human medicine and agriculture has led to the emergence of multi-resistant bacteria, mainly pathogens. In multi-resistant bacteria, individual genes coding for antibiotic resistance are transferred horizontally as a part of operon. Therefore, in order to identify horizontally transferred genes, the prediction of operon units has to be assesses in sequenced genomes. This study is focused on the identifying of genes that are associated with antibiotic resistance as well as accompanied genes simultaneously present in the same operon units. A total of 5 genes were examined in 5 draft genomes *Bacteroides spp*., that were previously extracted from the chicken fecal and cecal samples and sequenced.**

Keywords—Microbiota, In-silico Analysis, Cluster of Orthologous Groups, Operon, Antibiotics Resistance

I. INTRODUCTION

Horizontal gene transfer (HGT) is the transfer of genetic information between different organisms. In bacteria, HGT is responsible for the transfer of antibiotic resistance genes between different microbiota members. However, not only genes coding for resistance to antibiotics are disseminated via HGT, but also genes responsible for resistance to heavy metals, genes directly associated with the transfer and recombination processes, or even genes of unknown function. In general, multi-gene unit represents the core of mobilized genetic element. Genes within such genetic elements are not only simultaneously co-transferred, but they have to also be concomitantly co-expressed in the host. Therefore, it seems that the transferable genes are usually arranged in operons [1].

In our previous studies [2][3], we already identified genes suspected to HGT among chicken gut microbiota. It is assumed that genes belonging to COG1131 (gene coding for ABC-type multidrug transport system, ATPase component), COG1961 (site-specific DNA recombinase SpoIVCA), COG0385 (Na^+-dependent transporter YfeH), COG0582 (integrase/recombinase), and COG4974 (site-specific recombinase XerD) clusters of orthologous groups (COG) [4] are transferred horizontally. However, only sole genes were predicted. Therefore, in the current study, we extended our study on the identification and characterization of the whole operons. As *Bacteroidetes spp*. are common members of chicken gut microbiota and are prone to HGT [1][5][6], we investigated draft genome sequences of 5 bacteria for the presence of transferable operons.

II. MATERIALS AND METHODS

A. Dataset

The dataset represents 5 draft genome sequences of *Bacteroides spp*. – *Bacteroides ovatus* An161, *Bacteroides* sp. An279, *Bacteroides caecigallinarum* An428a. *Bacteroides mediterraneensis* An793 and *Bacteroides mediterraneensis* An801. The samples have been collected from healthy chicken ceca and their gDNA was sequenced on the Illumina NextSeq500 platform in the paired-end modus (2 x 150 bp) [7]. Trimmed paired-end reads were assembled *de novo* assembly [7] All the data are available from NCBI under the BioProject number PRJNA377666 [7].

B. Methods

The whole pipeline of methodology is shown in Figure 1. Firstly, operons were predicted using the online tool Operon-mapper [8]. Operon predictions are based on the intergenic distance of neighboring genes as well as the functional relationship of their protein-coding products [8]. Draft

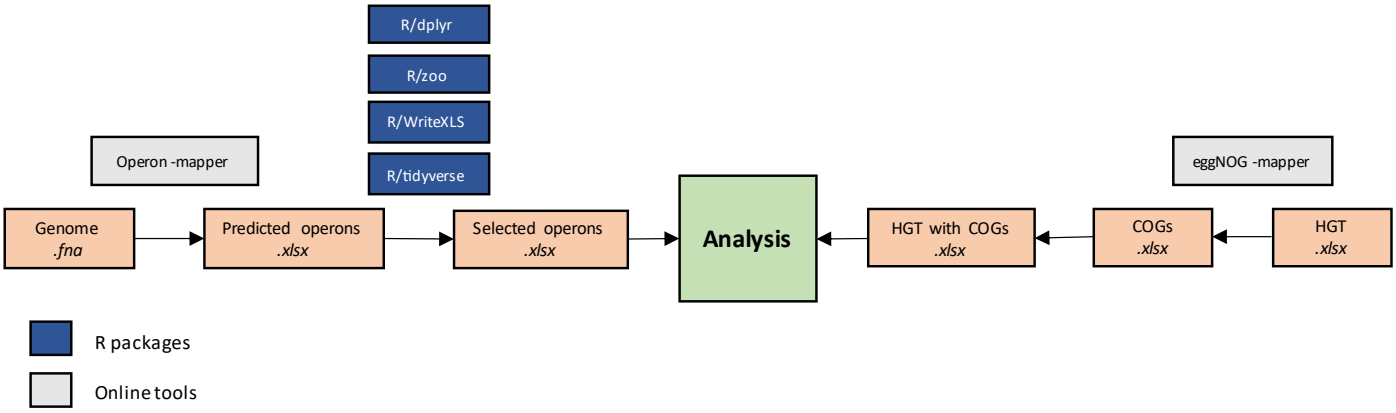


Figure 1: Methodology used in our study. Orange rectangles represent files; grey represent online tools; blue represents R packages.

genomesequences of each individual genome (*.fna* files) represent input file, whereas text files (*.txt*) with *in-silico* identification of operonwere the output files.

In the next step, an R script was constructed to post-process the results of the Operon-mapper. The script uses R libraries tidyverse [9], zoo[10], dplyr[11], and WriteXLSX [12] and this script is available on github.com/JulNejez/Operon-mapper-post-processing . Using the script sequences of genes of interest were extracted (genes belonging to COG1131, COG1961, COG0385, COG0582, COG4974), as well as other genes co-localized in the same operons. Then, we analyzed the length of possible operons and searched for the presence of the same operons in other examined genome sequences. The identified co-present genes were further characterized using the COG database [13], *e.g.* identified genes were categorized into COG categories according to their biological function within bacterial cell.

Next, we investigated whether other genes co-localized on operons harboring genes belonging to COG1131, COG1961, COG0385, COG0582, and COG4974 were transmitted horizontally. As mentioned above, the list of horizontally transferred genes between gut microbiota member has been detected recently [2]. Therefore, we predicted the function of these genes by eggNOG-Mapper [14]. Similarly to Operon-mapper, eggNOG-Mapper allows to assign COG function and COG categories to examined genes. Finally, the information about HGT from previous study [1] and the table with COGs from the current study were merged and summarized.

III. RESULTS AND DISCUSSION

A. Gene occurrence in operons

The summaries of the genetic features of the examined operons are shown in Table 1. All genes examined are present at least in four *Bacteroides spp.* It is evident that some genes are more likely to occur independently (e.g. genes belonging to COG1961) and some occur exclusively in operons (e.g. genes belonging to COG1131 – part of an ABC-type transporter consisting of at least 3 genes). Genes belonging to COG4974 are present in 47 operons. This suggests that this gene allows incorporation into the host genome. Interestingly, the operons with specific genes also varied considerably in length. For example, operons containing genes belonging to COG0385 consisted of 2-14 genes, but operons containing genes belonging to COG1961 were composed of only 2 genes. Moreover, genes belonging to COG1961 was found as a part of operon only in *B. sp*. An279. Therefore, it is possible that the Operon-mapper prediction may not have been correct and this gene may not be present in operons at all.

In bacteria *B. mediterraneensis* An801, *B. sp.* An279 and *B. ovatus* An161, not all of our investigated genes were found to be present. Genes belonging to COG1961 were absent in *B. mediterraneensis* An801, genes belonging to COG0385 were absent in *B. sp.* An279 and genes belonging to COG0582 were absent in *B. ovatus* An161. However, genes belonging to COG1131 and COG4974 were present in all bacteria. This confirms the assumption [1] that these genes are transmitted horizontally.

Table 1: Basic statistics of operons harboring genes belonging to COG1131, COG1961 and COG0385, COG0582, COG4974

	COG1131	COG1961	COG0385	COG0582	COG4974
GENE OCCURANCE IN *Bacteroides spp.*	5	4	4	4	5
AVERAGE NUMBER OF OPERONS PER BACTERIUM	2	0	1	2	9
NUMBER OF OPERONS PER GENOME	11	2	3	7	47
NUMBER OF SINGLE GENES PER GENOME	0	14	4	10	18
AVERAGE NUMBER OF GENES IN THE OPERON	5	2	6	3	3
MAXIMUM NUMBER OF GENES IN THE OPERON	8	2	14	5	5
MINIMUM NUMBER OF GENES IN THE OPERON	3	2	2	2	2

B. *The extended analysis of operon structure*

In addition, we analyzed operons in depth in order to compare operon structure within and between individual genome sequences. Genes adjacent to the known transferable genes were detected, characterized and gene synteny was compared. More frequent occurrences of the same genes in operons of different bacteria were observed for operons harboring genes belonging to COG1131 and COG4974 (Table 2).

Genes belonging to COG1131 are a part of ABC-type multidrug transport system that catalyzes drug efflux from inside of the cell into the periplasmic space in Gram-negative bacteria [14]. Genes belonging to COG1131 are only one of the 5 genes studied that belongs to the ABC-type transporter family. In the operons harboring genes belonging to COG1131 there are other genes from ABC-type transporter family (e.g. COG0842, COG1277). Genes belonging to COG1277 allow resistance to copper and it is often used as a toxic agent [13]

Genes belonging to COG4974 represent site-specific recombinase XerD. Adjacent to the COG4974 genes, numerous Remained Orthologous Groups (ROG), conserved hypothetical genes - e.g. ROG9437, ROG5764 were observed. Also adjacent to the COG4974 genes, genes belonging to COG4804 were observed. Genes belonging to COG4804 represents predicted nuclease of restriction endonuclease-like. Figure 2 shows the composition of two operons with COG1131 and COG4974. Also, the genes belonging to COG0582 and COG4974 were found to occur in the same operons.

Table 2: Genes frequently found in operons in conjunction with COG1131, COG1961, COG0385, COG0582, COG4974

COG1131	COG1277, COG0842, COG1566, ROG0497, ROG7278, ROG0722, ROG4288
COG1961	-
COG0385	COG1951
COG0582	COG4974
COG4974	COG4804, COG0582, ROG6557, ROG8437, ROG5764

The functional annotation based on COG database also allows us to assess gene function from the perspective of cell/biochemical/biological processes [13]. In general, genes localized in investigated operons were predominantly associated with L category (DNA replication and repair) (Figure 3). Also genes belonging to COG1961, COG0385, COG0582, and COG4974 play role in DNA repair mechanism, however

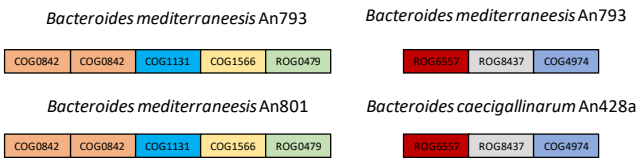


Figure 2: Operons harboring same gene structure.

genes belonging to COG1131 are involved in the defense system (category V). Although adjacent genes cannot be yet directly associated with any known function, they are likely co-transferred, and thus they might be involved in mechanism such as DNA repair or defense.

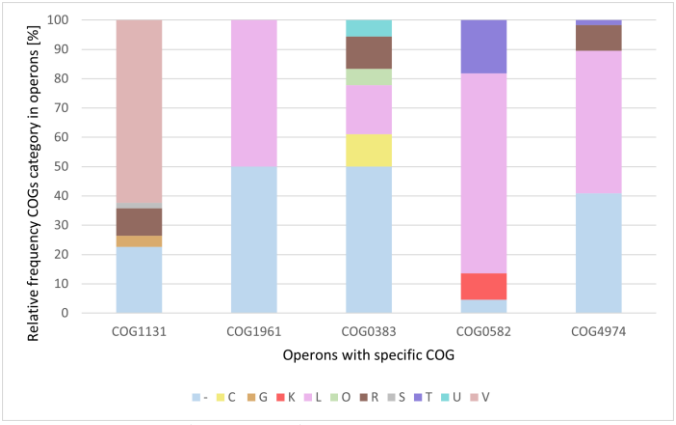


Figure 3: Relative frequency of COGs category in operons with genes COG1131, COG1961, COG0385, COG0582 and COG4974. "-" represents unknown genes, C category of energy production and conversion, G carbohydrate transport and metabolism, K transcription, L replication, O posttranscriptional modification, G gene function prediction only, S function unknown, T signal transduction mechanism, U ultracellular trafficking, secretion, and vesicular transport and V defense mechanism [13].

C. *HGT*

Table 3 shows that most genes are predicted based to be HGT, but for some genes, horizontal gene transfer is not expected. Our study confirmed that the entire operon harboring genes belonging to COG0582 was horizontally transmitted in all cases.

Although HGT is also predicted for other COG genes in the COG4974 operon, we cannot say whether the entire operon is horizontally transmitted because HGT information was not available for the hypothetical genes that were frequently found in these operons.

Table 3: Presence or absence of HGT

GENE	HGT	GENE	HGT
COG1131	Yes	**COG1277**	Yes
COG1470	No	**COG1668**	Yes
COG1277	Yes	**COG4152**	No
COG0842	Yes	**COG3451**	Yes
COG1566	Yes	**COG3772**	Yes
COG1961	Yes	**COG1951**	No
COG0358	Yes	**COG0582**	Yes
COG4974	Yes	**COG4804**	Yes
COG0582	Yes		

IV. Conclusion

In present study, we investigated 5 genes belonging to COG1131, COG1961, COG0385, COG0582, and COG4974 in 5 representatives of *Bacteroides spp. - Bacteroides ovatus* An161, *Bacteroides* sp. An279, *Bacteroides caecigallinarum* An428a. *Bacteroides mediterraneensis* An793 and *Bacteroides mediterraneensis* An801. These genes were predicted to be transferred horizontally between different bacteria, therefore we investigated whether other genes co-located in the same operons as well as whole operons may also be horizontally transferred.

Genes belonging to COG1131, COG1961, and COG4974 are frequently found in operons. On contrary, genes belonging to COG0582 and COG1961 genes are more likely to occur alone. In different bacteria, the length of the operon and gene synteny within the homologous operon were in concordance. The genes belonging to COG1131 often formed operons congaing genes involved in defense mechanism. Genes belonging to COG0582 and COG4974 were accompanied by genes associated with DNA repair.

In conclusion, most genes from HGT operons are horizontally transmitted; however, are not able to conclusively determine whether all compact operons can be horizontally transmitted, as HGT has not been demonstrated for all genes.

Acknowledgment

Computational resources were provided by the e-INFRA CZ project (ID:90140), supported by the Ministry of Education, Youth and Sports of the Czech Republic.

This work has been supported by the Czech Science Foundation (grant number 22-16786S)

References

[1] PÁL, Csaba; HURST, Laurence D. Evidence against the selfish operon theory. *Trends in genetics*, 2004, 20.6: 232-234.

[2] SCHWARZEROVÁ, Jana, et al. Systems biology approach for analysis of mobile genetic elements in the chicken gut microbiome. In: *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2022. p. 2865-2870.*)*

[3] SCHWARZEROVÁ, J.; ČEJKOVÁ, D. *Identification of horizontal genes transfer elements across strains inhabiting the same niche using pan-genome analysis.* Proceedings I of the 28th Conference STUDENT EEICT 2022 General Papers. Brno: Vysoké učení technické v Brně, Fakulta elektroniky a komunikačních technologií, 2022. s. 416-420. ISBN: 978-80-214-6029-4

[4] GALPERIN, Michael Y., et al. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic acids research*, 2015, 43.D1: D261-D269.

[5] PATRICK, Sheila. Bacteroides. In: *Molecular medical microbiology*. Academic Press, 2015. p. 917-944.

[6] JOBLING, Kelly Louise. Horizontal gene transfer in Bacteroides fragilis. 2014.

[7] 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.

[8] TABOADA, Blanca, et al. Operon-mapper: a web server for precise operon identification in bacterial and archaeal genomes. *Bioinformatics*, 2018, 34.23: 4118-4120.

[9] WICKHAM, Hadley, et al. Welcome to the Tidyverse. *Journal of open source software*, 2019, 4.43: 1686.

[10] GROTHENDIECK, Gabor; ZEILEIS, Achim. zoo: S3 infrastructure for regular and irregular time series. *Journal of Statistical Software*, 2005, 14.6: 27.

[11] YARBERRY, William; YARBERRY, William. Dplyr. *CRAN Recipes: DPLYR, Stringr, Lubridate, and RegEx in R*, 2021, 1-58.

[12] DRAGULESCU, Adrian A.; DRAGULESCU, Maintainer Adrian A.; PROVIDE, R. Package 'xlsx'. *Cell*, 2020, 9.1.

[13] TATUSOV, Roman L., et al. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic acids research*, 2000, 28.1: 33-36.

[14] 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.

[15] LUBELSKI, Jacek; KONINGS, Wil N.; DRIESSEN, Arnold JM. Distribution and physiology of ABC-type transporters contributing to multidrug resistance in bacteria. *Microbiology and molecular biology reviews*, 2007, 71.3: 463-476.

Joint audio denoising and inpainting with plug-and-play proximal algorithm

Michal Švento
Brno University of Technology, FEEC,
Department of Telecommunications,
Technická 12, 616 00 Brno, Czech Republic
212584@vut.cz

Ondřej Mokrý
Brno University of Technology, FEEC,
Department of Telecommunications,
Technická 12, 616 00 Brno, Czech Republic
xmokry12@vut.cz

***Abstract*—We propose plug-and-play variant of the Douglas–Rachford proximal algorithm for audio inpainting, which replaces a proximal step with a denoiser. In our situation, the observed samples are further degraded by noise. We demonstrate that the plug-and-play aproach has potential to succeed in this joint task of inpainting and denoising. Objective metrics show that the new method outperforms a conventional counterpart and that it is less sensitive to model hyperparameters.**

***Index Terms*—speech enhancement, deep learning, denoising, Douglas–Rachford algorithm, inpainting**

I. Introduction

Audio enhancement tasks mostly face problems like missing or damaged samples, noise, or clipping. Considering speech signals, we are not only interested in restoring the degradation sample by sample, but we also aim at improving the intelligibility of the recorded speech. Each restoration problem has developed its own way of enhancing the signal. Nowadays, the best way to differentiate algorithms is into two categories: conventional, e.g. using autoregressive (AR) modeling or sparsity-based optimization, and solutions using deep learning. The present paper focuses on the case of restoring a partially observed signal whose observed samples are further degraded by noise, i.e., the aim is to perform simultaneous inpainting and denoising of the speech signal.

In conventional approaches to inpainting, the AR-based Janssen [1] and Etter [2] algorithms dominate in terms of restoration quality. A more recent, successful class of methods is based on sparsity. The key idea is that after performing proper time-frequency analysis of an audio signal, most of the information is concentrated in a few coefficients, i.e., it is sparse. This can be applied as fitting the sparsest possible restoration either to the reliable observed samples [3]–[6], or to a signal not much diverging from the observation in the case of denoising [7].

Deep learning algorithms have also made their own progress in this area. The most efficient neural network models are autoencoders, recurrent neural networks (RNN) and Generative Adversial Networks (GAN). Current state-of-the-art deep learned algorithms are base on Speech Enhancement GAN (SEGAN) [8], NSNet [9], FullSubNet [10]. While learning-based algorithms allow to adapt to real-world signals, rather than to rely on hand-crafted priors like sparsity or the AR nature of signals, they need large datasets for training. Furthermore, neural networks are usually trained for a specific problem, lacking universal applicability on similar restoration tasks, in contrast to sparsity-based methods [11]–[13].

As a compromise between the conventional and learning-based methods, the plug-and-play method for image restoration was introduced in [14] and subsequently studied in [15], where part of each iteration of an optimization algorithm is replaced by a (learned) denoiser. In the present paper, we propose a hybrid algorithm based on the same paradigm, aiming at restoration of degraded speech. While [15] focused on adapting the Alternating Direction method of Multipliers (ADMM) and a recent declipping approach used the learned element only partially [16], we choose an opposite approach by working with a simple Douglas–Rachford algorithm (DRA) and exploring the trade-off between data fitting and denoising in the algorithm.

The paper is organized as follows. In section II we introduce the task from mathematical point of view and we define the restoration as a minimization task. Section III presents the plug-and-play method and its challenges. Section IV discusses the results and further improvements of algorithm. Finally, Section V concludes the paper.

II. Prerequsities

In this section, we formalize the task of inpainting and denoising and propose algorithmic solutions based on DRA. First, using frequency coefficients as input explained in II-B [17]. Second, using samples in time domain is described in subsection II-C [18].

A. Task formulation

We consider column vector $\mathbf{s} \in \mathbb{R}^N$ as our observed damaged single-channel signal of length N. We have set of sample indices $\{1, 2, \ldots, N\}$, which has two disjunctive subsets: I_{M} for missing positions and I_{R} stands for reliable positions. Usually, samples $\mathbf{s}(I_{\mathrm{R}})$ are considered reliable (undamaged) and $\mathbf{s}(I_{\mathrm{M}})$ are samples, which we are looking for. It is common to rewrite it in matrix form:

$$\mathbf{s}_{\mathrm{R}} = \mathbf{M}_{\mathrm{R}}\mathbf{s},$$

where $\mathbf{M}_\mathrm{R} \in \mathbb{R}^{|I_\mathrm{R}| \times N}$ is the mask matrix, selecting rows from identity matrix corresponding to the indices in I_R [3], with $|I_\mathrm{R}|$ denoting the number of indices in the subset I_R. In words, $\mathbf{M}_\mathrm{R}\mathbf{s}$ represents choosing the samples from $\mathbf{s}$ on positions I_R.

It is convenient to define the set of signals fitting the observation as

$$\Gamma = \{\mathbf{x} \in \mathbb{R}^L \mid \mathbf{M}_\mathrm{R}\mathbf{x} = \mathbf{M}_\mathrm{R}\mathbf{s}\}. \tag{1}$$

In the noise-less case, we would search for a suitable signal in Γ. On the other hand, when the observed samples are distorted by noise, we only require the solution $\mathbf{x}$ to be close to Γ.

B. Synthesis-based formulation

We define the task of finding a suitable signal in (or close to) Γ as a sparsity-based problem, minimizing ℓ_0-norm of the Gabor coefficients, i.e., the number of non-zero coefficients of the time-frequency representation of the signal. However, this task leads to an NP-hard problem and is hardly solvable [17]. The closest redefinition is to use the ℓ_1-norm as follows:

$$\arg\min_{\mathbf{c}} \|\mathbf{c}\|_1 \quad \text{subject to} \quad \mathbf{D}\mathbf{c} \in \Gamma, \tag{2}$$

where $\mathbf{D}$ is synthesis operator (inverse discrete Gabor transform), which is the adjoint of the analysis operator $\mathbf{A}$, i.e., $\mathbf{D} = \mathbf{A}^*$. Therefore the restored signal corresponds to $\mathbf{x} = \mathbf{D}\mathbf{c}$.

In the presence of noise, the condition $\mathbf{D}\mathbf{c} \in \Gamma$ is not beneficial, since it forces the solution to still contain the noise. Its reduction can be included in the formulation by rather minimizing the distance of the solution from the set Γ. This can be computed as the difference of the solution $\mathbf{D}\mathbf{c}$ and the observation $\mathbf{s}$ on the positions I_R. Typically, squared ℓ_2-norm is used in this context:

$$\arg\min_{\mathbf{c}} \|\mathbf{c}\|_1 + \frac{\alpha}{2}\|\mathbf{M}_\mathrm{R}\mathbf{D}\mathbf{c} - \mathbf{M}_\mathrm{R}\mathbf{s}\|_2^2. \tag{3}$$

The parameter $\alpha > 0$ manages the trade-off between sparsity of the coefficients and the data-fidelity. Both (2) and (3) could be solved using DRA [13], [17]. However, the algorithms use the Gabor coefficients as the main variable, which makes the incorporation of a denoiser (working with a signal as the input) problematic.

C. Analysis-based formulation

Second approach uses time domain samples as input and also as variables of the optimization task. In case the transform is redundant, i.e., we have more coefficients than signal samples, this second approach leads to different solutions in general [17].

The main minimazation task (4) is reformulated as,

$$\arg\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{x} \in \Gamma, \tag{4}$$

or, in the denoising case,

$$\arg\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x}\|_1 + \frac{\alpha}{2}\|\mathbf{M}_\mathrm{R}\mathbf{x} - \mathbf{M}_\mathrm{R}\mathbf{s}\|_2^2. \tag{5}$$

If we want to employ DRA also on the analysis-based task, a problem occurs that we do not know proximal operator of ℓ_1-norm after analysis, which is essential for the DRA. However, we can resort to the so-called approximal operator without significant loss of restoration quality [18]. The algorithm is summarized in Alg. 1.

Algorithm 1 DRA for (4) or (5).

Input: $\lambda_n > 0$, $\gamma > 0$, $\widetilde{\mathbf{x}}_0 \in \mathbb{R}^N$, $\beta \in [0, 1]$
1: **for** $n = 0, 1, \ldots$ **do**
2: $\quad \mathbf{x}_n = (1-\beta)\widetilde{\mathbf{x}}_n + \beta\,\mathrm{proj}_\Gamma(\widetilde{\mathbf{x}}_n)$
3: $\quad \widetilde{\mathbf{x}}_{n+1} = \mathbf{x}_n + \lambda_n\left(\mathbf{D}\left(\mathrm{soft}_\gamma\left(\mathbf{A}\left(2\mathbf{x}_n - \widetilde{\mathbf{x}}_n\right)\right)\right) - \mathbf{x}_n\right)$
4: **end for**
5: **return** $\mathbf{x}_n$

The operator proj_Γ is projection onto convex set Γ and soft_γ is soft thresholding operator [19]. Projection, in this case, means replacing restored samples in positions considered reliable by the observed samples. The parameter β allows to differentiate between simple inpainting (4) and the joint problem (5). The case of $\beta = 1$ corresponds to performing projection in the update, in line with [17]. For $\beta < 1$, it can be derived (see e.g. [19, Sec. 4 and Tab. 1]) that the update corresponds to the proximal operator of $f(\mathbf{x}) = \frac{\alpha}{2}\|\mathbf{M}_\mathrm{R}\mathbf{x} - \mathbf{M}_\mathrm{R}\mathbf{s}\|_2^2$ with $\alpha = \frac{\beta}{1-\beta}$.

III. Plug-and-play inpainting

The plug-and-play method was proposed in [14] and [15] for image restoration. The main idea is to replace a proximal operator inside an iterative algorithm with an denoiser. Since this approach does not affect the data-fitting part of the minimization, it should be suitable for different resotaration tasks. We rewrite this task to solve our audio inpainting problem using DRA with approximal operator (analysis-based), i.e., we modify Alg. 1, leading to Alg. 2:

Algorithm 2 Plug-and-play DRA

Input: $\lambda_n > 0$, $\gamma > 0$, $\widetilde{\mathbf{x}}_0 \in \mathbb{R}^N$, $\beta \in [0, 1]$
1: **for** $n = 0, 1, \ldots$ **do**
2: $\quad \mathbf{x}_n = (1-\beta)\widetilde{\mathbf{x}}_n + \beta\,\mathrm{proj}_\Gamma(\widetilde{\mathbf{x}}_n)$
3: $\quad \widetilde{\mathbf{x}}_{n+1} = \widetilde{\mathbf{x}}_n + \lambda_n\left(\mathcal{D}\left(2\mathbf{x}_n - \widetilde{\mathbf{x}}_n\right) - \mathbf{x}_n\right)$
4: **end for**
5: **return** $\mathbf{x}_n$

Convergence of the plug-and-play approach can be proven in case $\mathcal{D}$ is a non-expansive operator [15]. However, this is hard to prove in practice with off-the-shelf denoisers. Thus, the only requirement for our denoiser $\mathcal{D}$ was primarily good results in subjective metrics (discussed more in IV-A). The implementation of the denoiser of our choice is based on the *mayavoz* toolkit [20]. This tool provides us simple use of learned model with downloadable checkpoints. Our selection is the WaveUnet learned on the Valentini dateset [21] (the variant with 28 speakers).

IV. Testing data and evaluation

As introduced before, we use discrete Gabor transform as the analysis operator $\mathbf{A}$ with following setup parameters:

Gauss window with length $w = 1024$, hop length $a = 256$ and number of frequency channels used in Fast Fourier transform $n_{\text{FFT}} = 1024$. To satisfy the Parseval tight frame condition, the synthesis operator $\mathbf{D} = \mathbf{A}^*$ is configured the same way [17].

A. Metrics

A standard metric for restoration quality is the signal-to-noise ratio (SNR), where the *signal* represents the clean ground truth and *noise* is the difference of the restoration and the ground truth. We calculate SNR either for whole signal or only on I_{M}, i.e, on the indices of missing samples. The higher the SNR, the better the restoration.

A drawback of SNR is that it measures sample-wise difference, which does not take into account human perception of sound quality. Valuable metrics for evaluating speech signals are those simulate measure human perception. We use two objective metrics: Perceptual Evaluation of Speech Quality (PESQ) [22] and Short-Time Objective Intelligibility (STOI) [23]. PESQ has scale from -0.5 to 4.5, with higher score meaning better result. STOI measures correlation of two signals on a scale from 0 to 1.

B. Dataset and degradation setup

We have chosen to work with the Valentini dataset [21] used commonly in speech enhancement challenges (the dataset is split to two subsets: clean and noisy). For each signal, we used the clean sample as a reference and its noisy variant as the input to the algorithm. The average SNR of our files was approximately 11 dB. As a simulation of inpainting task, we used random drop-out of 40% of samples as used in [18]. This means that the reliable set I_{R} consisted of 60% of all the sample indices.

C. Choice of the soft threshold

To reduce the number of variables in the comparison between the conventional and plug-and-play approach, we decided to fix the γ parameter in DRA. We examined DRA with values of $\gamma \in \{0.001, 0.01, 0.1\}$ for a single restoration task. At the same time, we examined the effect of the trade-off parameter β in Alg. 1, since it has a significant effect on the results in the case of signal denoising. Tested signals are pair of clean and noisy signal from the dataset described above with random drop-out of 40% samples. The parameters of DRA were: $\lambda_{n+1} = 0.9 \cdot \lambda_n$, starting with $\lambda_0 = 1$, and 50 iterations. The results are shown in Fig. 1:

The best results were obtained with $\gamma = 0.01$, which reached the best restoration quality, especially with increasing ratio of projected signal.

D. Test setup

We have chosen randomly 10 audio files for final evaluation from the Valentini dataset [21] described before. As justified in IV-C, we have chosen $\gamma = 0.01$. The simulation was split to three restoration variants: two conventional ones and one including the learned denoiser. The conventional ones were

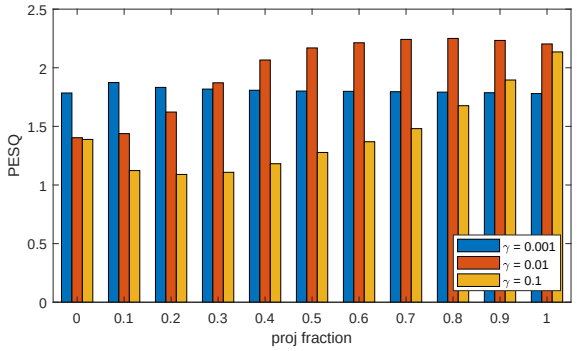


Fig. 1: PESQ metric results made to choose the best value of γ in DRA for testing with multiple audio files (these values are shown in the legend). The results in terms of STOI and SNR were principally the same. The label *proj fraction* refers to parameter β from Alg. 1.

two versions of DRA 1 with same parameters except for the number of iterations (50 or 500) and the choice of λ_n (decreasing or constant). The DRA with denoiser proposed in Alg. 2 was tested with 50 iterations, as in the shorter conventional method. In all the algorithms, the initial λ_0 was set to 1. For the case of DRA 1 with 500 iterations, we kept constant $\lambda_n = 0.1$, which provided better results compared to the choice of $\lambda_n = 1$. In the case of the plug-and-play version 2, and the reference DRA 1 with 50 iterations, the parameter was set to decrease in every iteration as $\lambda_{n+1} = 0.9 \cdot \lambda_n$. The motivation for decreasing λ is clear from Fig. 2. The test was performed using DRA and every option should eventually converge to the same optimal point in time [19]. However, the choice has an impact in practice, either in particular choice of the argument of the minima, or due to numerical reasons and finite number of iterations. If λ decreases, we minimize the impact of soft threshold operator in latter iterations which appears to be effective in our case.

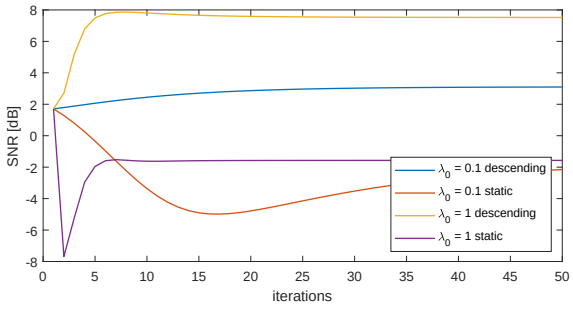


Fig. 2: SNR from DRA tested for 50 iterations with static and decreasing choice of $\lambda \in \{0.1, 1\}$. Tested signals are same as in experiment with γ.

E. Comparison results

Graphs with STOI, PESQ and SNR are presented in Fig. 3. We see that plug-and-play is less sensitive to the fraction

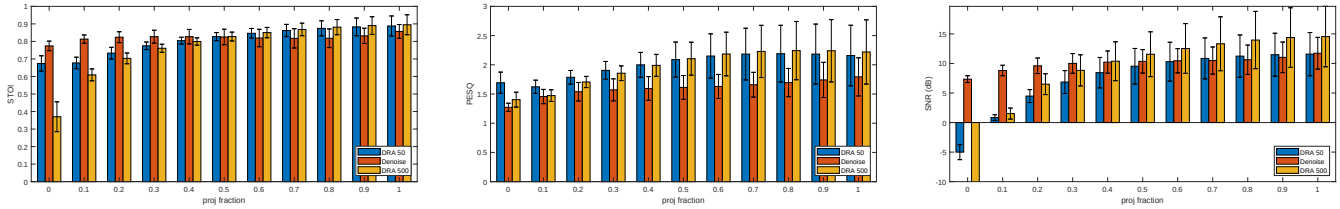


Fig. 3: Average results from 10 tested signals. Metrics are from left to right: STOI, PESQ, SNR.

of projection compared to DRA, and in comparison with DRA (50 iterations, decreasing λ_n), it reaches better result. DRA with 500 iterations has higher variance of the results, but globally, it outperformed the plug-and-play method with optimal choice of the parameters such as the ratio β.

V. Conclusion

Our reformulation of plug-and-play method for audio inpainting with noisy data was demonstrated to be successful with properly chosen ratio of projection and λ. Next steps could be to extend the test for other similar tasks (declipping, dequantization), following the pattern of root method [15], solving multiple tasks with one setup of the algorithm. We also suggest suggested not to use off-the-shelf denoiser, but learn own model on data related to our problem (e.g. specific noise due to dropout of a certain percentage of samples). In addition, we might be able to make the self-learned denoiser non-expansive, thus satisfying the theoretical convergence guarantees [14], [15]. A compromise might be to use transfer learning, i.e., to take a successful denoiser and continue learning from a checkpoint with our dataset, with idea to reduce artifacts generated in the inpainting task.

References

[1] A. Janssen, R. Veldhuis, and L. Vries, "Adaptive interpolation of discrete-time signals that can be modeled as autoregressive processes," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, pp. 317–330, Apr. 1986.

[2] W. Etter, "Restoration of a discrete-time signal segment by interpolation based on the left-sided and right-sided autoregressive parameters," *IEEE Transactions on Signal Processing*, vol. 44, pp. 1124–1135, May 1996.

[3] A. Adler, V. Emiya, M. Jafari, M. Elad, R. Gribonval, and M. Plumbley, "Audio Inpainting," *IEEE Transactions on Audio Speech and Language Processing*, vol. 20, pp. 922–932, Mar. 2012.

[4] S. Kitić, N. Bertin, and R. Gribonval, "Sparsity and cosparsity for audio declipping: a flexible non-convex approach," in *LVA/ICA 2015 – The 12th International Conference on Latent Variable Analysis and Signal Separation*, (Liberec, Czech Republic), pp. 243–250, Aug. 2015.

[5] P. Záviška, P. Rajmic, O. Mokrý, and Z. Průša, "A Proper Version of Synthesis-based Sparse Audio Declipper," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 591–595, May 2019. ISSN: 2379-190X.

[6] O. Mokrý, P. Záviška, P. Rajmic, and V. Veselý, "Introducing SPAIN (SParse Audio INpainter)," in *2019 27th European Signal Processing Conference (EUSIPCO)*, pp. 1–5, Sept. 2019. ISSN: 2076-1465.

[7] M. Kowalski, K. Siedenburg, and M. Dörfler, "Social sparsity! neighborhood systems enrich structured shrinkage operators," *IEEE Transactions on Signal Processing*, vol. 61, no. 10, pp. 2498–2511, 2013.

[8] S. Pascual, A. Bonafonte, and J. Serrà, "SEGAN: Speech Enhancement Generative Adversarial Network," tech. rep., June 2017. arXiv:1703.09452 [cs] type: article.

[9] Y. Xia, S. Braun, C. K. A. Reddy, H. Dubey, R. Cutler, and I. Tashev, "Weighted Speech Distortion Losses for Neural-network-based Real-time Speech Enhancement," tech. rep., Feb. 2020. arXiv:2001.10601 [cs, eess] type: article.

[10] X. Hao, X. Su, R. Horaud, and X. Li, "FullSubNet: A Full-Band and Sub-Band Fusion Model for Real-Time Single-Channel Speech Enhancement," tech. rep., Jan. 2021. arXiv:2010.15508 [cs, eess] type: article.

[11] C. Gaultier, N. Bertin, S. Kitić, and R. Gribonval, "A modeling and algorithmic framework for (non)social (co)sparse audio restoration," 2017.

[12] O. Mokrý, P. Rajmic, and P. Záviška, "Flexible framework for audio reconstruction," in *Proceedings of the 23rd International Conference on Digital Audio Effects (DAFx2020)*, vol. 1, (Vienna, Austria), Sept. 2020-21.

[13] P. Záviška, P. Rajmic, and O. Mokrý, "Audio dequantization using (co)sparse (non)convex methods," in *2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Toronto, Canada), pp. 701–705, 2021.

[14] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *2013 IEEE Global Conference on Signal and Information Processing*, IEEE, Dec. 2013.

[15] S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-Play ADMM for Image Restoration: Fixed Point Convergence and Applications," tech. rep., Nov. 2016. arXiv:1605.01710 [cs] type: article.

[16] T. Tanaka, K. Yatabe, M. Yasuda, and Y. Oikawa, "APPLADE: Adjustable plug-and-play audio declipper combining DNN with sparse optimization," in *2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, may 2022.

[17] O. Mokrý and P. Rajmic, "Audio Inpainting: Revisited and Reweighted," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 2906–2918, 2020.

[18] O. Mokrý and P. Rajmic, "Approximal operator with application to audio inpainting," *Signal Processing*, vol. 179, p. 107807, Feb. 2021.

[19] P. Combettes and J. Pesquet, "Proximal splitting methods in signal processing," *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, vol. 49, pp. 185–212, 2011.

[20] A. Kiran, "shahules786/mayavoz," Mar. 2023. original-date: 2022-08-20T04:44:35Z.

[21] C. Valentini-Botinhao, "Noisy speech database for training speech enhancement algorithms and TTS models," *http://parole.loria.fr/DEMAND/*, Aug. 2017.

[22] A. Rix, J. Beerends, M. Hollier, and A. Hekstra, "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221)*, vol. 2, pp. 749–752 vol.2, May 2001. ISSN: 1520-6149.

[23] C. H. Taal, R. C. Hendriks, R. Heusdens, and J. Jensen, "A short-time objective intelligibility measure for time-frequency weighted noisy speech," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 4214–4217, Mar. 2010. ISSN: 2379-190X.

Visualization of system real-time data using software modules

Martin Kubásek
Department of Control and Instrumentation
Faculty of Electrical Engineering and Communication, BUT
Brno, Czech Republic
xkubas04@vutbr.cz

Václav Kaczmarczyk
Department of Control and Instrumentation
Faculty of Electrical Engineering and Communication, BUT
Brno, Czech Republic
kaczmarczyk@vut.cz

***Abstract*—This paper describes the creation of a web visualization that will be able to display the states of any system in real time and that could be deployed at various data sources and locations. At the beginning, the existing visualization is described, in which it was possible to look for inspiration when creating a new system. The shortcomings and limitations of the existing solution were identified, on the basis of which it was decided that main areas to focus on are higher modularity and manageability of the solution code. These goals were achieved with the use of software modules, focus on clean code and leveraging the capabilities of .NET. The new system should be easy to deploy on any data source and should be easy to expand thanks to the clear separation of concerns. At the same time the system is able to use the same code for accessing data and similar data model for configuration. This should ensure an easy upgrade for installations that use the old visualization.**

***Index Terms*—Visualization, P&ID, Web, .NET, C#, DotVVM, ASP.NET**

I. Introduction

This paper deals with the creation of visualization for displaying system states in real time and for displaying data from its history. It will be implemented as a web application on the DotVVM platform, written in C#. The new visualization will be designed with an emphasis on modularity, extensibility and simple manageability in future expansion. The new visualization will be based on an already existing visualization system, which is custom created for a specific process and therefore it is not possible to simply deploy it to visualize new processes without modifications of the code. The new system should be modular enough to overcome this problem. Deployment on each new installation should require minimum programming work. At the same time, it should be as compatible as possible with the data model used by the existing system and should implement all its existing functions, or be prepared for their simple implementation.

II. Existing visualisation

The specific system that served as an inspiration was created as a one-off project for a manufacturing facility, where it provides overview of a compressor room and cooling system technologies. At the same time, it allows to adjust some of the technologies. For each of the parts of the technology, the so-called **scenes** are defined. Visualization scene is a P&ID diagram that shows the current values of individual devices. Example scene is shown in the image 1.

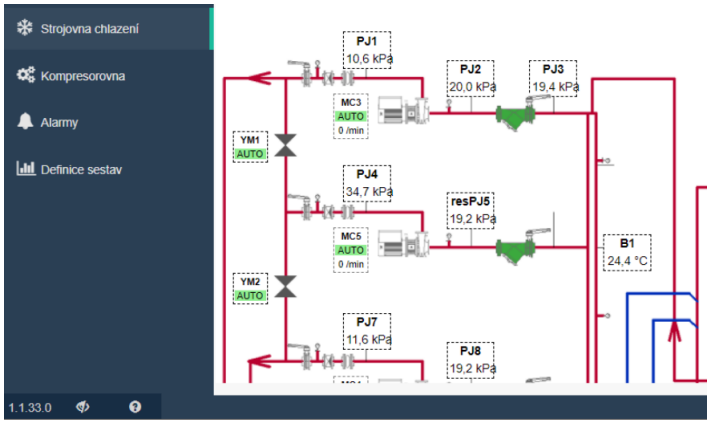


Fig. 1. Visualization scene in the old system

On every scene there are several types of elements that display specific values of the monitored system, for example:

- Valve - shows status (open/closed/intermediate position/fault) and control mode
- Pump - shows the speed and control mode
- Sensor - shows the current analog value

It is necessary to log in to access the visualization. Each part of the web application is tied to a permission that the user needs to own.

A. Maintainibility problems

The existing system is built in such a way that it is always tailored to the specific location where it is to be deployed. If it was necessary to deploy the system to another site, another copy was always created, containing specific modifications for the installation. As a result, there are multiple versions of the same system that are not completely compatible with each other and are difficult to maintain.

Another problem is the design of the application itself. Each element on the scene is statically defined ”side by side” in the page code as multiple lists of elements that are then rendered on top of each other on the page. They also contain a lot of

cross-references between themselves and the scene page itself, which is problematic from the point of view of manageability of the code [1]. Elements on the scene are always tailored to the location and expect a specific data model.

Another advantage of creating a new version of the visualization is updating to new technologies. A newer version of DotVVM will be used, which already uses the much faster and cross-platform ASP.NET Core and .NET6. [2] [3] [4].

III. Used technologies

A. Web application

The old and new systems are based on the same technologies and will be connected to the same type of database. The web application is implemented using the DotVVM platform, which has been offered free of charge as an open-source project since 2016 by the Czech company RIGANTI s.r.o. [5]. The platform builds on the proven ASP.NET Core technology and allows most of the page to be written in C# from the server to the client part. The front-end of the application itself is then implemented using HTML and CSS. Alternatively, in parallel with DotVVM, you can also use code running directly in the browser like JavaScript or React. JavaScript components can directly integrate with DotVVM and extend its functionality. There is also native support for creating of React base DotVVM components [2].

DotVVM is built on the MVVM (Model-View-ViewModel) design pattern, which can also be seen in other .NET frameworks such as WPF, Xamarin, and MAUI. The structure of the application is divided into three parts, where each part has a precisely defined responsibility [6].

Pages generated through DotVVM are common HTML pages that are linked by generated JavaScript. This makes up the View part and its connection to ViewModel. As needed, it sends JSON queries to the REST API, which in turn returns JSON objects. These JSON queries usually contain serialized ViewModel objects with altered states either by View (from client) or by Model (from server). The API is then generated automatically in ASP.NET. The client-side uses the KnockoutJS library, which solves all interactivity between elements linked through binding and communication with ViewModel [7] [8] [2].

B. Data storage

A MySQL database is used for data storage, where the configuration data of the visualization and the data of the monitored system are located.

C. Dependency injection

In the new system, the Dependency injection (DI) method was used to propagate dependencies between objects. This method allows you to manage all dependencies in the system in one place and have full control over their life cycle.

Two DI layers were used in the design. The top layer uses a built-in method in the ASP.NET Core to ensure dependency propagation for classes that are created by ASP.NET (DotVVM). The second layer uses the Autofac library and provides services that exist within the scope of user's login session.

IV. New system

The architecture and basic concepts in the new system will be now described. Visually, it is built so that there is a navigation menu on the left side and selected content on the rest of the screen. The site is built responsively, so it should also be accessible on mobile devices.

A. Autorization

Authorization is provided with using standard cookie authentication. Authorization is checked on each request. If the user is not logged in, he is redirected to the login page. After logging in, he is granted access to scene management or to one of the defined scenes.

B. Scene management

Scene management page is intended for creating, deleting and overall management of scenes. It is possible to set parameters such as name, icon and background. Available scenes are further displayed as menu items in the visualization navigation menu.

C. Scene

Scene is a page that provides a view of the current data of a part of the monitored system (technology). The page can be switched to edit mode, where you can add and edit tiles, move them to any location, or re-size them. Any changes to the scene can be saved to the database or reverted to the state before editing. This is one of the places where it was necessary to create a JS working closely with DotVVM. Changes to tile size and positions are handled on client side in JS and only sent to server after the interaction has finished.

D. Tiles

Tiles are general DotVVM components placed on the scene, the content of which consists of individual software modules. Their main function is to create a common interface between the scene and the modules. The tile has a common user interface regardless of its content, and provides functions such as loading and saving settings for the module, displaying the software module settings in a dialog and the interface for periodic refresh. Because of this abstraction, there is no direct software link (reference) between the scene and the module or between other parts of the application.

E. Software modules

These modules contain the graphic elements that the user will see on the scene and contain the logic for processing the retrieved value from the database. They are completely controlled from tiles using an unified interface and contain no code to access the database. They can have their own specific settings defined, which is stored over the tile interface as JSON. The type of software module that will appear in the tile is identified by a string that corresponds to the module namespace.

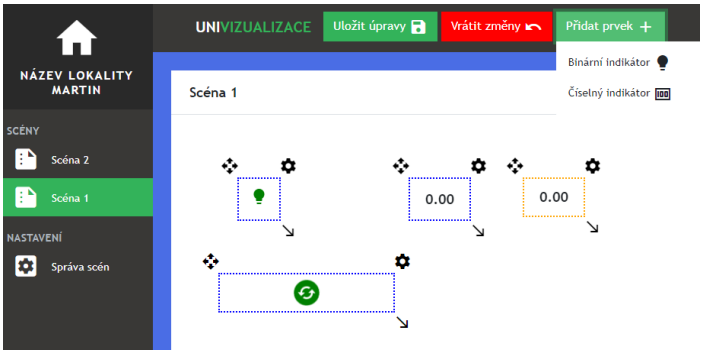


Fig. 2. Tiles placed on scene in the new system

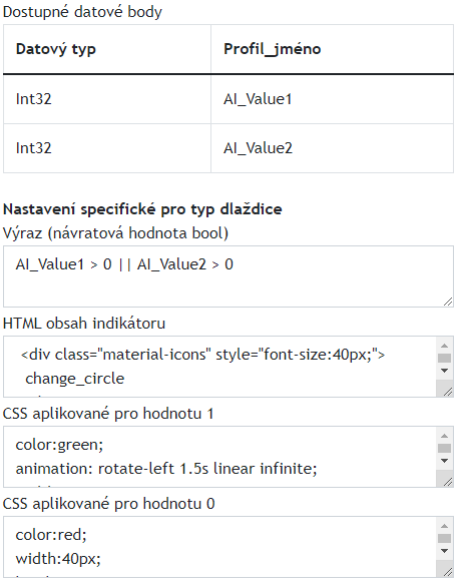


Fig. 3. Example of tile settings

They have no references to other parts of the system or other modules. With a unified interface, the development of the new module is simplified to the definition of two View and ViewModel objects for the tile and settings user interface. Alternatively, it is now possible to simply develop a new module while the visualization is deployed and perform a one-time update using a namespace change and application restart.

An example of several placed tiles that contain the Binary Indicator and Numeric Indicator software modules can be seen in the figure 2.

F. Loading of data

As indicated earlier, all access to data takes place through services that are provided via Dependency Injection. Thanks to layered separation of concerns, the entire application works with independent POCO classes that are not dependent on the data model. Access to the database is limited to two services, one providing loading and saving the visualization configuration and the other providing access to the monitored system data.

In each tile settings, you can choose a data source (monitored device). Each data source has defined profiles in the database that determine which data points are available for that device. These data points are then presented to the user and the software module, who can use them as a source of value that will impact the state of shown tile in some way. The values are provided as processed values of C# data types. It is then up to the developer (and user) of the module what will happen with the value. It can be directly bound to some property (state) of the module or a built in support for expressions can be used.

In the software module settings, it is possible to write so-called expressions, which are parts of C# code that use provided data points as variables. As an example of settings with expressions the Binary indicator tile can be seen in the image 3. Here a binary state of the indicator is evaluated as logical OR that one of the analog values is greater than zero. Thanks to the ability to insert your own HTML, to which any CSS can be applied using expressions, this type of module is very flexible.

Updating elements on the scene is done in following sequence:

- JS timer running on scene page initiates the scene update
- All data point expressions and references from all modules that are present in the scene and have a refresh interval are collected
- References to data points are grouped so that the database is read only once for each object
- Data is read for each data point used
- All modules are gradually updated - the module lets evaluate its expressions in the tile and works only with the resulting value

This solution uses extensive reflection and interpretation of expressions using the DynamicExpresso library. This will definitely make this approach slower and more computationally intensive than reading directly from variables in the database and could cause problems in complex scenes. Performance impact is reduced by caching data requests and expressions in one central location for all scenes and logged-in users. At the same time, the database is loaded only once. The read data within the scene update is therefore always the last available data that was read from the database.

G. Login scoped services

ASP.NET supports these three basic types of services, provided via DI. Services differ in their life cycle within HTTP request and application runtime [9]:

- Transient - each request for this service creates a new object
- Scoped - the request for this service creates a new object only once within the lifecycle of the HTTP request
- Singleton - there is only one service within the runtime of the entire application

For some features, such as scene and tile editing, it was necessary to ensure that some services existed for the duration

of the user's login. To do this, an additional DI layer was defined using the Autofac library. It is a Singleton service that maintains a set of additional services for each user, which are reserved only for that user. Services are disposed only when the user logs out. The diagram describing this feature so-called LoginScopedService is shown in 4.

All mentioned services exist only on the web server so we need to mind the memory utilization. Each user login reserves a part of memory with LoginScopedService that would be reserved indefinitely. In addition to disposing of the instance on user log out a timeout was implemented. This timeout tracks user's requests and disposes the services after a long time of inactivity. The time interval for this timeout is much longer than timeout of login cookie so it should not triggered with everyday usage.

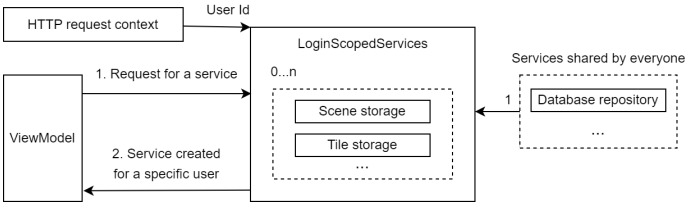


Fig. 4. Diagram of LoginScopedServices lifecycle

V. Conclusion

The new system is a functional prototype of visualization, which is written in a completely modular way and allows to implement visualization elements as separate software modules. These modules have minimal reference in the code to their surroundings, which simplifies their development and maintenance. The system also allows you to read data using expressions bound to data points, which can be freely used according to the selected data profile. This simplifies module development, configuration and possible changes in the data model. The current version of visualization does not implement all the visual elements that were available in the old version, but provides a complete functional basis for their future implementation.

Acknowledgment

The completion of this paper was made possible by the grant No. FEKT-S-23-8451 - "Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence, automation and measurement" financially supported by the Internal science fund of Brno University of Technology.

References

[1] R. C. Martin, *Clean Code: A Handbook of Agile Software Craftsmanship*. Prentice Hall.
[2] Riganti, "DotVVM | Documentation." [Online]. Available: https://www.dotvvm.com/docs
[3] Microsoft, "Choose an ASP.NET Core UI." [Online]. Available: https://learn.microsoft.com/en-us/aspnet/core/tutorials/choose-web-ui
[4] S. Toub, "Performance Improvements in .NET 6." [Online]. Available: https://devblogs.microsoft.com/dotnet/performance-improvements-in-net-6/
[5] "Dotvvm | github." [Online]. Available: https://github.com/riganti/dotvvm
[6] D. Britch, N. Schonning, C. Dunn, and J. Osborne, "The Model-View-ViewModel Pattern - Xamarin." [Online]. Available: https://learn.microsoft.com/en-us/xamarin/xamarin-forms/enterprise-application-patterns/mvvm
[7] S. Sanderson, "Knockoutjs documentation." [Online]. Available: https://knockoutjs.com/documentation/observables.html
[8] Riganti, "DotVVM | FAQ." [Online]. Available: https://www.dotvvm.com/faq
[9] Microsoft, "Dependency injection in ASP.NET Core." [Online]. Available: https://learn.microsoft.com/en-us/aspnet/core/fundamentals/dependency-injection

Influence of dataset on anomaly detector performance estimation

Jakub Lukaszczyk
Master's Degree Programme (2), FEEC
Brno University of Technology
Brno, Czech Republic
e-mail: 211157@vut.cz

***Abstract*—This paper deals with anomaly detection in image data of an industrial product. An existing Deep Support Vector Data Description method (Deep SVDD) is used to conduct experiments to determine whether the anomalous class can be modelled with the same success rate as the normal class for the purposes of anomaly detection. The paper also studies the relationship between the model performance estimation and the diversity of anomalies used in the testing phase.**

***Index Terms*—Deep SVDD, Anomaly detection, Image anomaly, Performance metrics**

I. Introduction to anomaly detection

As the variability and complexity of products increases, so do the demands for quality control. The issue of detecting anomalies arising in production is a very topical problem and the approach to its solution is not always straightforward, as the nature of possible anomalies cannot be effectively predicted in advance or otherwise described analytically. The aim of this paper is to determine whether the diversity of the anomalous test data used is relevant for the model performance estimation. As an anomaly detector a one-class classifier based on the Deep SVDD method [1] is used. Two sets of experiments will be conducted. The first will explore whether anomalous samples are as suitable for training the model as the normal samples are. The second set of experiments will aim to investigate the effects of the diversity of anomalies in the testing dataset on the evaluation and estimation of the performance of the model.

In machine learning there are three basic types of classification tasks. These are multi-class classification, binary classification and one-class classification [2] [3]. Depending on the nature of the problem to be solved, it is necessary to choose the right type of classifier before choosing the classification method itself as not all classifiers are suitable for every application.

For example a popular approach to anomaly detection is to use a binary classifier with normal samples and anomalies chosen as the two distinct classes [4]. The issue with this approach is that it does not meet the prerequisites for using a binary classifier. Some of the requirements are for the two classes to be roughly equally represented, homogeneous and clearly definable [4]. For the normal class these requirements are met. As for the anomalies, not all types of anomalies are known a priori, nor do they form a homogeneous clusters, since there is virtually unlimited number of possibilities how a sample can deviate from the normal.

For anomaly detection a different approach must be taken, such as one-class classifiers. These methods are based on the assumption that only one class is well defined and the goal is to detect whether or not the presented sample satisfies the characteristics of this class. With this approach the fact that the anomalies are non-homogeneous and not well-defined is no longer an issue because only deviation of the samples from the normal class is observed.

It is also necessary to consider the different severity of classification errors when detecting anomalies opposed to other types of classification tasks. In the case of binary or multi-class classification the classes often have the same or very similar relevance. However, in the area of anomaly detection emphasis is placed on correctly detecting anomalies, often at the expense of a slight increase in the number of mislabeled normal samples [5].

Thus, when setting the operating point the goal is not to find a region where the number of false positive detections and false negative detections is approximately equal. In the case of anomaly detection task models should be rather compared by an AUC metric and by comparing the increase in mislabeled normal samples when choosing a certain true positive rate on the ROC curve (90 % was chosen for demonstration purposes within this paper).

II. Deep SVDD - method for anomaly detection

Nowadays, the most commonly used methods for anomaly detection in images often use principles that are adopted from other similar tasks and are exploiting some of their properties [6]. Deep SVDD aims to address this problem by defining an optimization criterion that is directly tailored to the anomaly detection task and thus has a potential to achieve better results than the methods using adopted principles [1].

Deep SVDD specifies a neural network for feature extraction and an optimization rule that attempts to minimize the hypersphere that surrounds the data representation. Minimizing the volume forces the network to extract only the common features from the input data which causes them to be mapped near the center of the hypersphere [1].

Training then takes place simultaneously for both the extraction of the features and their mapping to the volume-

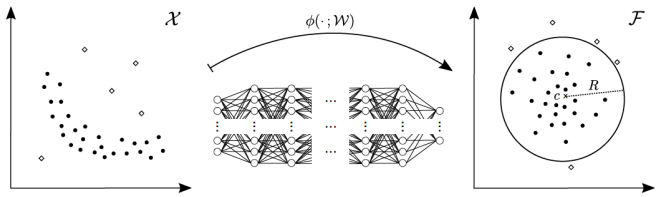


Fig. 1. Deep SVDD - Basic idea of mapping input features into multidimensional sphere of minimum volume using network Φ [1]

minimized hypersphere. This criterion can be described by the following formula [1]:

$$\min_{R,\mathcal{W}} R^2 + \frac{1}{\nu n}\sum_{i=1}^{n} max\{0, ||\phi(x_i;\mathcal{W}) - c||^2 - R^2\} + \\ + \frac{\lambda}{2}\sum_{\ell=1}^{L}||\boldsymbol{W}^{\ell}||_F^2 \tag{1}$$

The aim is to set the weights $\mathcal{W}$ so all the samples are mapped inside the hypersphere of radius R. The aim is also to minimize the radius R at the same time as setting the weights $\mathcal{W}$. The rest of the parameters in the formula above have the following function:

- c - center of the hypersphere (its position can also be a hyperparameter)
- n - number of training samples
- $\nu \in (0;1\rangle$ - hyperparameter that sets the weight of the penalty for elements that are mapped outside the hypersphere at the expense of its minimum radius
- ϕ - denotes the network itself. It uses the weights $\mathcal{W}$ to map input samples x from the input space $\boldsymbol{X}$ of a dimension $\mathbb{R}^d$ to the output space $\mathcal{F}$ of a dimension $\mathbb{R}^p$
- λ - hyperparameter controlling the rate of weight decay
- L - network layer count

The individual terms of the formula have the following function:

- First term - indicates the radius of the hypersphere that surrounds the transformed data
- Second term - Penalty for a sample that is mapped outside the hypersphere after passing through the network
- Third term - controls the rate of weight decay [7]

This optimization rule forces normal samples to be mapped close to the center of the hypersphere, while mapping anomalies farther from the center or completely outside of the hypersphere.

If it is assumed that all (or the vast majority) of the training data are normal samples, the optimization criterion can be simplified as follows [1]:

$$\min_{\mathcal{W}} \frac{1}{n}\sum_{i=1}^{n}||\phi(x_i;\mathcal{W}) - c||^2 + \frac{\lambda}{2}\sum_{\ell=1}^{L}||\boldsymbol{W}^{\ell}||_F^2 \tag{2}$$

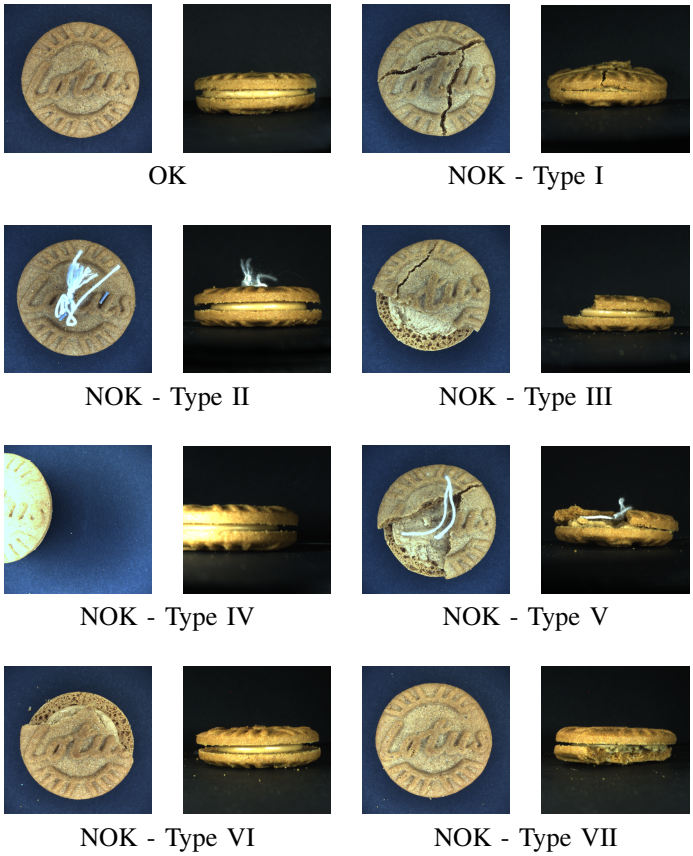


Fig. 2. Dataset Preview - NOK cookies types are as follows: Type I - Cracks, Type II - Foreign Objects, Type III - Missing Parts, Type IV - Poor Scene Composition, Type V - Combined Anomalies, Type VI - One-sided from the Top, Type VII - One-sided from the Side

After this modification, the radius of the hypersphere is no longer minimized. Now, only the distance of the mapped sample to the center of the hypersphere is minimized.

III. Dataset

For the purposes of examining the effects of different testing and training data on the final model performance a custom dataset was created to represent an industrial product on a production line. The experiments conducted in the section IV were conducted using this custom dataset.

The dataset was captured from two perspectives simultaneously so it can be explored whether adding more perspectives can improve the robustness of the anomaly detection system.

The dataset consists of more than 1300 unique instances, each consisting of a top and side view. The dataset contains 650 instances of normal samples and 650 anomalies, which are further divided into 7 categories (foreign objects, missing parts, poor scene composition, cracks, combined anomalies, one-sided top and one-sided side).

IV. Experiments

The first experiment was designed to test the null hypothesis. The aim of the null hypothesis was to test whether the network is able to model anomalies with the same level of



Fig. 3. Anomalous samples undetected by the anomaly modelling approach and at the same time detected by the normal modelling approach

accuracy as it does for the normal class. This hypothesis is based on the theoretical assumption that while the normal class is relatively well defined and individual instances of this class differ only slightly from each other, anomalous samples on the other hand can differ from each other very significantly and thus cannot be used to train the Deep SVDD model, the basic idea of which is to map similar elements as close to each other as possible while at the same time map dissimilar elements as far away from the center of the hypersphere as possible.

The ROC curve depicting model trained for this experiment can be found in the figure 4, performance metrics are located in the table I. Looking at the ROC curve, notice that with increasing number of detected anomalies (true positive rate), the number of normal instances that have been flagged as anomalous increases very rapidly (false positive rate). Note that by setting the operating point to the area of high percentage of detected anomalies (for example 90 %), the amount of normal samples flagged as anomalies climbed close to 60 %.

The figure 4 also shows ROC curve for a model that was tested on the same data, but trained using normal instances (for comparison purposes, the same amount of samples was used for training as in the previous case). This model is much more capable at detecting anomalies as it mislabels only 2 % of normal instances when set to the same true positive rate value (90 %).

In the figure 3 are shown some of the anomalies, that were mislabeled by the model that was trained with anomalous samples and also were correctly detected by the model trained with normal samples.

The experiment provided a strong evidence that anomalous features cannot be considered to be a complete class that would be as suitable for training the model as the normal class is.

TABLE I
PERFORMANCE METRICS FOR THE NULL HYPOTHESIS EXPERIMENT

Model	Normal modelling	Anomaly modelling
AUC on test dataset	0.993	0.581
Distance to the center threshold for TPR=0.9	0.742	66.134
Accuracy	0.95	0.66
Precision	0.93	0.81
Recall	0.98	0.42

Note: Performance metrics shown in the table are for the operating point defined by distance to the center threshold at TPR=0.9

The second experiment was aiming to evaluate whether the resulting estimate of the performance of the model is affected

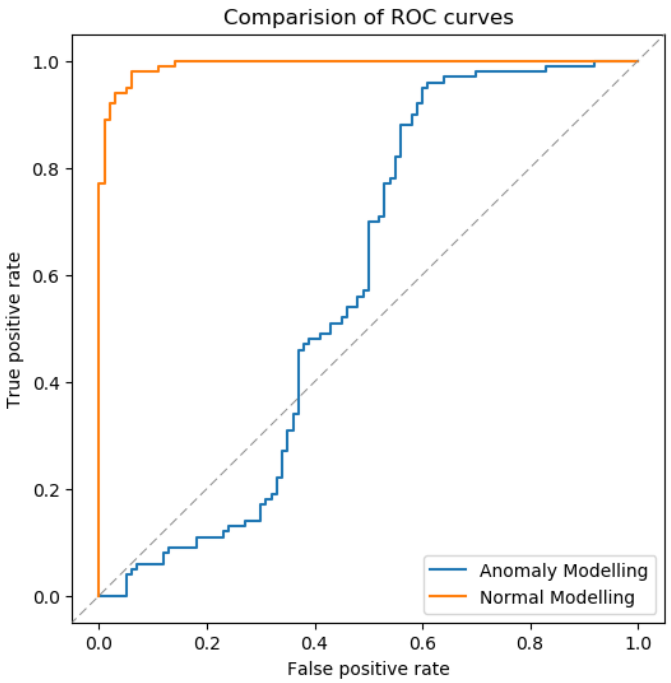


Fig. 4. ROC curve for the null hypothesis. Blue line represents the model trained on anomalous samples. Orange line represents the model that was trained to model normal samples. The variable parameter for plotting the ROC curves is threshold distance from the hypersphere center.



Fig. 5. Normal samples that were falsely detected as anomalous during the second experiment.

by the nature of the anomalous samples that are used in the testing phase, or whether the model is robust enough for the nature of the anomalies to have no effect on the resulting estimate.

For this experiment, the model was trained with 500 normal samples and then tested in three different configurations. The first configuration of the test data contained 100 normal samples and 100 anomalies equally drawn from two different anomaly classes. The second and third configurations contained the same 100 normal samples, but this time the 100 anomalies were drawn from just one of the anomalous classes.

The ROC curves can be seen in the figure 6 and the performance metrics can be found in the table II. The results suggest that using only one anomalous class tends to overestimate the resulting model accuracy, but these differences are very small and the method is still quite robust. Based on these results, drawn from a limited number of experiments, it would seem that the types of anomalies used in the testing phase do not

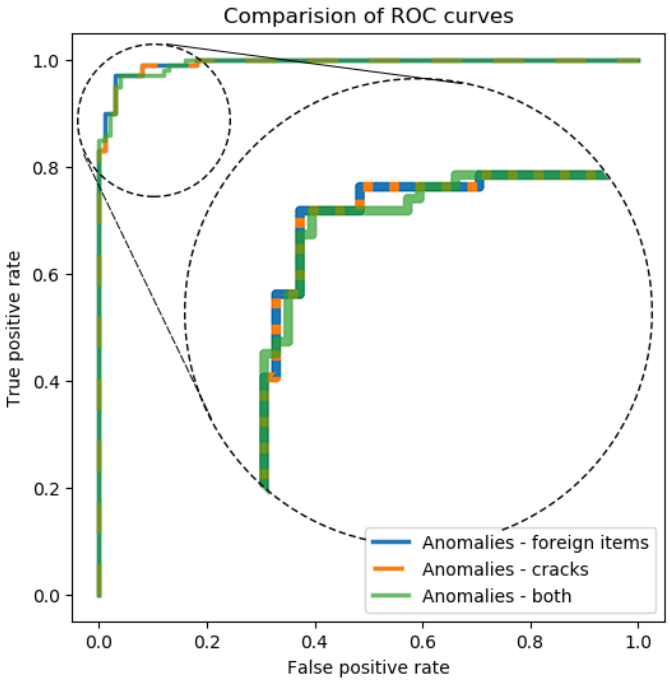


Fig. 6. The figure shows 3 different ROC curves for different testing datasets. Blue line uses foreign item anomalies, orange line uses cracked cookies as anomalies. Both ROC curves are the same in this situation. Green curve depicts model with testing anomalies of both types. It can be noted, that AUC of this setup is lower than in the previous cases but not by a great degree. The data points for the curves were obtained by changing the threshold distance from the hypersphere center.

play a significant role in the reliable performance estimation.

TABLE II
PERFORMANCE METRICS FOR THE SECOND EXPERIMENT

Used test anomalies	Both	Foreign items	Cracks
AUC on test dataset	0.993	0.994	0.994
Distance to the center threshold for TPR=0.9	0.624	0.794	0.794
Accuracy	0.94	0.95	0.95
Precision	0.91	0.91	0.91
Recall	0.98	0.99	0.99

Note: Performance metrics shown in the table are for the operating point defined by distance to the center threshold at TPR=0.9

V. Conclusion

In this paper, anomaly detection task was briefly introduced alongside with the Deep SVDD method that was used as the one-class classifier in the subsequent experiments.

Using the Deep SVDD method an experiment was conducted to test the null hypothesis that anomalies cannot be modelled as effectively as the normal class since their nature is not homogeneous as in the case of the normal class where only small differences between the samples occur. This hypothesis was confirmed and so any further efforts to train Deep SVDD using only anomalous samples were ruled out.

The second experiment was focused on the effect of the anomalies used in the testing phase on the final estimate of the performance of the model. Here, the results suggest that although reducing the diversity of anomalies leads to an overestimation of the model accuracy, this overestimation is not substantial. The results suggest that the nature of the anomalies present during the testing is not essential for calculating a reliable estimate of the performance of the model.

Acknowledgment

I would like to express my sincere gratitude to Karel Horák, my supervisor from the Machine Vision Group at the Faculty of Electrical Engineering and Communication at Brno University of Technology. His guidance, suggestions, and support were extremely valuable when I was writing this paper.

References

[1] RUFF, Lukas, Robert VANDERMEULEN, Nico GOERNITZ, Lucas DEECKE, Shoaib Ahmed SIDDIQUI, Alexander BINDER, Emmanuel MÜLLER a Marius KLOFT. Deep One-Class Classification. *Proceedings of the 35th International Conference on Machine Learning* [online]. PMLR, 15 Jul 2018, **2018**, 4393-4402 [cited 2022-11-14].

[2] HALVANI, Oren, Christian WINTER a Lukas GRANER. Unary and Binary Classification Approaches and their Implications for Authorship Verification. *CoRR*. **2019**(abs/1901.00399). Available at: http://arxiv.org/abs/1901.00399

[3] KUMAR, Ajitesh. *Difference: Binary, Multiclass & Multi-label Classification* [online]. In: . May 16, 2022 [cited 2022-12-29]. Available at: https://vitalflux.com/difference-binary-multi-class-multi-label-classification

[4] OLIVERI, Paolo. Class-modelling in food analytical chemistry: Development, sampling, optimisation and validation issues — A tutorial. *Analytica Chimica Acta*. 2017, **982**, 9-19. ISSN 00032670. Available at: doi:10.1016/j.aca.2017.05.013

[5] KHAN, Shehroz S. a Michael G. MADDEN. One-class classification: taxonomy of study and review of techniques. *The Knowledge Engineering Review* [online]. 2014, **29**(3), 345-374 [cited 2022-12-29]. ISSN 0269-8889. Available at: doi:10.1017/S026988891300043X

[6] PANG, Guansong, Chunhua SHEN, Longbing CAO a Anton Van Den HENGEL. Deep Learning for Anomaly Detection. *ACM Computing Surveys* [online]. 2022, **54**(2), 1-38 [cited 2022-11-14]. ISSN 0360-0300. Available at: doi:10.1145/3439950

[7] VASANI, Dipam. *This thing called Weight Decay* [online]. In: . Apr 29, 2019 [cited 2023-01-02]. Available at: https://towardsdatascience.com/this-thing-called-weight-decay-a7cd4bcfccab

Mutual Coupling Reduction Techniques in Microstrip Patch Phased Arrays

1st Milan Hornik
Department of Radio Electronics
Brno University of Technology
Brno, Czech Republic
xhorni15@vutbr.cz

2nd Jan Puskely
Retia a.s,
Pardubice, Czech Republic
jpuskely@retia.cz

3rd Tomas Mikulasek
Department of Radio Electronics
Brno University of Technology
Brno, Czech Republic
mikulasekt@vut.cz

***Abstract*—This paper deals with the problem of mutual coupling between antenna elements, especially in the phased antenna arrays applications. The paper includes analysis of the mutual coupling and the design of the techniques for its reduction. Two versions of the patch antenna were designed differing in feeding method, namely coaxial probe feeding and aperture feeding. Both polarizations were also considered for the analysis of the mutual coupling between the patch elements. The paper presents the state of the art of the current methods for the mutual coupling reduction, as well as the technique that we have proposed**

Keywords—Mutual coupling, Probe fed antenna, Aperture fed antenna, Phased array antenna, EBG.

I. Introduction

The paper deals with the investigation of the mutual coupling between antenna elements in the phased antenna arrays. The design of the current antenna systems becomes a much more complex issue as the demands for channel transmission capacity and speed increase. The properties of a single antenna element are insufficient in many applications, and the final antenna system is usually realized by an antenna array. In the most complex applications, each element of the antenna array is supplemented by a phase shifter that allows to steer the main beam in both axes (in azimuth and elevation). This type of antenna system is called a phased antenna array.

Modern phased antenna arrays must fulfil the requirements for the impedance bandwidth and the radiation properties (gain, polarization, sidelobe suppression, scanning range). The limiting factor for achieving acceptable antenna parameters is the mutual coupling between the individual antenna elements. The aim of this work is to analyse the mutual coupling between the antenna elements and design a suitable method for its reduction.

Many methods have already been published to reduce the mutual coupling between patch elements, such as defected ground structure (DGS) [1] , metamaterials [2]based on the principle of negative permittivity the resonant structures, such as the split ring resonators (SRR) or slots [5]. Other methods are based on Electromagnetic band-gap (EBG) structures, which exhibit a band gap for certain frequency ranges [3, 4] and supress the propagation of the EM waves. The technique of the multilayer substrate can be used too to capture the wave in the substrate [6]. There are many other methods and the final choice depends on the particular application.

II. Analysis Of Mutual Couplings Between Two Coaxially Fed Patch Antennas

The target frequency band, to study the mutual coupling, is the S-band. The designed antenna should operate at a frequency of 3 GHz with a bandwidth of at least 5%. That's why we have chosen the probe fed patch antenna as a basic element. The antenna was designed on a Rogers-5780 substrate with a relative permittivity of 2.33 and a thickness of 3.2 mm.

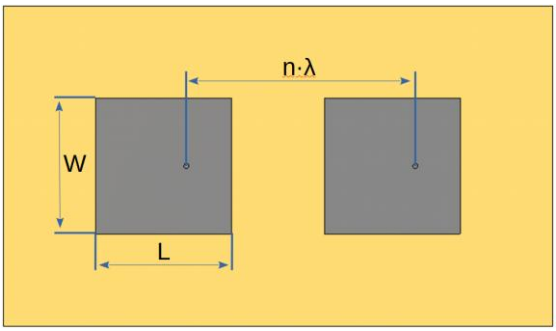


Fig. 1. The small array of the probe fed patch antennas in E-plane (W = 30 mm, L = 30 mm, probe position = 6 mm).

Firstly, to get the overview how the mutual coupling is affected by the spacing between the elements we did the parametric

study. The spacing was changing from 0.4λ to 1λ, where λ = 10 cm is the wavelength in free space at center frequency of 3 GHz. The design of the patch antennas is shown in Figure 1

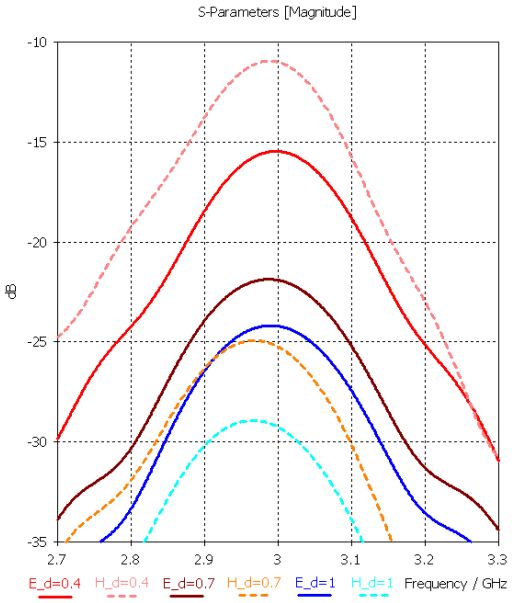


Fig. 2. The offect of the element spacing on mutual coupling.

Figure 2 shows the effect of the spacing *n* on the mutual coupling MC. Obviously, the coupling decrease with increasing the element spacing from -11 dB to -29 dB in E-plane and from -16 dB to -24 dB in H-plane, respectively. . In practice, the required MC should be lower than -20 dB depending on application.

Moreover, it should be noted that the antenna polarization plays a significant role in selecting the appropriate technique, as techniques that work well for the E-plane may not be effective for the H plane [7].

The effect of the antenna feeding methods on the mutual coupling was also investigated considering the coaxial probe method and the slot fed method. Table 1 presents the results of the MC between two antennas of 0.6lambda spacing obtained by the comparison of both methods in both planes. The obtained values were similar in all cases. The strongest coupling was observed for the slot feeding in the E-plane, whereas the weakest mutual coupling was observed for the coaxial probe feeding in the E-plane.

TABLE I. S21 IN DEPENDING ON THE TYPE OF POWER FOR 0.6 Λ

Power type	Plane	S21 [dB]
Slot fed antenna	E	-20.8
Slot fed antenna	H	-21.6
Probe fed antenna	E	-22.3
Probe fed antenna	H	-21.1

To be able to effectively reduce the effect of the MC, we need to study the origin of MC and its mechanisms of the creation. The passive patch in the small array was replaced by 3 probes as shown in Figure 3 to figure out which mechanisms of the MC play an important role..

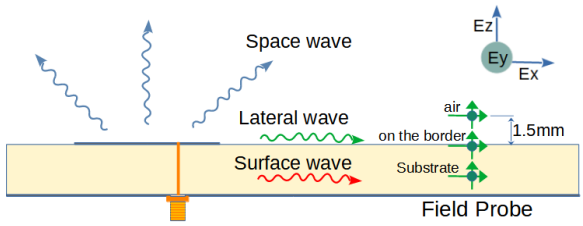


Fig. 3. The mechanisms of wave propagation between two antennas. The positions of the probes to measure the el. field.

The intensity of the electric field (V/m) was investigated using field probes located in the substrate (surface wave), on the border of air/substrate (lateral radiation), and 1.5mm above the substrate surface (space wave radiation). The analysis was performed for both planes, the E- and H-plane, and both feeding methods. The numerical results show that the horizontal component E-X is similar for both feeding methods and both planes. In case of the vertical component E-Z. It is possible to see differences in the electric field strength, particularly in the E-plane, we may notice that the Z component is significant due to lateral radiation. The most substantial distinction between the two power supply methods is noticeable in the H-plane, where the impact of the Z component is negligible for the slotted power supply and comparable to the X component for the probe power supply.

Based on the results, it can be concluded that the slot fed antenna is advantageous from the mutual coupling viewpoint. On the other hand, the coaxial fed antenna is less complex and easier to manufacture. Therefore the probe fed antenna was selected for the experimental verification of the developed technique for the MC reduction.

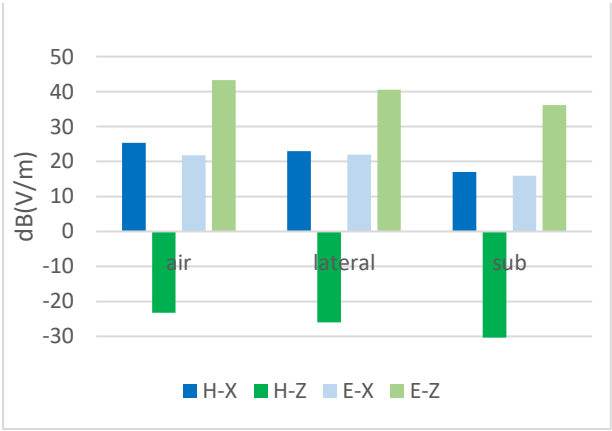


Fig. 4. Evaluation of probes to power for the slot

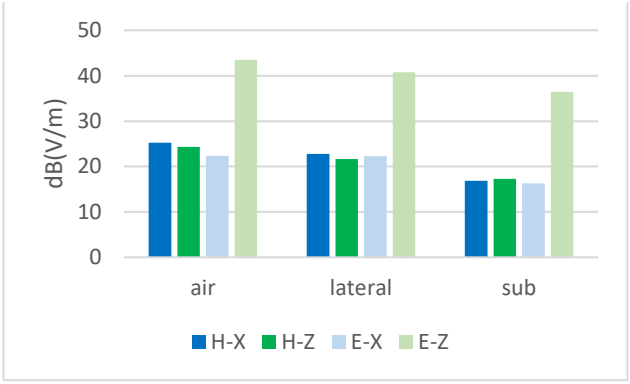


Fig. 5. Evalutaion of probes to power for the probe

III. Available Methods Of Reduction Of Mutual Coupling

According to the results of the electromagnetic field propagation analysis, it is possible to consider an appropriate technique for suppressing mutual coupling. It should be noted that for the use of phased array antennas it is necessary to avoid placing any structures directly above the patch position, therefore excluding use of the wall technique [7], which is otherwise very effective even if the spacing between patches is very small. The first investigated approach was inspired by the article [6]. It involves a technique that is utilizing a multi-layered substrate to constrain electromagnetic waves within the substrate due to the differences in substrate permittivity. Results from the articles were successfully reproduced. However, it should be noted that this technique is not very effective, as the reduction of the transmission coefficient was only -3 dB, as shown by the article. Although it is a small reduction, this approach is beneficial as it can be combined with other techniques The mentioned technique is effective for the E-plane, but for the H-plane, the influence of this technique is minimal.

Dispersion diagrams are frequently used to analyse EBG (Electromagnetic Bandgap) structures, where they are illuminated by a plane wave with gradually changing radiation direction. Using this method, it is possible to observe the formation of a band gap in proposed EBG structure, which is characterised by a region where no mode is excited [7]. Firstly, the structures were tuned using Eigen-Mode in the CST Studio Suite program. However, after inserting the EBG structure between two patches, no influence was observed. It should be noted that EBG structures are classified as periodic structures and are therefore in interaction with each other's. Due to small spacing it was not possible to fit more rows of EBG structures between the patches.

To minimize the size of the EBG (Electromagnetic Band Gap) patches, two different techniques have been tested [3,8]. First technique consisted of inserting vias into a known square patch, resulting in a 30% decrease in structure size [8].

The second tested technique makes use of DUC-EBG [3], which significantly reduces the size of the structure. Originally designed for the frequency of 5.5 GHz, therefore structure had to be enlarged from 7.2 mm to 9.36 mm. While the structure itself did not exhibit any effect between the patches, when a second row was added, resonances were observed at the resonance frequencies. This slightly reduced the transmission coefficient and negatively impacted reflection coefficient of both patches, which resulted in oscillations.

IV. Proposed Method

Method used in this study was based on the reference article [4], where three substrates were considered instead of the two substrates used in this article, as shown in Figure 4. EBG structures were placed between the patches on the basic substrate, and the outermost EBG structures are provided with vias. The structure proposed in Figure 5 is valid for the E-plane, but it is not very effective for the H-plane. Simple square EBG patch was placed in the middle of the elements and it's resulting band gap is shown in Figure 6, ranging from 2.5 GHz to 3.2 GHz for mode 1 and 2 respectively, and 3.2 GHz to 3.75 GHz for mode 3. Second EBG structure with vias was implemented for the second band gap, and its band gap is shown in Figure 7, ranging from 2.95 GHz to 3.75 GHz. Both EBG structures have the same dimensions. Two small patches are inserted between the EBG patches to suppress the resonance that occurs at 2.7 GHz. Top view of the structure is shown in Figure 8. Figure 9 shows the reflection coefficient for patch 1 is displayed in red and Patch 2 in orange. Transmission between patches 1 and 2 is characterized by the green color, and the transmission between patch 2 and 1 is represented by the blue color. During the design process, emphasis was placed on achieving sufficient bandwidth for mutual coupling suppression, which was achieved with a value of 500 MHz.

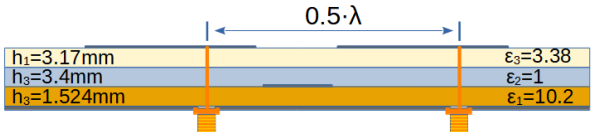


Fig. 6. Cut through of the design

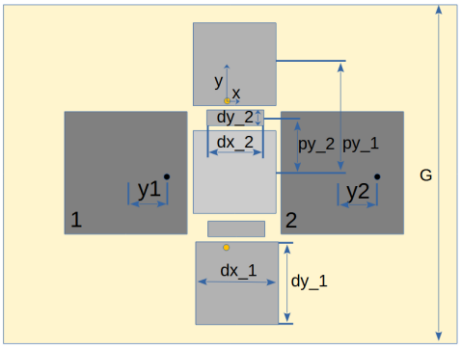


Fig. 7. Top view of small array of the probe fed patch antennas

Structure is composed of three substrates with the thickness and permittivity values of h1=3.17mm, ε1=10.5, h2=3.42mm, ε2=2, and h3=1.524mm, ε3=3.38. Dimensions of the patch element are W=L=30.6 mm, G= 100 mm, with a distance of y1=8.5 mm for the feeding probe from the center, and y2=9.9 mm for the other. Spacing between the elements is λ*0.5, and for the small EBG patch, dx_1= dy_1= 17.5 and py_1=26, and for the small patches, dx_2=9.6mm, dy_2=3.3mm, and py_2=13. Vias are placed on the outermost patches, which were shifted from the center of the patch by x=1.43mm and y=18.4mm.

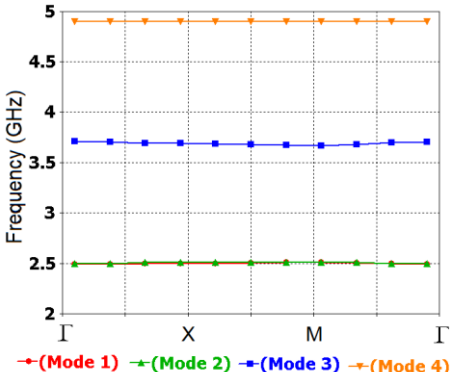


Fig. 8. Dispersion diagram for simple EBG

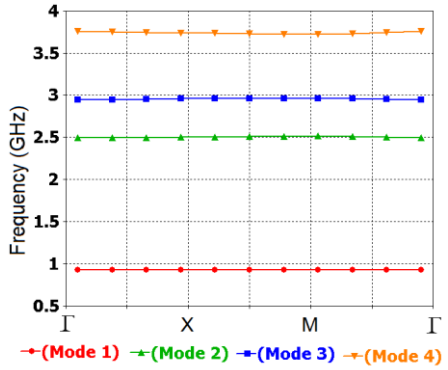


Fig. 9. Dispersion diagram for EBG patch with via

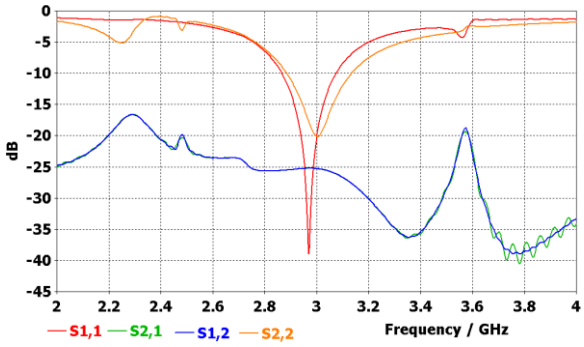


Fig. 10. Simulated S-parametrs of the probe fed patch antennas

V. Conclusion

Method presented in [4] has been further modified to meet the requirement of achieving a minimum reduction in reflection factor of -25 dB. It should be noted that this method is not applicable to H-plane orientation where achieving sufficient isolation for radar scanning purposes is more important. Future work will require experimental verification of results and validation through simulations. Additionally, further efforts will be made to develop a technique that enables adequate isolation in both the E and H planes, especially in the H-plane and to a lesser extent in the E-plane.

Acknowledgment

This work was supported by the Internal Grant Agency of the Brno University of Technology under project no. FEKT-S-23-8191.

References

[1] KIANI-KHARAJI, Mojtaba, Hamid Reza HASSANI and Sajad MOHAMMAD-ALI-NEZHAD. Wide-angle phase antenna with mutual coupling reduction. IET Microwaves, Antennas & Propagation [online]. 2018, 12(12), 1932-1938 [cited 2023-03-30]. ISSN 1751-8733. Available from: doi:10.1049/iet-map.2018.0155.

[2] KEDZE, Kam E., Wenyu ZHOU, Nima JAVANBAKHT, George XIAO, Jafar SHAKER, and Rony E. AMAYA. Implementation of complementary split-ring resonators for mutual coupling suppression in a dual differentially fed microstrip antenna. In: 2022 IEEE International Symposium on Phased Array Systems & Technology (PAST) [online]. IEEE, 2022-10-11, pp. 1-2 [cited 2023-03-30]. ISBN 978-1-6654-4166-7. Available from: doi:10.1109/PAST49659.2022.9975013.

[3] ALAM, Md. Shahidul, Norbahiah MISRAN, Baharudin YATIM, and Mohammad Tariqul ISLAM. "Development of Electromagnetic Band Gap Structures in the Perspective of Microstrip Antenna Design." International Journal of Antennas and Propagation, vol. 2013, pp. 1-22, 2013. Accessed March 7, 2023. ISSN 1687-5869. Available from: doi:10.1155/2013/507158.

[4] RAJO-IGLESIAS, Eva, Óscar QUEVEDO-TERUEL, and Luis INCLAN-SANCHEZ. "Mutual Coupling Reduction in Patch Antenna Arrays by Using a Planar EBG Structure and a Multilayer Dielectric Substrate." IEEE Transactions on Antennas and Propagation, vol. 56, no. 6, pp. 1648-1655, 2008. Accessed March 10, 2023. ISSN 0018-926X. Available from: doi:10.1109/TAP.2008.923306.

[5] ALSATH, M. Gulam Nabi, Malathi KANAGASABAI and Bhuvaneshwari BALASUBRAMANIAN. Implementation of slotted meander resonators for enhancing isolation in microstrip antenna arrays. IEEE Antennas and Wireless Propagation Letters [online]. 2013, 12, 15-18 [cited 2022-12-27]. ISSN 1536-1225. Available from: doi:10.1109/LAWP.2012.2237156.

[6] ASLAN, Yanki, and Alexander YAROVOY. "Reduction of mutual coupling between closely spaced patch antennas using dielectric stratification technique." In: 2017 47th European Microwave Conference (EuMC), IEEE, pp. 248-251, 2017. Accessed March 7, 2023. ISBN 978-2-87487-047-7. Available from: doi: 10.23919/EuMC.2017.8230846.

[7] TUNIO, Irfan Ali, Yann MAHE, Tchanguiz RAZBAN-HAGHIGHI, and Bruno FROPPIER. "Mutual Coupling Reduction in Patch Antenna Array Using Combination of Shorting Pins and Metallic Walls." Progress In Electromagnetics Research C, vol. 107, pp. 157-171, 2021. Accessed March 7, 2023. ISSN 1937-8718. Available from: doi:10.2528/PIERC20082803.

[8] RAJO-IGLESIAS, Eva, Luis INCLAN-SANCHEZ, Jose-Luis VAZQUEZ-ROY, and Enrique GARCIA-MUNOZ. Size Reduction of Mushroom-Type EBG Surfaces by Using Edge-Located Vias. IEEE Microwave and Wireless Components Letters [online]. 2007, 17(9), 670-672 [cit. 2023-03-29]. ISSN 1531-1309. Available from: doi:10.1109/LMWC.2007.903456.

Raspberry Pi Bee Health Monitoring Device

1st Jakub Nevláčil, 2nd Šimon Bilík, 3rd Karel Horák
Department of Control and Instrumentation
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
xnevla00@vutbr.cz, bilik@vut.cz, horak@vut.cz

***Abstract*—A declining honeybee population could pose a threat to a food resources of the whole world one of the latest trend in beekeeping is an effort to monitor a health of the honeybees using various sensors and devices. This paper participates on a development on one of these devices. The aim of this paper is to make an upgrades and improvement of an in-development bee health monitoring device and propose a remote data logging solution for a continual monitoring of a beehive.**

***Index Terms*—Appis melifera, Remote monitoring, Raspberry Pi, IoT, MQTT**

I. Introduction

Honeybees are crucial for our ecosystem and also for food production around the globe. One of the main purposes of the honey bees is to pollinate flowers and plants and produce honey. The pollination of the plants plays a key role in the agriculture, because without the bees the plants could not reproduce and grow its fruits or other crops. [1] Beekeeping is a thousands of years old craft and as times goes the bees appear to be more and more endangered. Whether it is endangering by pests and diseases or today especially by pesticides from nearby agricultural fields. [2]

This is the reason why it is important for the beekeepers to be aware of the honeybees health and activity in real time and in case of an irregular bee behavior react upon this danger as soon as possible. Therefore this project aims to modify a device that collects useful data from a beehive to be able to log this data to a database where the data could be visualized using a graphical interface. Later the data will be used to determine the health and the activity of the bees using machine learning algorithms.

II. Data acquisition device

A. Original data acquisition device

The device for data acquisition is based on a previous work of [4] and [3]. The original concept of the data acquisition device can be seen in the Fig. 1. This device is based on a Raspberry Pi platform with a Grove Base Hat for connecting sensors to capture weather and atmospheric conditions inside and outside of the beehive. Furthermore this device also monitors the traffic of the incoming and outcoming bees using the built-in camera module and takes sound samples from the inside of the beehive with a microphone.

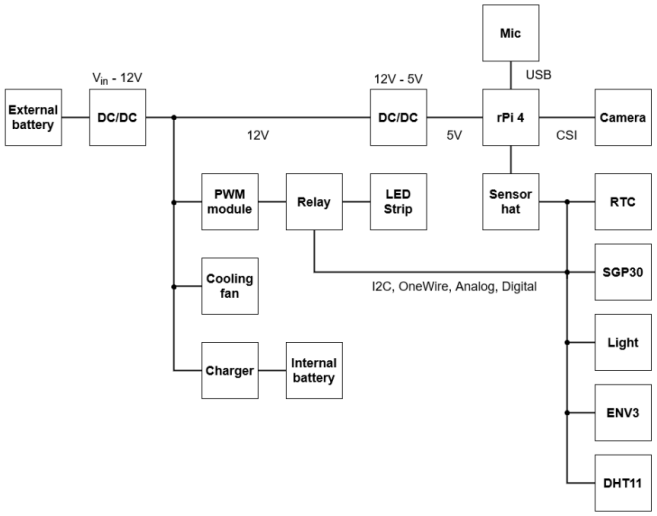


Fig. 1. Original data acquisition device

B. Modified data acquisition device

Although the data acquisition from the sensors and microphone usually occurs every 15 minutes (the interval can be adjusted, but there is most likely no use in collecting the data more frequently), the SGP30 sensor that measures CO2 concentration and TVOC (total volatile organic compounds) inside the beehive needs to be activated every second in order to provide correct readings. This task can be unnecessarily demanding on computing power, which can be later utilized for later implementation of the machine learning algorithms for determining health and activity of the bees.

For this purpose a Raspberry Pico was incorporated into the data acquisition chain. It was connected to the Raspberry Pi over USB port and serial communication between the two devices was established. This allows to transfer all the weather and atmospheric data from the sensors to the Raspberry Pico using the corresponding sensor Hat for this device. The Raspberry Pico is programmed in such a way that it acquires data from all the weather and atmospheric sensors every second. Then once the Raspberry Pi asks for the measured data, the Raspberry Pico calculates an average value over the inter-request period of the data from every sensor and sends them to the Raspberry Pi. The outcome of this modification is that the

Raspberry Pi does not have to dedicate an entire computational thread to call the SGP30 sensor every second and can use this thread for other purposes. It should also increase reliability of the camera capturing.

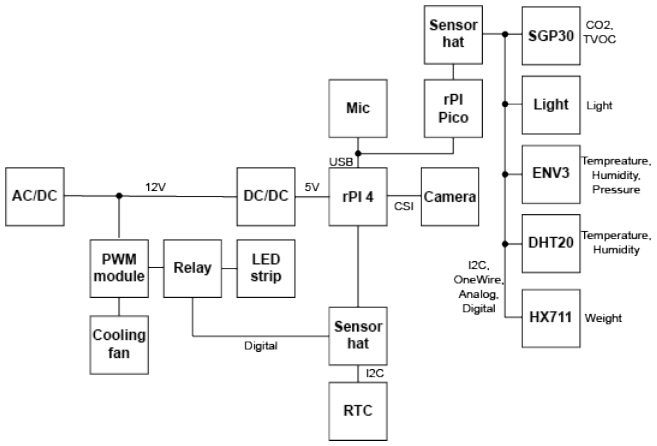


Fig. 2. Modified data acquisition device

In the process of modifying, the DHT10 sensor which measures the temperature and humidity inside of the hive was updated to a newer version DHT20. This modification should bring a higher measurement accuracy and a greater measurement range. In addition to the weather and atmospheric sensors was added a HX711 amplifier with 4 load cells in full bridge configuration. This sensor will later be used for measuring the weight changes of individual beehives. A scheme of our modified data acquisition device in its final configuration can be seen from Fig. 2. In the Fig. 3 we can then see a real implementation of the data acquisition device inside a box with a tunnels for the passage of the bees.

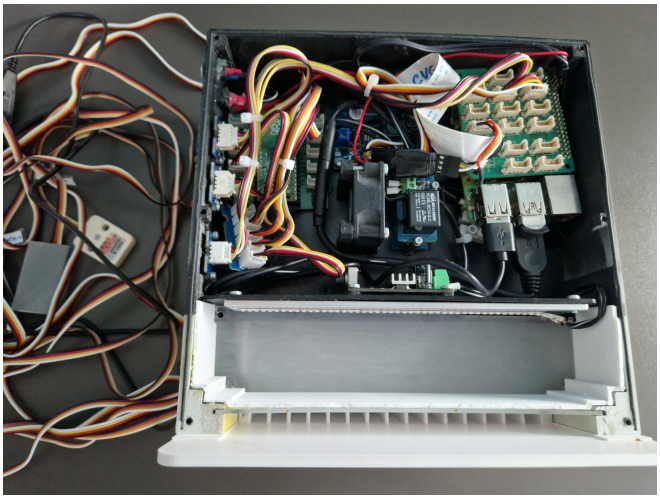


Fig. 3. Modified data acquisition device inside a box

III. Logging data to a database

A. Internet connection

After acquiring the useful data from the hive, we want to send them over the internet to a database where we can look at the and assess the state of the hive from anywhere we like. This of course requires an internet connection. Due to the fact that beehives are usually placed on remote locations, where Wi-Fi or wired internet connection is not an option, this can be a problem. Luckily the place of deployment of our devices should have sufficient Wi-Fi coverage. In situations where a Wi-Fi connection cannot be established, solutions as GSM module or LoRaWAN could be used easily.

B. MQTT messages

The device currently stores all the acquired data from the hive on an USB flash drive. This includes sound samples from a hive, pictures of the bees and data from the weather and atmospheric sensors. For the purpose of assessing the state of health of the beehive, only the data from the weather sensors needs to be logged to the database. In future this will also include the outputs from a machine learning models that should quantify the health of the beehive and activity of the bees. Due to this fact that only a few variables needs to be transmitted to a database, the MQTT messages were chosen for this purpose.

MQTT is a protocol used in IoT. One of the main advantages of the MQTT protocol is its lightweight design. Usually this protocol is used on a microcontrollers. This means that the MQTT protocol requires minimal computational resources and also the internet connection requirements are relatively low. [5] The MQTT space consists of two kind of devices, MQTT clients and a MQTT broker. MQTT clients can either subscribe or publish messages. MQTT broker controls the flow of the messages by receiving messages from publishing clients and forwarding them to the clients that are subscribed to a particular topic. For an easier understanding an example of a MQTT network can be seen in Fig. 4. The MQTT broker can be hosted on a cloud or it can be installed on a local device.

In our particular case the MQTT broker will be cloud based. This solution has been assess as more suitable for our particular problem. The upside of this solution is that it doesn't require a public IP address or a VPN server on the side of the subscribed client. In conclusion our setup will consist of a Raspberry Pi data acquisition device which will be configured as publishing client, then cloud based MQTT broker and one Raspberry Pi which will be configured as a subscribing client. The subscribing Raspberry Pi will also act as a database and will run a web application for displaying the logged data.

C. InfluxDB database

As mentioned above, the MQTT messages will be sent from the data acquisition device to an MQTT broker. All of the data from the sensors representing a specific period of time will be sent in a single message stored as a JSON object. The MQTT broker will the distribute every message to the subscribed clients. In our case the subscribed client will be a Raspberry

Pi device where the messages will be received a processed by Telegraf, which is a data collection agent. Telegraf the takes the message sorts the data from individual sensors and saves them into an InfluxDB database. The reason the InfluxDB database is used instead of other popular databases like MySQL is because The Influx DB database is a time series database. This means that this database is designed to store a data which are associated with time. This makes it the best choice for our application of plotting the measurements from sensors over time.

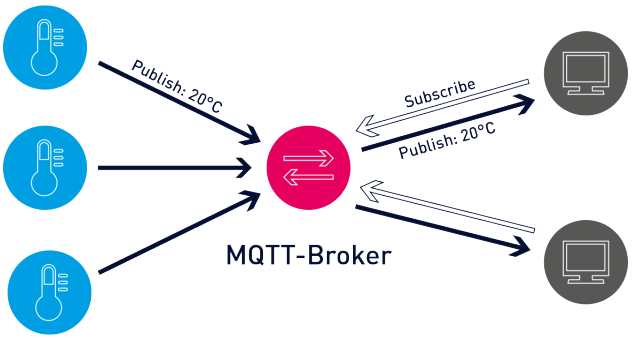


Fig. 4. MQTT architercture [6]

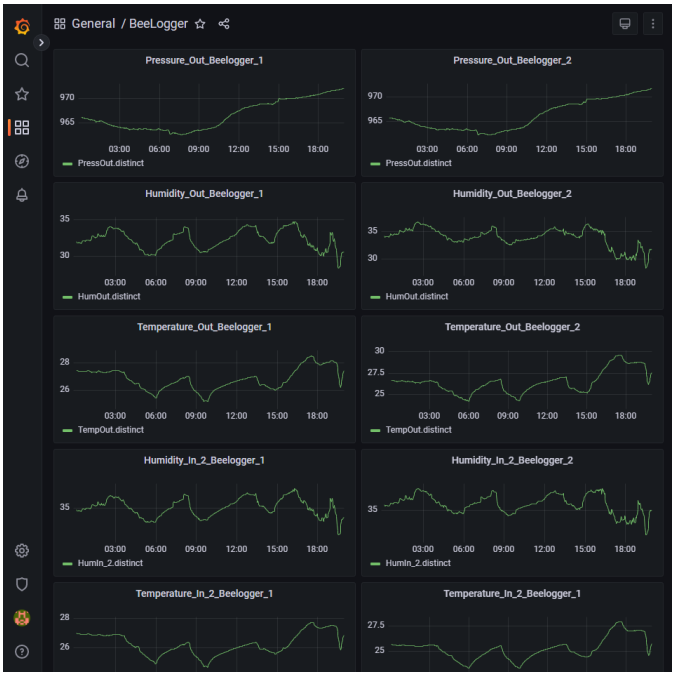


Fig. 5. Grafana dashboard

D. Grafana

Once the measurements from the sensors are saved in a database, we can them plot the values from individual sensors over time. For this purpose Grafana will be used. Grafana is a web application for visualizing data and metrics. This application will also run on the Raspberry Pi next to the InfluxDB database and Telegraf and will read metrics from the sensirs directly from the database and plot them into a graphs. The Grafana application can be accessed from a web browser on a device that is connected to the same local network as the device it is installed on. Grafana can also send notifications on Slack and email. This feature could be useful once the machine learning models for recognition of bee activity and health will be implemented to warn the beekeepers about emergency situations in the hive.

In the Fig. 5 we can see a Grafana dashboard with two sets of plots. This is a data from two prototypes of the data acquisition device that was described above. This data was collected during testing phase before deployment of the devices on a real beehives.

IV. Conclusion

In this paper we modified previous work of [4] to use a Raspberry Pico for collecting data from weather and atmospheric sensors to release some computational resources of a Raspberry Pi. An upgrade of a sensor was made for a newer version with better measuring accuracy and also a new weight sensor for measuring a weight changes of a beehive.

Then for an easier remote monitoring of the beehives a data logging solution was implemented. This solution starts at MQTT messages that are being sent from the data acquisition device to a cloud based MQTT broker from where the messages travel to a server with Influx DB database installed. As the last step the data is visualized using a web application Grafana, from where the visualized data can be accessed through a web browser.

In the future work an analysis of a sensor data fusion collected during a deployment on real beehives is planned. This analysis should include especially cluster analysis of a bee buzz audio samples, but also other machine learning techniques applied to both audio data and also sensor data.

Acknowledgment

The completion of this paper was made possible by the grant No. FEKT-S-23-8451 - ”Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence, automation and measurement” financially supported by the Internal science fund of Brno University of Technology.

References

[1] Why are Bees Important?. Friend of the Earth [online]. [cit. 2023-03-12]. Available at: https://foe.org/blog/why-are-bees-important/

[2] MEIKLE, William G., Guy MERCADIER, Fatiha GUERMACHE a Marie-Claude BON. Pseudomonas contamination of a fungus-based biopesticide: Implications for honey bee (Hymenoptera. Biological Control [online]. 2012, 60(3), 312-320 [cit. 2023-03-12]. ISSN 10499644. Available at: https://doi.org/10.1016/j.biocontrol.2011.12.004

[3] MVG-BVI: Bee Visual Inspector [online]. In: . s. 2 [cit. 2023-03-12]. Available at: http://vision.uamt.feec.vutbr.cz/data/MVG_2021_BVI_tech_spec.pdf

[4] Bee-Health-Monitor. GitHub[online]. 2022 [cit. 2022-12-15]. Available at: https://github.com/boortel/Bee-Health-Monitor.

[5] MQTT. MQTT [online]. [cit. 2023-03-09]. Available at: https://mqtt.org/

[6] IT Explained: MQTT. In: Paessler [online]. [cit. 2023-03-10]. Available at: https://www.paessler.com/it-explained/mqtt

NXP SEMICONDUCTORS CZECH REPUBLIC



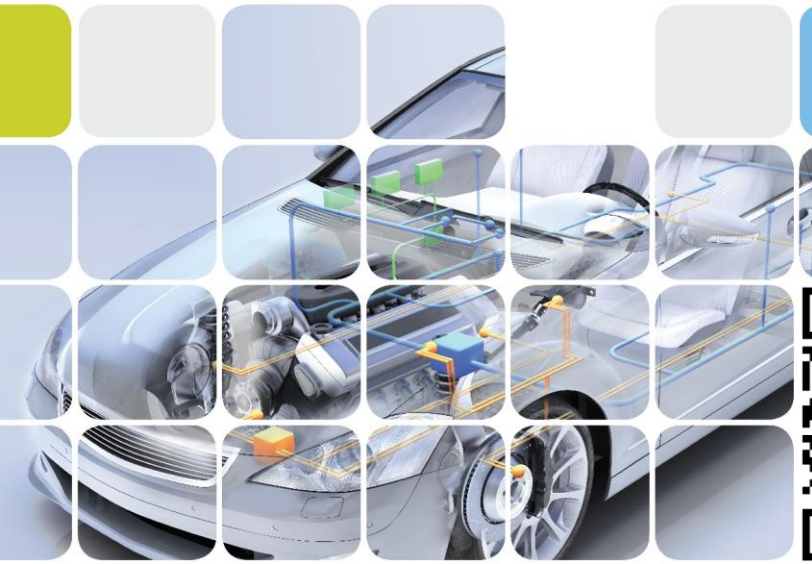
NXP Semiconductors se zabývá vývojem a výrobou mikroprocesorů a analogových součástek pro automobilový a spotřební průmysl, počítačové sítě a bezdrátové technologie. V České republice je společnost NXP Semiconductors zastoupena vývojovými centry **v Rožnově pod Radhoštěm** a v **Brně**. Nově jsme otevřeli i kanceláře na UTB **Zlín** a VŠB TU v **Ostravě**.

NXP **je světová špička v oblasti NFC technologií, Wireless charging či Automotive Electronics**. Kariéra ve firmě NXP Czech Republic znamená **příležitost spolupracovat s těmi nejlepšími odborníky**, v přátelském R&D prostředí.

V moderních laboratořích pracuje mezinárodní tým, čítající více než **400 embedded software odborníků**, z oblastí Wireless charging, Motor Control, Security and Connectivity, Functional safety, 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.

NXP SEMICONDUCTORS CZECH REPUBLIC

Hledáme talentované STUDENTY



Studentům nabízíme:

- Spolupráci s kapacitami svého oboru
- Práci na reálných projektech, technický mentoring
- Výborné mzdové ohodnocení a časovou flexibilitu
- Pomoc s bakalářskou, diplomovou či disertační prací
- NXP Cup – soutěž autonomních autíček s hlavní cenou 10 000 USD

Jak se zapojit?

Přihlaste se na studentské či absolventské pozice skrze QR kód výše, nebo nás kontaktujte na:

josef.halamik@nxp.com

www.nxp.com/czechrepublic

Intelligent Wattmeter With Network Communication Option

1st Daniel Haluška
Institute of Telecommunications
Brno University of Technology
Brno, Czech republic
xhalus00@vutbr.cz

2nd Martin Rusz
Institute of Telecommunications
Brno University of Technology
Brno, Czech republic
xruszm00@vut.cz

***Abstract*—The aim of the paper is to present the realization of a measuring instrument for measuring AC voltage, current and calculation of AC power. The created device uses a voltage divider with DC offset and the implementation of the ACS712 circuit. The obtained quantities are presented to the user using a web interface. The microcontroller sampling and electrical power calculation and display to the user is presented in this paper. The measurement of the fabricated device is summarized in the paper.**

***Index Terms*—ACS712, Arduino, Electric current, Electric power, Ethernet, LCD display, Mains voltage, Measuring device, Voltage divider with DC offset, Wattmeter, W5100**

I. Introduction

The purpose of the device is to monitor and streamline the electricity consumption of a common appliance. The measurement of electricity is carried out in a normal household by a billing meter, which provides information on the total electricity consumption. However, the total consumption figure is not usable for a single appliance overview which is mediated by our wattmeter. The aim of the device was not to achieve the highest possible measurement accuracy, but rather to implement the proposed solution on hardware and make a web application with the possibility for the user to interactively monitor the measured parameters.

II. Measurement of variables

The goal is to detect electrical quantities as accurately as possible using a robust and simple measurement method. The subsequent evaluation of the different types of performance is in digital circuits by calculation with voltage and current samples. When sampling the mains voltage with a microprocessor, it is necessary to reduce the input level of the analogue input to a level of 0 - 5 V. For space reasons, a voltage divider implenting offset is used in the measuring instrument due to the level of the mains sinusoidal voltage. The current measurement is mediated by the Allegro ACS712 circuit.

A. Current sensor

The ACS712 integrated circuit provides economical and accurate measurement of AC or DC current. It consists of a precision, low offset, linear Hall effect sensor with a copper conductive path placed adjacent to it. Current flowing through the conductive path with a resistance of typically 1.2 mΩ generates a magnetic field that is sensed by the Hall sensor. The manufacturer produces the circuit in three variants differing in the maximum value of the sensed current: 5 A, 20 A and 30 A. Depending on the variant, the sensitivity of the circuit varies for 5 A 182 mV/A, for 20 A 100 mV/A and for 30 A 60 mV/A. The circuit is powered by an external voltage of 5 V [1].

For the measurement of the current passing through the measuring device, a variant with a maximum sensed current value of 30 A was chosen, so that the current measurement of 16 A, which is given by the fusing element of a conventional socket, is possible.

B. Safety

The equipment operates with mains voltage and its failure may result in damage to the equipment or the risk of electric shock. Circuit simulation in PSpice will show us how the circuit and the individual circuit elements will behave. In the transient simulation we can observe the supply voltage 230 V in red curve, the output of the voltage divider in green and the power ripple caused by the output of the divider in blue curve. From the simulation, which can be seen in (Figure 1), we can see that the divider has the desired output and does not interpolate the mains frequency into the supply branch.

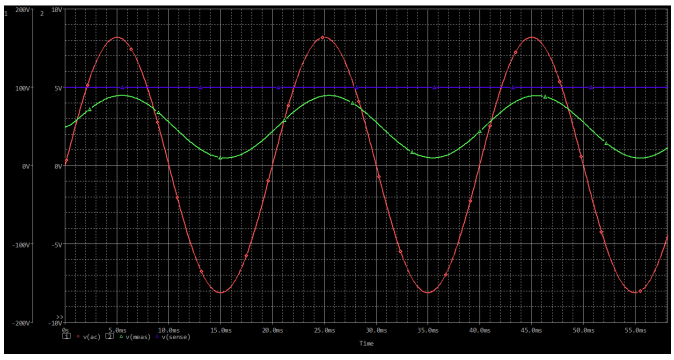


Fig. 1. Simulation mains voltage part

C. Calculation of variables

Sampling and calculation of electrical quantities on sensors is mediated by the EmonLib library. The library is designed for

different platforms (Arduino, Raspberry, ESP32). It provides convenient measurement of AC current, voltage and from it power calculation. It performs the measurement using the two A/D inputs of the microprocessor. One input is used for sensing the mains voltage from the voltage divider. The other one is connected to the output of the ACS712 integrated circuit. The signal at the input, both voltage and current, is sampled and the effective value is calculated from the samples. The effective RMS value is important because it represents the value of the effect of the DC voltage or DC current, which will give the same thermal effect in the resistor as the value of the AC current or voltage. An explanation of the calculation principle, which is mediated by the EmonLib.cpp library, is presented in the following [2].

D. Equations:

- **Calculation of effective RMS voltage:**
  The library calculates the effective voltage of any input signal using a graphical method. The equation of the graphical method of calculation:

$$U_{RMS} = \sqrt{\sum U_s^2}\,[V] \tag{1}$$

  The root-mean-square is calculated in the way the name suggests. First we square the quantity, then we calculate the mean and finally, the square-root of the mean of the squares, this is how it is done.
- **Calculation of effective RMS current:**
  The calculation proceeds in the same way as for the voltage calculation, but with current samples:

$$I_{RMS} = \sqrt{\sum I_s^2}\,[A] \tag{2}$$

  First, each current sample I_s is amplified, then the average of the amplified samples is calculated, and finally this average is subtracted .
- **Calculation of apparent power:**
  Apparent power is calculated from effective values in the library:

$$S = U_{RMS} \cdot I_{RMS}\,[VA] \tag{3}$$

  The apparent power value is obtained as the product of the effective voltage value U_{RMS} and the effective current value I_{RMS}.
- **Calculating of real power:**
  The real power is expressed from the voltage and current samples alone using the following equation :

$$P = \sqrt{\sum U_s^2 \cdot I_s^2}\,[W] \tag{4}$$

  It is calculated as the average of the sum of the products of the instantaneous values of the voltage U_s and the current I_s of each sample.
- **Calculating of power factor:**
  The factor expressing the ratio between real and apparent power is given by :

$$cos\,\varphi = \frac{P}{S}\,[-] \tag{5}$$

III. Communication system design

One of the main goals is to create a system for transferring data from the measuring device to the server that will process and store the data. For most of such systems, the specificity is the production of excessive amounts of data that need to be resolved.

The proposed communication system can be divided into several parts:

- Measuring instrument
  Implementation is device that measures electrical quantities and sends them via HTTP protocol.
- Connection of the device
  It is realized by W5100 modul connected to LAN internetwork with DHCP protocol for initializing IP adresses.
- Server
  The following block receiving data from the measuring device, serve as a data repository (database) and represent the measured data as a web application.
- User
  Represents the measured values as a web page. Connects to the server running the web application.

The logical connections between the different parts of the communication chain can be seen in the following model (Figure 2).

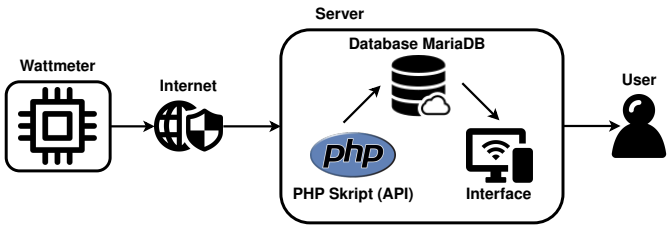


Fig. 2. Communication model

IV. Web interface

An important part of the interface for the user is the display of the measured values in an accessible and understandable form. Among the requirements for interactive visualization is the possibility of using two web applications Grafana and Nodered. These web applications are popularly used for creating graphs. Nodered was selected for the facility.

A. Nodered

Node-RED is an open source development tool for hardware device integration. It is developed by IBM Emerging Technology. Node-RED is a free JavaScript-based tool built on the Node.js platform, a visual web-based flow editor is used to edit the program. The flow editor contains nodes that are represented by their respective icons. Adding icons can work in two ways namely by dragging, moving and joining nodes or by importing JavaScript code as a function [3].

B. Device user interface

The user interface is built on top of the Node-RED web application. It consists of four pages.

On the first page (Fig. 3), called Measurement, it is possible to see the current measured values of voltage, current and power on the needle pointers and below them there is a graph plotting the measured values for the last hour. In addition to the measured values, the Load On and Load Off buttons on the first page control the load switching relays.



Fig. 3. Web application homepage.

The second Graph page (Figure 4) is used to represent the data in a graph. The top bar controls the source data and the timeline that is plotted. The first drop-down menu selects the variables that are displayed.

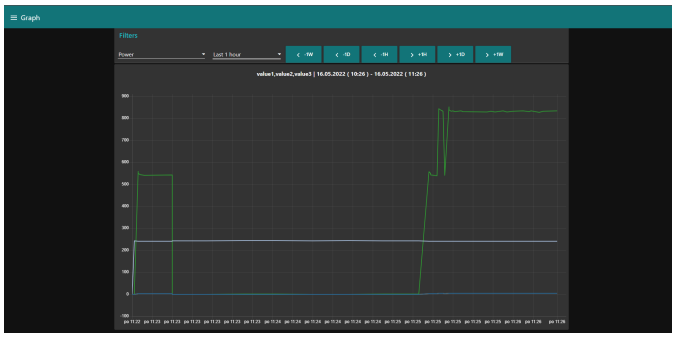


Fig. 4. Multifunctional graph of web application.

The third Trends page (Figure 5) shows trends in electricity consumption. The page shows the total consumption and partial consumption for 30 days, 7 days, day and hour. The sub-consumption is compared to the same period and the percentage of consumption matching the period is evaluated according to the consumption.

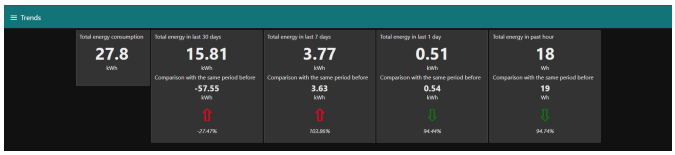


Fig. 5. View electricity consumption trends.

The last Alerts page (Figure 6) is dedicated to setting the trigger levels for alerts. If a measurement exceeds the set value, the user is alerted by a pop-up window with a text message informing what the exceedance is.

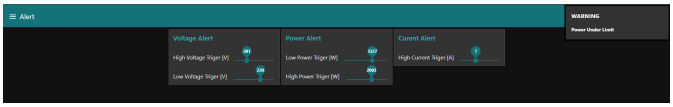


Fig. 6. Alert and warning settings.

V. Hardware

A. Component wiring block diagram

This figure shows the logical interconnection of the individual hardware components (Figure 7). The SPI and I2C buses are used to connect the dysplay and ethernet modul peri. The entire device is powered by a MeanWell 5V switching power supply.

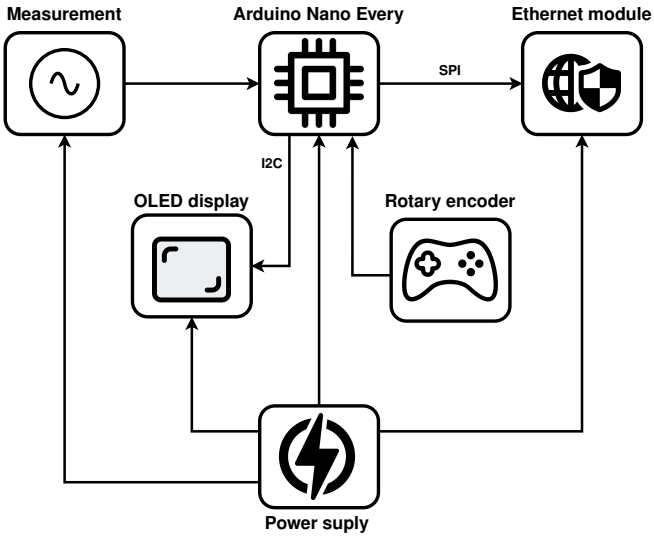


Fig. 7. Block diagram

B. Printed circuit board

The design of the PCB (Figure 8) was done in the PCB editor Altium, where the complete schematic of the measuring instrument was also processed. The PCB was designed so that the complete measuring device with all parts on one PCB including the source part, display and encoder. The only expansion device that will be connected by a connector is the Ethernet module brought out by the SPI bus. In PCB design we tried to create the largest possible gaps between coper traces to maintain safety [4].

C. Device box

The wattmeter was mounted in a universal plastic box as shown in (Figure 8). The semi-finished cabinet consists of two parts connected at each corner by a self-tapping screw. The top part consist of display, encoder and 16A panel socket.



Fig. 8. A prototype of the device and PCB

D. Mesurement precision

Measurement precision is defined as the amount of agreement between the measurement result and the conventionally true value of the measurand. The accuracy of the measurement was compared against a Voltcraft VC890 reference multimeter and the subsequent calculated deviations can be seen in the table below (Table I).

The absolute measurement error $\triangle_x$ is expressed as the difference between the measured value X_m and the conventionally true value X_p. The true value cannot be detected for physical reasons, so it is replaced by the conventionally true value. The difference between the true value and the conventionally true value is that the conventionally true value is close to the true value with a negligible difference [5].

$$\triangle_x = X_m - X_p[unit\,of\,X] \tag{6}$$

The relative measurement error δ_x is expressed as the fraction of the absolute measurement error $\triangle_x$ compared to the measured value X_m. The relative measurement error is used to evaluate the accuracy of the measurement method [5].

$$\delta_x = \frac{\triangle_x}{X_m} \cdot 100[\%] \tag{7}$$

TABLE I
COMPARISON OF WATTMETER AND MULTIMETER READINGS

Load	Absolute meas. deviation			Relative meas. deviation		
	$\triangle_U[V]$	$\triangle_I[I]$	$\triangle_S[VA]$	δ_U [%]	δ_I [%]	δ_S [%]
LED 9 W	-0,42	-0,004	-0,60	-0,181	-8,400	-4,976
Bulb 105 W	-0,18	0,000	0,78	-0,077	0,053	1,029
Bulb 35 W	-0,20	0,006	0,71	-0,085	3,563	1,915
Bulb 50 W	-0,10	0,004	0,00	-0,043	1,565	0,000
Dryer I.step	-0,05	-0,004	-2,97	-0,021	-0,377	-1,210
Dryer II.step	-0.06	0,005	2,03	-0,026	0,258	0,448
Dryer III.step	-0,14	-0,012	-3,14	-0,060	-0,354	-0,398
Dryer IV.step	-0,02	-0,043	-3,48	-0,009	-0,987	-0,345

VI. CONCLUSION

The main objective of the work was to make a model of a meter that measures the electricity consumption of a common appliance. From the analysis of the measurement problem, it is clear that it is not trivial to measure AC electrical power and requires sampling and calculation by a microcontroller. However, even with these methods, we only get to accuracies of 5 % at low currents. We can evaluate the developed user interface as a success and future extensions with new features are possible.

REFERENCES

[1] Allegro MicroSystems, Inc. Datasheet for ACS712. *Fully Integrated, Hall Effect-Based Linear Current Sensor with 2.1 kVRMS Voltage Isolation and a Low-Resistance Current Conductor* [online]. 2006 - 2007, Allegro MicroSystems: 115 Northeast Cutoff Worcester, Massachusetts U.S.A . Available from: <https://lastminuteengineers.com/rotary-encoder-arduino-tutorial/>.

[2] HUDSON, Glyn. *AC Power Theory - Arduino Maths* [online]. last actualization 7. 2. 2021 [cit. 20. 11. 2021]. Available from: <https://learn.openenergymonitor.org/electricity-monitoring/ac-power-theory/arduino-maths>.

[3] Lekić M. , Gardašević G. *IoT sensor integration to Node-RED platform* [online]. 2018,International Symposium INFOTEH-JAHORINA (INFOTEH) [cit. 9. 4. 2022]. Available from: <https://ieeexplore.ieee.org/abstract/document/8345544>. ISBN 978-1-5386-4907-7.

[4] J. Wang and T. -T. Liu. *The textbook compilation for Altium Designer should highlight its practical characteristics*, 2018 Chinese Control And Decision Conference (CCDC), Shenyang, China. CCDC.2018.8407721.

[5] BEJČEK Ludvík, ČEJKA Miloslav, REZ Jiří, GESCHEIDTOVÁ Eva a STEIBAUER Miloslav. *Skripta k předmětu. Měření v elektrotechnice (BMVE)*, 2. vyd. KOPP: České Budejovice, 2007. ISBN 978-80-7232-324-1.

Video Summarization Using Deep Learning Approach

Libor Matejek
Brno University of Technology
Brno, Czech Republic
xmatej53@vutbr.cz

Tomas Fryza
Brno University of Technology
Brno, Czech Republic

***Abstract*—This article is about a video summarization algorithm. The basic building block for image processing is a convolutional neural network. Subsequently, the output feature vector is processed using the k-means method and thresholding in parallel. The output is a video reduced by redundant data.**

***Index Terms*—Video summarization, CNN, K-Means, Python**

I. Introduction

Today's world is full of data. Multimedia is a large separate area, we will be interested in videos. It is possible to come across billions of gigabytes of video sequences on the Internet but some of these sequences are completely redundant. A clear type of these video sequences are, for example, redundant non-informative parts of recordings from security video cameras. It is appropriate to deal with the area of video summarization and develop it.

Video summarization is the process by which we are able to generate a shorter version of a long one and the informational value is sufficient. In general, there are many types of algorithm for summarizing different kinds of videos. One of them is the creation of movie trailers, another is the search for pathology in endoscopic video recordings. This article discusses an algorithm detecting general information change of any type and reducing redundant data.

The article presents a machine learning based summarization algorithm.

The rest of this document is organized as follows. Section II presents the basic concept of the algorithm. In Sections III, the used type of convolutional neural network is described. Subsequently, Section IV describes a more detailed description of the actual implementation of the summarization algorithm. Finally, Section VI gives the conclusions and future work.

II. Aplication

The concept of the application is based on convolutional neural networks and on the features vectors it generates. This type of networks is an essential part of any deep learning algorithm that works with the image. Feature vectors represent the extraction of information, on which other algorithms can easily be built. Feature vectors are further processed by the block of K Means clustering and the block based on the Euclidean distances of individual images.

III. Used CNN - VGG16

A convolutional neural network (CNN) is a class of neural networks that, by their nature, specialize in image data processing. The network works with the image as a vector of pixel values that determine color and brightness. CNN processes the image similarly to the human brain. Each neuron operates only in a limited receptive field. Layers detect first simpler patterns, then more complex objects. A convolutional neural network is the means by which a computer can "see".

For our algorithm there is used VGG16 CNN. Its a CNN from Visual Geometry Group from Oxford, 16 means a number of layers in network. We can see the structure of the network at Figure 1 [1].

A lower number of layers has an advantage in the form of speed, at the same time accuracy sufficient for a video summarization application. The input format to the network is an image with dimensions 224x224x3. The first two layers architectures are convolutional, i.e. feature map extraction networks. Specifically, 3x3 filters with 64 channels are applied. At the same time, padding is performed to preserve co possibly the most information. Then, after dimension reduction by max-pooling, they are applied more convolutional 3x3 filters, this time 128. Then another max-pooling, 256 3x3 filters. Further the number of convolutional layers is further increased, followed by 3 sets of 512 filters, then again max- pooling. The feature map extraction is completed by the last 3x3 filters. The output is 7x7x512 18 feature vector. However, such a size is unsuitable for further processing. That's why we use the flatten function, i.e. we transform the dimension to 1x25088. This dimension further reduces fully connected layer. These are generally found in the VGG16 architecture 3, they first reduce to 1x4096, the second layer does not change the size and the last layer represents the reduction to the required number of classes into which we classify. All hidden layers contain the ReLU activation function, the last one is the implementation of the softmax function. The main advantage of the ReLU activation function is the ability to avoid the phenomenon called vanishing gradient. This phenomenon can be caused by mutual negation of the negative and positive outputs of the neuron. At the same time, ReLU shows faster learning. More about VGG16 can be found on [2].

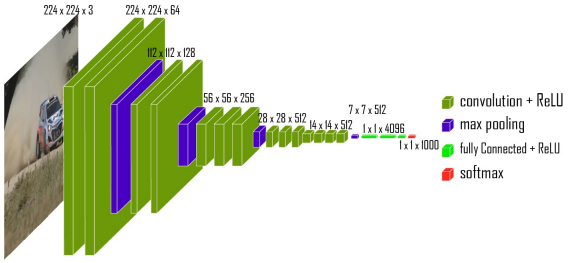


Fig. 1. VGG16 Ilustration.

IV. Implementation

The basic tools for implementing the program are the Python programming language, which is the most widely used and user-friendly in the field of machine learning. The powerful framework is Google's Tensorflow with Keras API. Tensorflow with Keras API is one of the most intuitive frameworks in the field of machine learning. At the same time, its performance is comparable to Facebook's PyTorch. Implementation is inspired by [3].

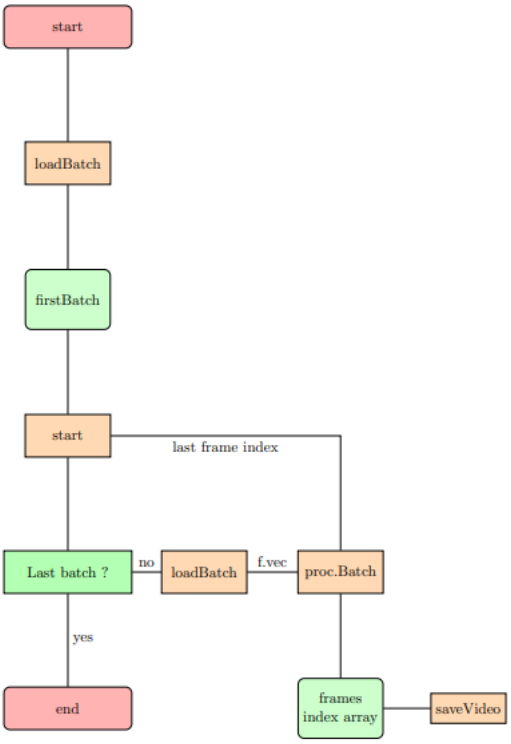


Fig. 2. function `sumarize`

A. *function* `sumarize`

First, the VGG16 model is initialized. This model however, it is fundamentally reconfigured. In the next step, the last two are "cut off" classification layers. The result is a significant change in the output of the convolutional neural network. It now has 4096 dimensions, which is an order of magnitude more than in the default setting. This step is a necessary measure to increase the resolution of feature vectors and the consequent better sensitivity to information change in different frames. Next, it is iterated after video batches and the algorithm for image reduction is executed. The flowchart on Figure 2 describes the function.

B. *function* `loadBatch`

The loadBatch function is a means of dividing a video into segments by a certain amount length, at the same time extracts frame rate information from the original video. In order to preserve the power of other algorithms, i.e. to start from sufficiently long ones segments, we introduce a relative minimum segment size condition here. If the set segment length cannot be met, an incomplete segment is added to the previous one. The last segment is therefore usually longer. The reason why the video is processed in batches is the size of the feature vectors and the large allocation of RAM in the case of the entire video. The default batch size is 1000 frames. However, in the developing GUI, this parameter will be optional.

If we consider a video 1 minute long, 24 FPS and a feature vector of size 4096 * float:

$$FPS{\cdot}t{\cdot}vector_{size}{\cdot}float_{size} = 24{\cdot}60{\cdot}4096{\cdot}4B = 23,59MB$$

It is certainly worth noting the theoretical amount of memory that should be needed to process such a video. The result is a relatively low value. In practice, however, the allocated memory of the entire process varies by orders of magnitude. Specifically, Python together with the libraries allocates about 900 MB during initialization and during the processing of a 30s video, it climbs tens to hundreds of MB higher. The complexity of the libraries used and the existence of overhead data are to blame.

Unfortunately, the disadvantage of batching is the slight deterioration of the K-Means algorithm, which is described below. It needs the entire matrix of vectors for its evaluation. The meaning of the labels assigned by the K-Means algorithm is therefore only relevant in the area of each batch separately.

C. *function* `processBatch`

The processBatch function is a platform for working with the output array from loadBatch function. Here, the entire vector is first reduced by the PCA algorithm. We are now working from the original 4096-dimensional feacture vector for each frame 100 dimensional. Next, the frames input field is transformed into a Euclidean field distance between images. At this moment, the solution to the problem is bifurcated into two separate directions. One of them is the application of K-Means clustering, which from 100 dimensional reduced feature vectors of the entire batch. We can see the flowchart on Figure 3.

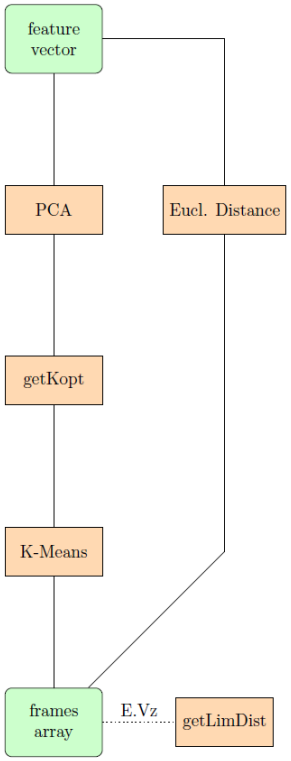


Fig. 3. function `processBatch`

1) function `getKopt`*:* In the field of unsupervised learning algorithms, K-means is the most widely used. The task of this method is to divide the data space so that the points in individual clusters are as similar as possible. At the same time, points from other clusters must be as different as possible.

In the K-means method, each cluster is represented by its center, which is marked as the centroid. The centroid corresponds to the arithmetic mean of the data points that are assigned to it. This is a point that is not necessarily a member of the cluster itself. During the method, the most "energy-saving" centroid positions are searched for by an iterative method.

A necessary parameter for K-Means clustering is the constant K. This is a value determining the number of resulting clusters. The correct resolution of the algorithm depends on the correct determination of the value of K. This task is performed by the function `getKopt`, see Figure 4.

The basic idea of the function is the initial plotting of Sum of squared distance (sum of distances centroid-data points) values for various constants. The plot of the dependence is further interpolated with a mathematical function:

$$y(x) = A * x^B \tag{1}$$

The next step is to find the tangents using the derivative at the first and last point interleaved functions. The tangent lines are calculated using Python operations. Optimal the value of K

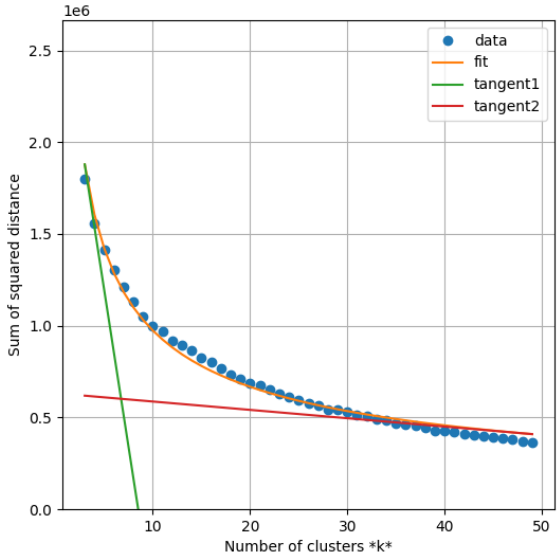


Fig. 4. function `getKopt`

is the rounded x coordinate of the intersection of the tangents determined above.

2) function `getLimDist`*:* The second way is to set the detection threshold of changes in the information on the images. This the method is based on the Euclidean distances of consecutive images, the vector of which we have prepared. A large change in information is easily detectable by a large one by increasing this distance, however, it is necessary to determine the correct threshold value. The information change is visible in the Figure 5.

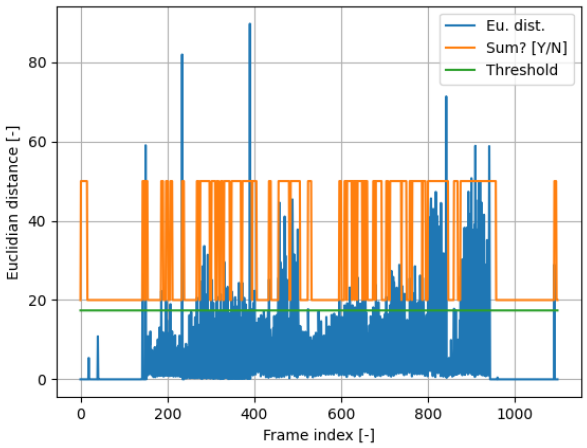


Fig. 5. function `getLimDist`

First, extremes that would potentially distort the mean and standard deviation need to be trimmed. This is achieved by calculating the initial average and the standard deviation.

Subsequently, all values smaller than 0.1 are deleted. Then we filter out values greater than the mean value + 2 standard deviations. This threshold calculation algorithm is found experimentally. It guarantees finding a threshold not degraded by extreme values.

Now it is possible to calculate the ”usable” mean and standard deviation. Threshold is then equal to the new mean + standard deviation.

D. function `filterFrames`

At the stage when we are able to locate the indexes of images where significant changes occur, it is necessary to define an algorithm for selecting indexes for the resulting summarization. In order to select the length of informatively advantageous segments, it is first necessary to count the number of frames in the area between two peaks of the Euclidean distances, i.e. peaks that activate the threshold. At the same time, a change in the cluster determined by the K-Means clustering method is also a trigger for evaluating the change in the informative area. This gives us the number and indices of frames with the same informative value. The image filtering method always stores at least 5 images, or 10 percent of the images from the informatively same area.

At the moment when we have marked the images that we want in the final summarization, we generate the video.

In addition to the video itself, the output is also a csv file with information about vector changes and the assignment of clusters to individual frames. However, the meaning of clusters always refers to only one image batch. In other words, cluster assignment internally in a batch always collects informatively similar images. The images at the beginning of the batch can therefore be similar to those at the end of the batch and are assigned a value of e.g. 7. However, if the algorithm ”starts” another batch, these values are reset and the cluster value 7 can indicate a completely different informative area.

V. Aplication testing

The implemented algorithm is tested on a video freely available on the web [4]. This is a video with footage of animals in motion. The video is preceded by a static intro and ends with a static outro. It is a slow-motion video, so there is room for a large data reduction. In the figure 6 we see the basic types of shots in the video.

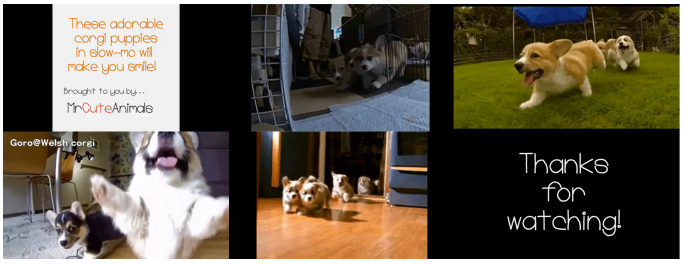


Fig. 6. Frame types in tested video

Before applying the algorithm, the video contains 1099 frames. The algorithm is then able to remove more than half of the redundant data from the video, resulting in 509/1099 frames. The resulting video is informative. It contains all the more distinctly unique images. The disadvantage is the loss of smooth transitions between frames, but this is an inevitable feature of reduced video sequences. In addition to the reduced video sequence, the output of the program is out.csv, where, in addition to the frame index, there is also information about the Euclidean distance from the previous frame, the number of the cluster in the batch to which it belongs, and True/False information about whether the frame is included in the resulting video.

The image processing speed was in the range of 100-150ms/image during calculations on an Intel Core i5-1135G7 CPU. When calculating on the NVIDIA GeForce 1050Ti GPU, the speed increased, frames were processed at a speed of 20-35ms/frame.

VI. Conclusion

The application, programmed in Python on the Keras/Tensorflow platform, relies on the VGG16 convolutional neural network, which is shortened by 2 classification layers. This shortening makes the convolutional network a universal tool for image processing. CNN now generates a vector of 4096 values. This is further reduced to 100 dimensions using the PCA method. The Euclidean distance is found between the vectors of the individual images and at the same time K-Means clustering is performed on them. The data we have at that moment is sufficient to evaluate the informative value of the images and the subsequent reduction of the video.

In order to be able to process longer videos and not fill up the RAM memory, the algorithm works with data batches, in other words, the video is processed in parts.

The work describes, of course, how the algorithms are set to determine the threshold when the image is informatively interesting and when it is not.

The algorithm was finally tested on a short video with moving animals. The original 1099 frame video was reduced to a 509 frame video. By subjective evaluation, I can say that the information was largely preserved.

In the future, the algorithm will be extended with a graphical interface and programmed using the pyQT library. The GUI will allow the user to modify the proposed threshold based on need. At the same time, the degree of image reduction will be shown graphically.

References

[1] Mayank Mishra *Convolutional Neural Networks, Explained* [online]. [cit. 19. 10. 2022]. Dostupné z URL: <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939>

[2] pawangfg *VGG-16 — CNN model* [online]. [cit. 19. 10. 2022]. Dostupné z URL: <https://www.geeksforgeeks.org/vgg-16-cnn-model/>

[3] Gabe Flomo *How to cluster images based on visual similarity* [online]. [cit. 19. 10. 2022]. Dostupné z URL: <https://towardsdatascience.com/how-to-cluster-images-based-on-visual-similarity-cd6e7209fe34>

[4] MrCuteAnimals *In 30 sec, These Hilarious Slow-Mo Corgi Puppies Will Make You Laugh & Smile!* [online]. [cit. 19. 10. 2022]. Available: <https://www.youtube.com/watch?v=F48_PVjQp4M&ab_channel=MrCuteAnimals>

Comparison of compression and decompression speeds for 360° images with resolutions up to 8K using conventional and deep learning algorithms

Adam Budac, Jan Kufa
Brno University of Technology
Faculty of Electrical Engineering and Communications, Dept. of Radio Electronics
Technicka 12, 616 00 Brno, Czech Republic
{xbudac02, kufa}@vut.cz

***Abstract*—Recently, due to the accessibility of virtual reality, there has been a great interest in the creation and distribution of omnidirectional (360° field of view) image and video content. To preserve high resolution and high visual quality of this content during storage and distribution to the final user, the use of advanced compression methods is essential. It is not only the quality of compression and bitrate savings that are important with advanced compression methods, but also the speed of compression and decompression. It is often necessary to compress and decompress images in real time, which requires sufficient computing power or specialized hardware. The demand for computing power is constantly increasing with new advanced compression algorithms. These compression algorithms are not only conventional image compression codecs, but thanks to progress in the field of artificial intelligence, it is also possible to compress the image using the deep learning methods. These methods are based on the JPEG AI framework. Four JPEG AI-based image compression algorithms (FRICwRNN, CAEMfE2EOIC, HiFiC, and LIC360) were compared against conventional image compression codecs (JPEG, JPEG XL, HEIC, AVIF, and VVC Intra) in terms of compression and decompression speeds and hardware resources required for compression and decompression. For this purpose, ten images from publicly available databases were compressed and decompressed.**

Keywords—360° omnidirectional images, compression and decompression speed, deep learning, image compression codecs, JPEG, JPEG XL, HEIC, AVIF, VVC Intra, JPEG AI

I. INTRODUCTION

Recently, a lot of images are produced and shared every day [1]. In mobile phones, personal computers, or in the cloud and data centers utilized by social networks and streaming services, images and videos consume a significant amount of storage [2], [3]. The bandwidth required to send this material to end users through the internet is both expensive and limited, in addition to storage. Advanced compression methods are necessary to reduce file sizes with the least amount of visual quality loss given the rising demand for high-quality material. Due to the cut out in the image, omnidirectional content requires greater resolutions than traditional planar content. This cut out must have a sufficient resolution. Higher resolutions make it crucial to minimize the data size without significantly sacrificing quality in terms of compression artifacts. Modern compression codecs are employed for this purpose. There is significant interest in machine learning (ML), neural networks (NN), and deep learning (DL) approaches that may be employed for compression [4] because of the increased computational capabilities of contemporary digital devices. These approaches are used for image compression by the JPEG AI framework [5], and several image compression solutions based on them are presented. Newer compression algorithms take many times longer to compress and decompress the image. Depending on the quantization parameter/quality of the compression, the time required for compression and decompression may vary considerably. It is often necessary to compress and decompress a high-resolution image in real time, which requires the use of a very powerful central processing unit (CPU) or specialized hardware in the form of hardware encoder and decoder, graphics processing unit (GPU) and tensor processing unit (TPU) used for ML and DL hardware acceleration.

The Joint Photographic Experts Group (JPEG) released JPEG [6], the first widely used image compression codec, in 1992. JPEG 2000, JPEG LS, JPEG XT, and JPEG XR are more recent and more efficient image compression codecs that have been standardized, but they have not gained as much interest and usage as JPEG. JPEG XL [7], the most recent image codec, is continuously being developed. On the other hand, intra-compression of video codecs like High Efficiency Video Coding (HEVC) [8], AOMedia Video 1 (AV1) [9], or Versatile Video Coding (VVC) [10] can be employed for still-image compression. As a result, image compression codecs including High Efficiency Image Codec (HEIC), AV1 Image File Format (AVIF), and VVC Intra were developed. Since 2018, more JPEG AI-based solutions have been put out, and more are continually being developed, increasing compression efficiency. Four DL-based algorithms that implement JPEG AI methods were chosen for the experiment. These algorithms are Full Resolution Image Compression with Recurrent Neural Networks (FRICwRNN) [11], Context-adaptive Entropy Model for End-to-End Optimized Image Compression (CAEMfE2EOIC) [12], High-Fidelity Generative Image Compression (HiFiC) [13], and End-to-End Optimized 360-Degree Image Compression (LIC360) [14].

Research described in this paper was supported by the Internal Grant Agency of the Brno University of Technology under project no. FEKT-S-23-8191.

Many studies examine the compression efficiency of the most recent conventional codecs [15] or the compression effectiveness of state-of-the-art codecs for omnidirectional images [2], [3], and [16]. Several papers investigate the efficiency of DL for image compression [4], [5], and [11]–[14], or the hardware and system resources needed to compress high-resolution images [17]. However, they do not compare the encoding and decoding speeds of conventional codecs and DL-based algorithms, even considering the demand for computing power and hardware resources.

The main contribution of this work is to provide a comparison of compression and decompression speeds between trained and optimized DL-based image compression algorithms and the most recent conventional codecs set to their near-maximum complexity, resulting in the most efficient compression. The comparison also shows the maximum image resolution that can be compressed and decompressed by individual codecs on hardware with limited processing power and memory. Ten omnidirectional images with 8192 × 4096 (8K) resolution in the equirectangular projection (ERP) were selected for this experiment.

The image compression algorithms used are explained in Section II. The experiment is described in Section III, where the dataset of images used is mentioned in Section III/A, the hardware used is mentioned in Section III/B, and compression and decompression are described in Section III/C. The results are presented in Section IV, and the conclusion with future work are presented in Section V.

II. Image Compression Algorithms

Chroma planes in JPEG are subsampled using a 4:2:0 ratio. Next, 8 × 8 transform blocks are created and subjected to a 2-dimensional (2D) discrete cosine transform (DCT). Frequency coefficients are quantized using different quantization tables for different quantization settings. Huffman and run-length encoding (RLE) are used for entropy encoding of the quantized coefficients.

JPEG XL achieves higher efficiency with fewer compression artifacts by using adaptive quantization and variable block size DCT (VarDCT) for transform coding. It also uses more effective entropy encoding, chroma from luma prediction, and image feature extraction.

The image is divided into 64 × 64-pixel large coding tree units (CTU) for HEIC, which are then downscaled as required for efficient transform coding using DCT, discrete sine transform (DST), or no transform in the case of artificial images such as computer graphics. Deblocking filters and sample adaptive offset (SAO) are used to maintain the highest quality. Entropy encoding is context adaptive binary arithmetic coding (CABAC).

For transform and prediction, AVIF divides the image into 128 × 128-pixel large superblocks, which are then downscaled. DCT, asymmetric DST (ADST), flip ADST or no transform are used for transform coding. Compared to a pure diagonal readout, the coefficient readout has a quarter-circle shape, which is more effective. Deblocking filters and chroma from luma prediction are also used. Multi-symbol binary arithmetic coding is used for entropy encoding.

Large CTUs of 128 × 128 pixels, chroma from luma prediction, more accurate spatial prediction, DCT and DST transforms, as well as no transform, adaptive quantization, deblocking filters, CABAC entropy encoding, and 360° edge detection are all implemented by VVC Intra.

The gated recurrent unit (GRU) neural network architecture of the FRICwRNN combines a recurrent autoencoder (AE) with convolutional layers. The input image is divided into 32 × 32-pixel large blocks, which are processed independently and encoded using entropy encoding to create a latent code.

The CAEMfE2EOIC algorithm uses adaptive entropy encoding over 256 × 256-pixel large blocks after learning the AE with a convolutional neural network (CNN). Mean squared error (MSE) and multiscale structural similarity (MS-SSIM), two objective metrics, are optimized for two trained models.

The MSE measurement is optimized in the HiFiC algorithm, which is based on a generative adversarial network (GAN) with three trained models for three different bit rates. It divides the input image into 256 × 256-pixel large blocks.

LIC360 uses a deep neural network (DNN) architecture specifically designed for 360° images in ERP. The input images are cut outs with a resolution of 1024 × 512 (1K) pixels, sampled from a viewport in an omnidirectional projection sphere such as a head-mounted display (HMD). The models are tuned for viewport MSE (VMSE) and viewport structural similarity (VSSIM) metrics. The image is divided into 256 × 256-pixel large blocks.

III. Experiment

A. Dataset

Ten omnidirectional images in ERP are compressed using the DL-based FRICwRNN, CAEMfE2EOIC, HiFiC and LIC360 algorithms in addition to the conventional codecs such as JPEG, JPEG XL, HEIC, AVIF and VVC Intra. Publicly available databases[1, 2] provide source images in Tagged Image File Format (TIFF, lossless) and JPEG (near-lossless) formats in ERP. The images used are shown in Fig. 1. These images reflect different characteristics with respect to different spatial indices (SI) and regions of interest (ROI). The ERP has an aspect ratio of 2:1, and all images are in 8K resolution. Not all images had the exact resolution of 8192 × 4096 pixels. Some of these images had slightly smaller or much higher resolution, therefore they have been scaled to 8192 × 4096 resolution, as DL-based algorithms require input image dimensions to be multiples of 32 or 256, respectively. All images have a bit depth of 24 bits per pixel (bpp). Various codecs require input images and produce decoded output images in several formats. These file types include YUV, Windows Bitmap (BMP) and Portable Network Graphics (PNG). The FFmpeg[3] 5.1.2 x64 multimedia framework handled the processing and transcoding between the different formats.

[1] https://www.kandaovr.com/gallery/

[2] https://www.eso.org/public/images/archive/category/360pano/

[3] https://ffmpeg.org/



Fig. 1. 360° images used for compression and decompression speed evaluation.

B. Equipment

The ASUS ROG Strix SCAR III laptop with an Intel Core i7 9750H CPU, Nvidia GeForce RTX 2070 8 GB GPU, 32 GB of DDR4 RAM memory and a 2 TB SSD storage with a speed of 500 MB/s is the hardware used for this experiment.

C. Compression and decompression

Libjpeg-turbo[4] 2.1.4 x64 is used for JPG. Libjxl[5] 0.7.0 x64 is the version of JPEG XL used. HM[6] 17.0 x64 reference software is used for HEIC, while libavif[7] 0.11.1 x64 is used for AVIF compression and decompression. VVenc[8] 1.7.0 x64 is used for VVC Intra. The latest implementations of FRICwRNN[9] (24 Mar 2018), CAEMfE2EOIC[10] v2, HiFiC[11] v0.3, and LIC360[12] (25 Oct 2022) from GitHub, as well as the corresponding TensorFlow and PyTorch versions, were used for DL-based solutions. All conventional codecs were set to near-maximum complexity for the most effective compression, but where multithreading was available, it was used to accelerate compression. JPEG used Huffman tables optimization. JPEG XL optimizations are effort = 9 and Brotli effort = 11. HEIC optimization is 4:2:0 chroma subsampling and full range of output values. For AVIF, the same optimizations were used as for HEIC, with the addition of speed = 10, delta quantization in chroma channels and SSIM tuning. For VVC Intra, the same optimizations were used as for HEIC, with the addition of slower preset. Multithreading was used for JPEG XL, HEIC, AVIF and VVC Intra. For the DL algorithms CAEMfE2EOIC and LIC360, SSIM-based optimizations were preferred over MSE because the SSIM-based metrics consider the perception by human visual system (HVS).

[4] https://libjpeg-turbo.org/

[5] https://github.com/libjxl/libjxl

[6] https://github.com/AOMediaCodec/libavif

[7] https://vcgit.hhi.fraunhofer.de/jvet/HM

[8] https://github.com/fraunhoferhhi/vvenc

[9] https://github.com/tensorflow/models/tree/archive/research/compression/image_encoder

[10] https://github.com/JooyoungLeeETRI/CA_Entropy_Model

[11] https://github.com/Justin-Tan/high-fidelity-generative-compression

[12] https://github.com/limuhit/360-Image-Compression

IV. Results

All images were compressed by all codecs to all available qualities/quantization parameters. This produced a total of 4110 compressed images. Codecs such as JPEG, JPEG XL, HEIC, AVIF, and VVC Intra could compress images up to 8K resolution using only the CPU. The FRICwRNN and HiFiC algorithms could compress and decompress images up to 2048 × 1024 (2K) resolution due to the high memory requirements, and the compression and decompression were accelerated by the GPU. The CAEMfE2EOIC algorithm could compress and decompress images up to 4096 × 2048 (4K) resolution, but this algorithm did not use the GPU for hardware acceleration. The LIC360 algorithm compressed images at 1K resolution using GPU hardware acceleration. The individual compression times for minimum, maximum, and the average of all available qualities are shown in Table 1. It is the average of all images. The hardware resources used by each codec are shown in Table 2. In terms of compression speed, the fastest codec was LIC360 because it compresses images at 1K resolution. The CPU load was 25%, the GPU load was 100%, the GPU memory utilization was 3.5 gigabytes (GB), and the RAM utilization was 2.4 GB. The second fastest is JPEG. The CPU load was 17%, and the RAM utilization was 0.1 GB. Next is the HiFiC codec. The CPU load was 25%, the GPU load was 100%, the GPU memory utilization was 7.3 gigabytes (GB), and the RAM utilization was 2.8 GB. The fourth is the FRICwRNN codec, which has similar memory requirements to HiFiC. Only the CPU load was reduced by 10%. The next one is AVIF. The CPU load was 100%, and the RAM utilization was 0.7 GB. The sixth is the HEIC. The CPU load was 16%, and the RAM utilization was 0.1 GB. The next one is the JPEG XL codec. The CPU load was 25%, and the RAM utilization was 8.8 GB. The second slowest is the CAEMfE2EOIC codec. The CPU load was 16%, and the RAM utilization was 6.1 GB. Compression by the VVC Intra codec took the longest. The CPU load was 100%, and the RAM utilization was 0.8 GB. For the DL algorithms, the speed of decompression was similar to the speed of compression due to the architecture of neural networks. The same applies to the use of hardware resources. However, decompression by conventional codecs was much faster than compression, in the order of seconds. The fastest codec was JPEG, which decompressed images in less than a second. The use of hardware resources during decompression was also significantly reduced.

TABLE I. COMPRESSION TIMES

Quality	Time [s]		
	Minimum	*Maximum*	*Average*
JPEG (8K)	5	10	8.3
JPEG XL (8K)	120	2100	218
HEIC (8K)	100	275	175
AVIF (8K)	50	100	96
VV Intra (8K)	10	2580	1069
FRICwRNN (2K)	65	80	72
CAEMfE2EOIC (4K)	190	550	280
HiFiC (2K)	65	65	65
LIC360 (1K)	3.5	3.5	3.5

TABLE II. HARDWARE RESOURCES USED DURING COMPRESSION

Codec	Hardware resources			
	CPU [%]	*GPU [%]*	*GPU memory [GB]*	*RAM [GB]*
JPEG (8K)	17	0	0	0.1
JPEG XL (8K)	25	0	0	8.8
HEIC (8K)	16	0	0	0.1
AVIF (8K)	100	0	0	0.7
VV Intra (8K)	100	0	0	0.8
FRICwRNN (2K)	15	100	7.2	2.5
CAEMfE2EOIC (4K)	16	0	0	6.1
HiFiC (2K)	25	100	7.3	2.8
LIC360 (1K)	25	100	3.5	2.4

V. CONCLUSION

Compression with conventional codecs was performed on the CPU to achieve the highest compression efficiency. The codecs in order of compression speed, from fastest to slowest according to the average encoding time, are LIC360, JPEG, HiFiC, FRICwRNN, AVIF, HEIC, JPEG XL, CAEMfE2EOIC, and VVC Intra. Decompression with conventional codecs was almost instantaneous, while with deep learning codecs it took as long as compression. The maximum image resolution for compression and decompression with deep learning codecs was 2K and 4K, respectively. Compression at higher resolutions used up all available RAM, GPU memory, and the system page file and caused a crash of the compression framework. With sufficient memory, it would be possible to compress higher-resolution images using deep learning algorithms. Also, not all codecs were able to use 100% of the CPU. Codecs that also used hardware acceleration were limited by the speed of the GPU. If it were possible, their compression speed would be even higher. Codecs that also used hardware acceleration were limited by the speed of the calculations on the GPU. In future work, the impact of compression on the visual quality of omnidirectional images will be explored through objective and subjective quality evaluation, considering aspects such as high dynamic range (HDR) and wide color gamut (WCG).

REFERENCES

[1] Cloudinary, “2022 State of Visual Media,” 2022.

[2] M. Simka, J. Kufa and L. Polak, "Picture Quality of 360° Images Compressed by Emerging Compression Algorithms," 2022 32nd International Conference Radioelektronika (RADIOELEKTRONIKA), Kosice, Slovakia, 2022, pp. 1-4, doi: 10.1109/RADIOELEKTRONIKA54537.2022.9764941.

[3] M. Xu, C. Li, S. Zhang and P. L. Callet, "State-of-the-Art in 360° Video/Image Processing: Perception, Assessment and Compression," in IEEE Journal of Selected Topics in Signal Processing, vol. 14, no. 1, pp. 5-26, Jan. 2020, doi: 10.1109/JSTSP.2020.2966864.

[4] F. Hussain and J. Jeong, "Exploiting deep neural networks for digital image compression," 2015 2nd World Symposium on Web Applications and Networking (WSWAN), Sousse, Tunisia, 2015, pp. 1-6, doi: 10.1109/WSWAN.2015.7210294.

[5] J. Ascenso, P. Akyazi, F. Pereira, T. Ebrahimi, "Learning-based image coding: early solutions reviewing and subjective quality evaluation," Optics, Photonics and Digital Technologies for Imaging Applications VI, SPIE Photonics Europe, 2020, Online Only, April 2020, doi: 10.1117/12.2555368.

[6] "Joint Photographic Experts Group," World Wide Web Consortium, Cambridge, MA, Sep. 1992, https://www.w3.org/Graphics/JPEG/itu-t81.pdf.

[7] "JPEG XL," JPEG XL, Geneva, CH, ITU, 2022, https://docs.google.com/presentation/d/1LlmUR0Uoh4dgT3DjanLjhlXrk_5W2nJBDqDAMbhe8v8/edit#slide=id.g821b9c186b_0_34.

[8] "High Efficiency Video Coding," ITU, Geneva, CH, Nov. 2019, https://www.itu.int/rec/T-REC-H.265-202108-I/en.

[9] "AOMedia Video 1," Alliance for Open Media: Next Generation, Open-Source Digital Media Technology for Everyone, Wakefield, MA, 2019, http://aomedia.org/av1/specification/.

[10] "Versatile video coding," ITU, Geneva, CH, Apr. 2022, https://www.itu.int/rec/T-REC-H.266-202204-I/en.

[11] G. Toderici et al., "Full Resolution Image Compression with Recurrent Neural Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 5435-5443, doi: 10.1109/CVPR.2017.577.

[12] J. Lee, S. Cho, S. Beack, "Context-adaptive Entropy Model for End-to-end Optimized Image Compression," Image and Video Processing (eess.IV): ICLR 2019, Montreal, CA, May. 2019, doi: 10.48550/arXiv.1809.10452.

[13] F. Mentzer, G. Toderici, M Tschannen, E. Agustsson, "High-Fidelity Generative Image Compression," NeurIPS 2020: Image and Video Processing (eess.IV), Montreal, CA, Oct. 2020, doi: 10.48550/arXiv.2006.09965.

[14] N. Mahmoudian Bidgoli, R. G. de A. Azevedo, T. Maugey, A. Roumy and P. Frossard, "OSLO: On-the-Sphere Learning for Omnidirectional Images and Its Application to 360-Degree Image Compression," in IEEE Transactions on Image Processing, vol. 31, pp. 5813-5827, 2022, doi: 10.1109/TIP.2022.3202357.

[15] S. Naveen Kumar, M. V. Vamshi Bharadwaj and S. Subbarayappa, "Performance Comparison of Jpeg, Jpeg XT, Jpeg LS, Jpeg 2000, Jpeg XR, HEVC, EVC and VVC for Images," 2021 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2021, pp. 1-8, doi: 10.1109/I2CT51068.2021.9418160.

[16] Z. Chen, Y. Li, Y. Zhang, "Recent advances in omnidirectional video coding for virtual reality: Projection and evaluation," Signal Processing, May. 2018, vol. 146, pp. 66-78, doi: 10.1016/j.sigpro.2018.01.004.

[17] M. T. Pourazad et al., "Comparison of Emerging Video Compression Schemes for Efficient Transmission of 4K and 8K HDR Video," 2021 IEEE International Mediterranean Conference on Communications and Networking (MeditCom), Athens, Greece, 2021, pp. 209-214, doi: 10.1109/MeditCom49071.2021.9647504.

Exploration of potential inhibitors of CrtM enzyme of Staphylococcus aureus using squalene synthase inhibitors: An in-silico study

Inderjeet Bhogal
Department of Biomedical Engineering,
Faculty of Electrical Engineering and Communication
Brno University of Technology, Brno, 616 00, Czech Republic
234190@vut.cz

Vaishali Pankaj
Department of Biomedical Engineering,
Faculty of Electrical Engineering and Communication
Brno University of Technology, Brno, 616 00, Czech Republic
234084@vut.cz

Sudeep Roy
Department of Biomedical Engineering,
Faculty of Electrical Engineering and Communication
Brno University of Technology, Brno, 616 00, Czech Republic
roy@vut.cz

***Abstract*—In present study, compounds with cholesterol lowering activity viz., squalestatin analogs were screened for inhibitory effect towards CrtM enzyme of *S. aureus*. We used structure-based drug design approach including molecular docking studies followed by molecular dynamics simulations and binding free energy calculations for the identification of cholesterol lowering agents against CrtM. The investigation reveals squalestatin analog 25 and squalestatin analog 36 as potential candidates. The in-depth analysis revealed important amino acids critical for binding into the active site of CrtM. The compounds under investigation were further screened down based on their *in-silico* ADMET analysis. This discovery serves as a proof-of-concept against *S. aureus* virulence factor-based treatment.**

Keywords: CrtM, docking, MM-GBSA, molecular dynamics, staphyloxanthin, virtual screening

I. Introduction

Staphylococcus aureus is one of the main pathogens causing a variety of hospitals- and community-acquired infections that can lead to various diseases like pneumonia, meningitis and toxic shock syndrome [1]. Dehydrosqualene synthase (CrtM) is involved in the synthesis of staphyloxanthin which is the most recognized membrane-associated carotenoid pigment and contributes to golden pigment that provides anti-oxidant property to the *S. aureus* and enables its survival within the host cell [2]. It has been reported that there is 30% sequence identity between the human squalene synthase (SQS) and dehydrosqualene synthase (CrtM) *S. aureus*, and they share substantial structural features [3]. The structural similarity between human SQS and CrtM suggests that human SQS inhibitors available as cholesterol-lowering drugs might also be active against CrtM. Drugs that prevent the formation of staphyloxanthin can avoid bacterial infections from starting and continuing without favoring antibiotic resistance. Hence, there was a possibility to explore the the biosynthetic pathway of staphyloxanthin by various SQS inhibitors. Accordingly, an *in-silico* approach was used to study the molecular interactions between the statins selected into the binding site of *S. aureus dehydrosqualene synthase*.

In the present study, we used combined molecular modeling approach on human SQS inhibitors to investigate the mechanism of inhibitor binding with *S. aureus* CrtM enzyme. The structural interactions between the protein (PDB ID: 3W7F) and five inhibitors (squalestatin analogs) were docked separately by using Induced Fit Docking (IFD) for the prediction of ligand binding mode and estimation of docking score. To investigate further, the functionalities of active molecular interaction between the inhibitor and the catalytic pocket residues, an explicitly solvated 100 ns molecular dynamics (MD) simulation was performed. The binding free energy of the protein-ligand complex was calculated by the Molecular Mechanics/Generalized Born Surface Area (MM-GBSA). However, to understand the role of individual amino acid involvement in the mechanism of binding; residue energy decomposition analysis was performed. Furthermore, the test compounds were screened on the basis of ADMET properties by ADMETlab 2.0 and pre-ADMET. The current work is an extension of the research conducted by the same group Kahlon *et al.*, 2010 [4]. We carried out in-depth analysis and came up with additional findings.

II. Material and Methods

A. Protein and ligand preparation

The crystal structures of the C (30) carotenoid dehydrosqualene synthase from *S. aureus* (PDB Id: 3W7F) was retrieved from RCSB protein databank [5]. The protein structure was prepared using *Protein Preparation Wizard* in (Schrodinger's Maestro 2022-4). During protein preparation, heteroatoms, and crystal waters beyond 5Å from the binding sites were removed. Polar hydrogens were added and partial charges were assigned using Optimized Potentials for Liquid Simulations (OPLS4) force-field. Similarly, 3D conformers of Squalestatin analog 2 (PubChem ID: 460617), Squalestatin analog 15 (PubChem ID: 460618), Squalestatin analog 25 (PubChem ID: 460628), Squalestatin analog 32 (PubChem ID: 460634), and Squalestatin analog 36 (PubChem ID: 460638) were downloaded from PubChem database [6] in SDF format. Ligands were prepared using *LigPrep* module of Schrodinger's Maestro that corrects the protonation and ionization states of

the compounds and assign proper bond orders. Subsequently, the tautomeric states and ring conformations /ionization states were generated for each ligand.

B. Molecular docking studies

The molecular docking studies were conducted via *Induced Fit Docking (IFD)* module of (Schrodinger's Maestro 2022-4). A grid box was generated by selecting the co-crystallized ligand and subsequently, the native inhibitor of the 3W7F and all the selected compounds were docked into the active site of the protein using the standard precision (SP) Glide method to obtain the docking score and glide energies for different docking poses.

C. MD simulation studies

The validation of the docking results, top-scoring poses were assessed by all-atom molecular dynamics simulations using *Desmond*. The system was solvated in an orthorhombic box of SPC water model and 0.15 M NaCl was added to neutralize its charge. The system was equilibrated under NPT ensemble class for 100 ns at 300 K and 1.01325 bar and all other parameters were set as default.

D. Binding energy calculation and residue decomposition

The binding energy calculations were carried out by Prime MM-GBSA (Molecular Mechanics/Generalized Born Surface Area) tool [7]. The net binding energy was computed using *ther-mal_mmgbsa.py* script and residue-wise free energy calculations were performed through *breakdown_MMGBSA_by_residue.py* script included in Prime module [8]. The overall binding free energy calculation is expressed as:

$$\Delta G(bind) = Gcomplex - (Gprotein + Gligand)$$

The *breakdown_MMGBSA_by_residue.py* script uses the output obtained from MM-GBSA and calculates the average free energy contribution of each residue.

E. ADMET analysis

To eliminate the failure at later stages of drug discovery all the standard and test compounds were analyzed based on drug likeliness properties using *in-silico* ADMET analysis. The ADMET module is based on atomic logP98, polar surface area (PSA), blood-brain barrier (BBB), cytochrome P450, hepatotoxicity, bioactivity and human oral absorption. The ADMET properties were predicted through *ADMETlab 2.0* online platform [9]. On the other hand, the toxicity prediction of the compounds was done through *PreADMET* [10] based on the properties such as mutagenicity, Carcinogenicity, Acute algae/ daphnia/fish toxicity.

III. RESULTS

A. Molecular Docking Studies through IFD

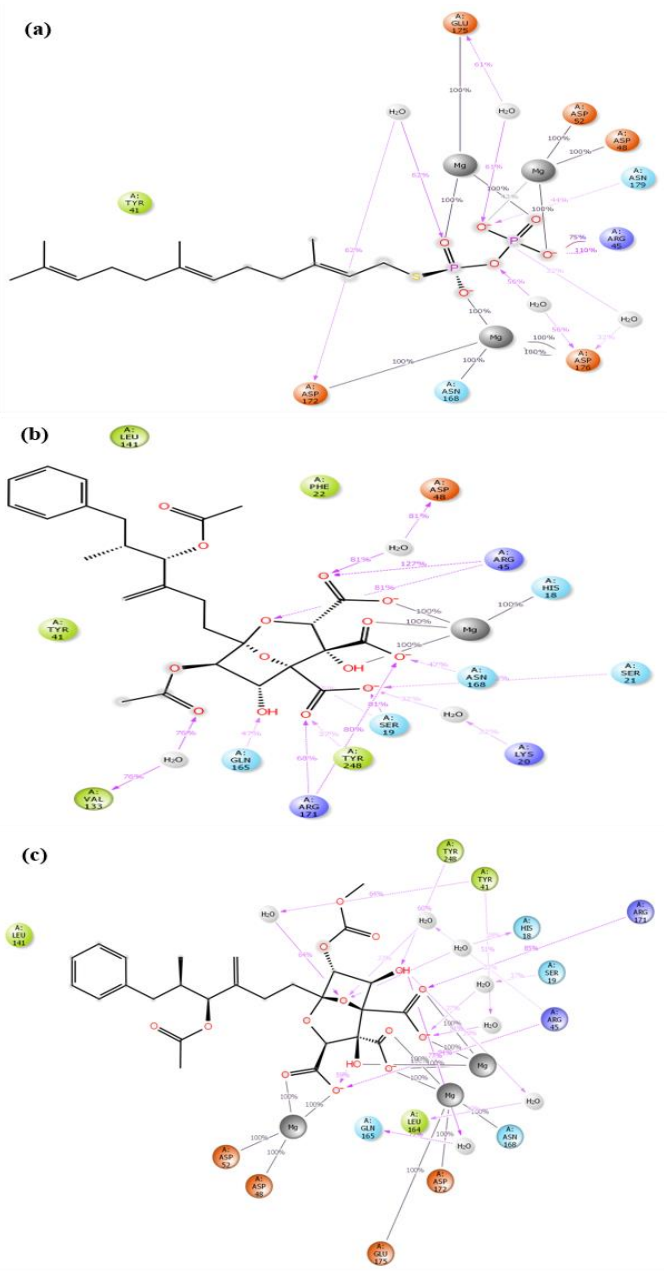
In order to understand the dynamics of interacting residues, all the selected ligands viz., squalestatin analogs were docked against the active site of dehydrosqualene synthase (CrtM) (Protein Id 3W7F) as shown in Fig. 1 (a-c). The interaction of CrtM with cholesterol lowering drugs was proposed through IFD module of Schrodinger's Maestro, and the best pose obtained in terms of docking score with important interacting residues involved in the binding are given in Table 1.

The positive control farnesyl thiopyrophosphate was redocked with CrtM and we obtained docking score (-10.600 kcal/mol) along with formation of three hydrogen bonds with ASN_168, ARG_171 and ARG_265 and two salt bridges with ARG_171 and ARG_265. Out of five selected test compounds, Squalestatin analog 25 exhibited the lowest docking score (-14.399 kcal/mol) and formed five hydrogen bonds with SER_19, ARG_45, ASN_168, ARG_171 and TYR_148; three salt bridges with ARG_45, ARG_171 and ARG_265. On the other hand, Squalestatin analog 36 showed the second lowest docking score (-13.740 kcal/mol) and formed five hydrogen bonds with SER_19, ARG_45, ASN_168, ARG_171 and ARG_265; three salt bridges with ARG_45, ARG_171 and ARG_265. However, ASN_168 was the only common hydrogen bond involved in the binding of standard and test compounds. Due to high binding affinity values of these two compounds and involvement of important amino acids, we proceeded with further evaluation of these 2 compounds through molecular simulations.

TABLE I. Induced fit docking results of standard and test compounds

SI. No.	PDB ID	Compound Name	Docking Score (kcal/mol)	Glide Energy (kcal/mol)	H-Bond interacting residues	Other interacting residues
1	3W7F	Farnesyl thiopyrophosphate	-10.600	-99.513	ASN_168; ARG_171; ARG_265	Salt Bridges: ARG_171; ARG_265
2	3W7F	Squalestatin analog 2	-13.203	-115.259	GLN_165; ASN_168; ARG_171	Salt Bridges: ARG_171; ARG_265 Pi-pi: PHE_22; PHE_26
3	3W7F	Squalestatin analog 15	-13.508	-119.523	ARG_45; GLN_165; ASN_168; ARG_171	Salt Bridges: ARG_45; ARG_171
4	3W7F	Squalestatin analog 25	-14.399	-104.739	SER_19; ARG_45; ASN_168; ARG_171; TYR_248	Salt Bridges: ARG_45; ARG_171; ARG_265
5	3W7F	Squalestatin analog 32	-13.485	-124.578	LYS_20; ARG_45; ASP_48; GLN_165; ASN_168; ARG_265	Salt Bridges: LYS_20; ARG_45; ARG_265
6	3W7F	Squalestatin analog 36	-13.740	-125.499	SER_19; ARG_45; ASN_168; ARG_171; ARG_265	Salt Bridges: ARG_45; ARG_171; ARG_265

Fig. 1. 2D interaction diagrams of (a) farnesyl thiopyrophosphate, (b) squalestatin analog-25 and (c) squalestatin analog-36



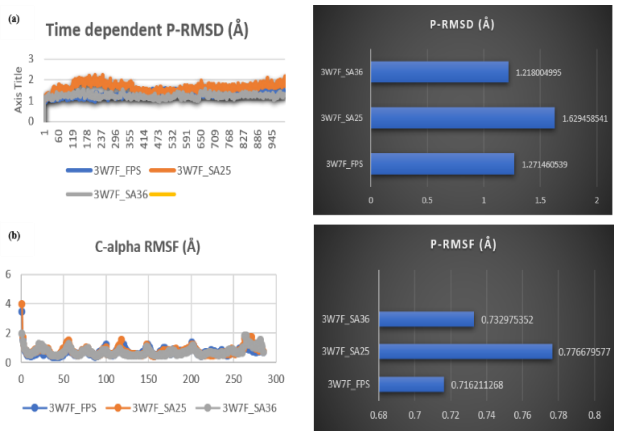
B. Molecular Dynamic Simulations

To obtain more accurate confirmations of the protein due to the flexible nature of the amino acid residues, various thermodynamic properties and stability parameters were evaluated by molecular dynamic simulations using Desmond module of Schrodinger's Maestro [11]. The Root mean square deviation (RMSD), Root mean square fluctuation (RMSF) and various intermolecular interactions (Hydrogen bonds, hydrophobic, ionic and water bridges) for the standard drug and top 2 test compounds viz., squalestatin analog 25 and squalestatin analog 36 were examined with respect to the initial structure through MD simulations for 100ns time frame.

a. ***Protein-ligand RMSD*:** The P-RMSD values must be within the range of 1-4 Å. The prorein-RMSD of all the complexes were calculated and the order is Squalestatin analog 36 (1.218004995) < Farnesyl thiopyrophosphate (1.271460539 Å) Squalestatin analog 25 < (1.629458541 Å). Wherein, the ligand RMSD values should be less than the protein RMSD which indicate that ligand is stable with respect to the protein and its active site. The order of ligand-RMSD is Squalestatin analog 25 (1.418911089 Å) < Farnesyl thiopyrophosphate (1.845926074 Å) < Squalestatin analog 36 (2.13754046 Å). The stability in terms of RMSD values are within the limits of 2 Å (Fig. 2 a).

b. ***Protein-ligand RMSF*:** Protein RMSF represents the local changes along the protein chain. Generally, the tails (N- and C- terminals) fluctuate the most as compared to other part of the protein. The C-alpha RMSF of all the proteins from the complexes were calculated and the results showed that Farnesyl thiopyrophosphate (0.7162113 Å) exhibited the lowest C-alpha RMSF followed by Squalestatin analog 36 (0.732975352 Å) and Squalestatin analog 25 (0.776679577 Å). On the other hand, L-RMSF results showed that Farnesyl thiopyrophosphate (1.231291666 Å) exhibited highest entropic value followed by Squalestatin analog 25 (1.026512195 Å) and Squalestatin analog 36 (0.895166667 Å). (Fig. 2 b)

c. ***Protein-ligand contacts:*** Our results revealed that there are several amino acid residues which are involved in the binding of standard and test compounds apart from those obtained from molecular docking studies. Standard compound i.e., Farnesyl thiopyrophosphate interacted with CrtM by forming strong ionic bonds with ASP_48, ASP_52, ASP_172, GLU_175 and ASP_176; strong H-bonds with ARG_45 and ASN_179; and hydrophobic interactions with TYR_41 and CYS_44. In case of Squalestatin analog 25, it interacted with CrtM by establishing strong ionic bonds with HIS_18; strong H-bond with SER_21, GLN_165 and ARG_265; and hydrophobic interactions with PHE_22 and TYR_41. Similarly, Squalestatin analog 36 interacted with CrtM by forming strong ionic bonds with ASP_48, ASP_52, ASN_168, ASP_172, and GLU_175; strong H-bonds with GLN_165, and TYR_248; and hydrophobic interactions with TYR_41, LEU_141 and LEU_164. (Fig. 2 (c-e))



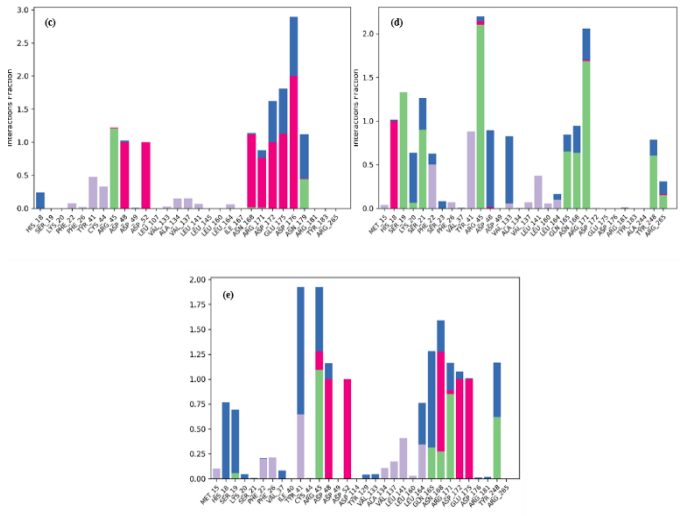


Fig. 2. Molecular dynamics simulation results over 100ns (a) Time dependent protein-RMSD displayed with average mean variation, (b) C-alpha RMSF of the complexes with respect to the residues and mean variation plot of average RMSF values Protein-ligand contacts of (c) farnesyl thiopyrophosphate, (d) squalestatin analog-25 and (e) squalestatin analog-36

C. *Binding free energy calculation*

The calculated binding free energy of the screened compounds ranged from 0.538613618 Kcal/mol to 87.74813056 Kcal/mol (Table 2). The MM-GBSA results illustrates that, Squalestatin analog 25 showed the lowest binding free energy (0.538613618 Kcal/mol) wherein, Squalestatin analog 36 highest binding free energy (87.74813056 Kcal/mol). MM-GBSA results are well correlated with induced fit docking results that showed Squalestatin analog 25 (-104.739 Kcal/mol) and Squalestatin analog 36 (-125.499 Kcal/mol) exhibited better binding affinities. We have also provided involvement of different bonds and their specific involvements during the simulation run. MM-GBSA calculations did not match the exact order as docking energies of complexes. There was slight change in the order of docking energy and binding energy.

TABLE II. MM-GBSA binding free energy calculations (kcal/mol) of the standard and test compounds

S. No.	PDB ID	Compound Name	ΔG Bind	ΔG Bind Coulomb	ΔG Bind H-bond	ΔG Bind vdw
1	3W7F	Farnesyl thiopyrophosphate	41.35093351	-264.9015262	-1.972024881	-40.02498242
2	3W7F	Squalestatin analog 2	43.72861713	-291.5742432	-3.776407168	-40.45544445
3	3W7F	Squalestatin analog 15	68.92768293	-332.5843704	-3.691859329	-54.01321557
4	3W7F	Squalestatin analog 25	0.538613618	-266.1973454	-8.376411491	-49.06439964
5	3W7F	Squalestatin analog 32	70.20593807	-227.8190179	-3.640353706	-50.46958792
6	3W7F	Squalestatin analog 36	87.74813056	-317.5786729	-3.344115849	-46.42899653

D. *Per-residue Energy Decomposition Analysis*

To understand the role of individual amino acid residues per-residue energy decomposition analysis was performed [12]. In this methodology, the total interaction energy is sub-divided into the contributions from individual amino acid residues allowing the quantification of their interaction energy contributions. The residue decomposition analysis shows that there are several other amino acid residues apart from residues obtained from molecular docking and MD simulations that are essential for protein-ligand interaction. These can be considered as potential amino acid candidates for further studies in this domain. Some of the key residues in standard and test compounds; namely, LYS _17, LYS_20, LYS_46, ARG_85, ARG_181 and ARG_265 among the highest contributing residues across the binding process through coulombic interactions.

E. *ADMET Analysis*

To predict the druggability of the compounds more accurately, *ADMETlab 2.0* web tool was used for the ADME analysis. All the standard and test compounds were tested based on the properties such as BBB penetration, MDCK permeability, Caco-2 permeability, HERG blockers, PAINS, Lipinski rule, skin sensitization, eye corrosion, T1/2 and human intestinal absorption (HIA). However, toxicity of all the compounds were predicted using *PreADMET* tool based on the properties such as mutagenicity, Carcinogenicity, Acute algae/ daphnia/fish toxicity. After careful inspection of the compounds it was noticed that the standard compound Farnesyl thiopyrophosphate and test compounds Squalestatin analog 25 and Squalestatin analog 36 fulfilled most of the ADMET properties as compared to other Squalestatin analogs.

IV. Discussion

It is evident from previous studies that cholesterol lowering agents are powerful inhibitors of CrtM which provides a basis for rational drug design methods against *S. aureus*. In present investigation, molecular docking studies were performed by using Induced fit docking (IFD) module of Schrodinger's Maestro to find the binding site of human SQS inhibitors within the active site of CrtM of *S. aureus* to study molecular interaction amongst them. IFD is a process in which a receptor is allowed to move in order to bind to the ligand optimally. The docking scores of all the ligands with the protein ranged between -14.399 Kcal/mol to -13.203 Kcal/mol, which is considerably better than the standard drug-complex i.e., -10.600 Kcal/mol. Further, important amino acid residues involved in protein-ligand interactions were also noted as mentioned in Table 1.

Structure based virtual screening may produce false positives due to rigid body consideration of target protein and lack of solvation effects. Thus, molecular docking results were further validated through MD simulations by analyzing various thermodynamic properties. The P-RMSD values must be within the range of 1-4 Å. The lower RMSD values indicate that the protein structure is more stable and has equilibrated. Protein RMSF represents the local changes along the protein chain. Generally, the tails (N- and C- terminals) fluctuate the most as compared to other part of the protein. Ligand-RMSF signifies

ligand's atom-wise fluctuations, corresponding to its 2D structure. The L-RMSF may give you glimpse about binding of ligand fragments with the protein and their entropic role in the binding. Values up to 2 Å is an acceptable range. There are mainly four types of P-L interactions (or 'contacts') viz., Hydrogen bonds, Hydrophobic, Ionic and Water Bridges. The RMSD and RMSF values indicated that top test compounds i.e., Squalestatin analog 25 and Squalestatin analog 36 were relatively stable and no fluctuations were observed After careful inspection of the thermodynamic properties it was found that test compounds Squalestatin analog 25 and Squalestatin analog 36 were much better than the other test compounds. Based on different molecular interactions the most important amino acid residues were ARG_45, ASP_48, ASP_52, ASN_168 and ASP_172.

To further validate the simulation studies, the binding free energy between the protein and screened ligands were calculated through Prime MM-GBSA (Molecular Mechanics/Generalized Born Surface Area) (Schrodinger Release 2022-4). The MM-GBSA results ranged from 0.538613618 Kcal/mol to 87.74813056 Kcal/mol. Out of five tested compounds, Squalestatin analog 20 showed the lowest binding free energy (0.538613618 Kcal/mol) wherein, Squalestatin analog 36 highest binding free energy (87.74813056 Kcal/mol). Additionally, physicochemical profiling of a drug molecule is one of the major worries and most of the drug candidates fails due to poor physicochemical properties. *In-silico* ADMET analyses were performed to check the drug likeliness properties of the standard and test compounds. It was found that the test compounds Squalestatin analog 25 and Squalestatin analog 36 fulfilled most of the druggable like properties suggesting that these molecules can be absorbed orally, distributed well in the body, and could be considered into further stages of drug-design process. Although, all other analogs passed *in-silico* ADMET properties but overall Squalestatin analog 25 and Squalestatin analog 36 were selected based on showing favorable scores, binding affinity and stability amongst all the studies that we performed in our study.

CONCLUSION

The current research focuses on identifying possible therapeutic inhibitors of dehydrosqualene synthase (CrtM) which is a key enzyme in staphyloxanthin synthesis. An *in-silico* approach was employed to recognize potential inhibitors for CrtM of *S. aureus* to examine various interactions, important amino acid residues within the active site of the protein. Further, several amino acid residues (SER_19, LYS_20, ARG_45, ASP_48, GLN_165, ASN_168, ARG_171, ARG_265 and TYR_248) are essential for the molecules for binding into the active site. The thermodynamic parameters such as RMSD and RMSF were studied to assess the stability of the protein-ligand complex, fluctuations between the residues and target protein over the 100ns simulation run. To conclude, Squalestatin analog 25 and Squalestatin analog 36 shown best binding affinity amongst all other test compounds due to their strong binding within the active site. It is expected that these findings pave way to usage of Squalastatin analogs as potential inhibitors against CrtM. We believe that CrtN is a novel target for a *S. aureus* virulence factor-based treatment and this study would help into the development of virulence factor-based therapy against *S. aureus*.

ACKNOWLEDGMENT

This work has been supported by the Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology and JCMM.

REFERENCES

[1] V. Dalal, P. Dhankhar, V. Singh, et al., "Structure-Based Identification of Potential Drugs Against FmtA of *Staphylococcus aureus*: Virtual Screening, Molecular Dynamics, MM-GBSA, and QM/MM", The protein journal, 40, pp. 148-165, 2021.

[2] S. Banu, R. Bollu, M. Naseema, et al., "A novel templates of piperazinyl-1,2-dihydroquinoline-3-carboxylates: Synthesis, anti-microbial evaluation and molecular docking studies", Bioorganic & Medicinal Chemistry Letters, 28(7), pp.1166-1170, 2018.

[3] P. Gao, J. Davies, and R.Y.T. Kao, "Dehydrosqualene desaturase as a novel target for anti-virulence therapy against *Staphylococcus aureus*", MBio, 8(5), pp. e01224-17, 2017.

[4] A.K. Kahlon, S. Roy, A. Sharma, "Molecular Docking Studies to Map the Binding Site of Squalene Synthase Inhibitors on Dehydrosqualene Synthase of *Staphylococcus Aureus",* Journal of Biomolecular Structure and Dynamics, 28(2), pp. 201-210, 2010.

[5] H. M. Berman et al., "The protein data bank.", Nucleic acids research, 28(1) pp. 235-242, 2000.

[6] S. Kim et al., "PubChem substance and compound databases." Nucleic acids research, 44(D1) pp. D1202-D1213, 2016.

[7] R.A. Elmesseri, S.E. Saleh, H.M. Elsherif, I.S. Yahia and K.M. Aboshanab, "Staphyloxanthin as a potential novel target for deciphering promising anti-staphylococcus aureus agents", Antibiotics, 11(3), pp. 298, 2022.

[8] M.I. Choudhary, M. Shaikh, A. tul-Wahab and A. ur-Rahman, "In silico identification of potential inhibitors of key SARS-CoV-2 3CL hydrolase (Mpro) via molecular docking, MMGBSA predictive binding energy calculations, and molecular dynamics simulation", Plos one, 15(7), pp. e0235030, 2020.

[9] G. Xiong, Z. Wu, J. Yi, et al., "ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties", Nucleic acids research, 49(W1) pp. W5-W14, 2021.

[10] S.K. Lee, I.H. Lee, H.J. Kim, et al., "The PreADME Approach: Web-based program for rapid prediction of physico-chemical, drug absorption and drug-like properties", EuroQSAR 2002 Designing Drugs and Crop Protectants: processes, problems and solutions, pp. 418-420, 2003.

[11] C. Nayak and S.K. Singh, "In silico identification of natural product inhibitors against Octamer-binding transcription factor 4 (Oct4) to impede the mechanism of glioma stem cells", Plos one, 16(10), pp. e0255803, 2021.

[12] A. Gupta, N. Chaudhary and P. Aparoy, "MM-PBSA and per-residue decomposition energy studies on 7-Phenyl-imidazoquinolin-4 (5H)-one derivatives: Identification of crucial site points at microsomal prostaglandin E synthase-1 (mPGES-1) active site", International journal of biological macromolecules, 119, pp. 352-359, 20

Approaches to clustering of microbial colonies on agar plates

1st Michal Cicatka
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
0000-0002-8119-6679

2nd Jan Lancos
Department of Intelligent Systems
Brno University of Technology
Brno, Czech Republic
xlanco00@stud.fit.vutbr.cz

3rd Radim Burget
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
0000-0003-1849-5390

***Abstract*—Agar plates with cultivated microbes are frequently used in many laboratories. Many approaches to analysing images with agar plates have been developed and they aim to automate tasks such as colony counting and colony picking. Image processing is mostly used for detection and/or segmentation of microbes in the agar plates. More complex analysis such as grouping of microbial mixtures on the plates or retrieving additional information about the microbes is still for the most part unexplored. This article proposes three solutions for clustering colonies on agar plates based on their visual properties and suggests a suitable evaluation metric to this problem. The first solution naively clusters all pixels in the image based on its RGB values and serves as a baseline experiment. The second experiment utilizes clustering of manually extracted features calculated for each microbial colony detected separately. The third experiment clusters the features from the custom trained autoencoder used as a feature extractor. The naive RGB clustering, the clustering of manually extracted features and the autoencoder features gained V-measure of 0.870 and 0.911 and 0.912 respectively. The article also presents results from the solutions being applied to real-life agar plate images. The proposed solutions are aimed for use in a prototype of a colony picking robot to which it could bring a next level of autonomy and contribute to another step towards total laboratory automation.**

***Index Terms*—eagar plates, microbial colonies, clustering**

I. Introduction

Agar plates cultivated with microbes are a frequent part of many laboratory workflows. Laboratory automation of tasks related to agar plates came a long way but it is still far from perfect. Especially colony picking tasks are still commonly performed manually by skilled laboratory technicians. Even though some solutions for automated colony picking robots already exist they are still fairly uncommon - especially in smaller laboratories. A crucial feature of colony picking robots is preselection of detected colonies in the image of an agar plate. There are many approaches to localisation of the microbial colonies in the agar plate image. The foundation of such algorithms is usually segmentation of microbes and agar. Some algorithms use traditional image processing techniques [1]–[3] whereas others rely on machine learning [4]–[8]. For completeness, several papers [9], [10] also utilize object detection algorithms - usually when applied to colony counting tasks.

Any kind of complex or advanced analysis of microbes is still very rare. In the article [11] the team applied classification to obtain a colony count in clusters of microbial colonies. In the article [9] the team predicted virulent and avirulent colony forming units. In the article [12] the researchers classified two species of bacteria based on their image properties.

To our knowledge no research aimed at grouping colonies of the same microbes together into clusters has yet been performed. This feature would significantly improve e.g. preselection algorithms in the colony picking robots and allow for advanced analysis in the colony counting tasks. This paper presents three solutions for grouping colonies of similar image properties into clusters. The grouping algorithms are developed for a prototype of an automated picking robot developed by the company Bruker Daltonics GmbH & Co. KG. The target application of the algorithm is to localize all clusters and select one representative of each cluster. We use annotated synthetic images of agar plates with microbes for the development and evaluation of the algorithms. We measured their performance with the V-measure metric (presented in [13]). V-measure is a well-known extrinsic metric for the evaluation of clustering algorithms.

The rest of the paper is structured as follows. Section 2 describes the used materials and methods. Section 3 describes the results and section 4 concludes the subject.

II. Methods and materials

The algorithms for clustering of the colonies were developed and evaluated with synthetic annotated images of agar plates with microbes provided by the company Bruker Daltonics GmbH & Co. KG. The synthetic data generator uses base images of empty agar plates which were scanned with a prototype of a colony picking robot. All agar plates contains at least two groups of microbial colonies and are annotated with segmentation masks. Agar and visible instrument internals are represented with zeros whereas any microbial material is represented with scalar values of the corresponding cluster. Examples of the synthetic images with their annotation masks are in the fig. 1.

Since there is an unknown number of microbial species on each agar plate and the variability of all microbes is too high, solution such as image classification is not suitable. That is why the proposed solutions utilize the K-means clustering algorithm to sort through a set of extracted features. The

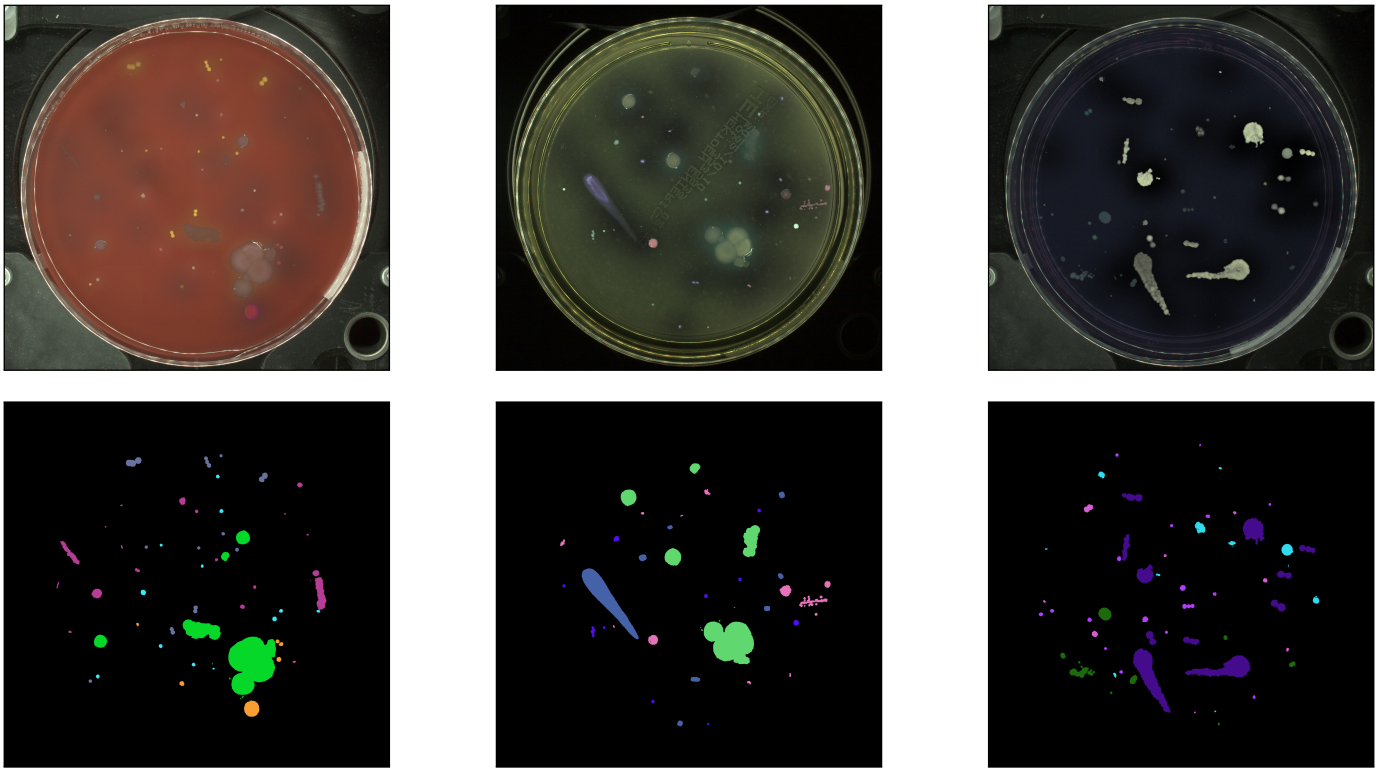


Fig. 1: Examples of synthetic images (top) and their corresponding annotation masks (bottom)

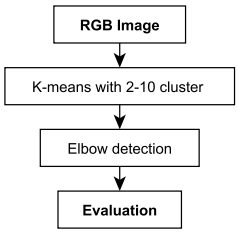


Fig. 2: Diagram of the naive RGB clustering experiment

clustering itself is performed with a two to ten cluster count. We then apply the *Kneedle* algorithm [14] in order to obtain the ideal number of clusters. We use both the inertia and distortion metrics for the elbow point detection. We use greater value of the obtained cluster counts. This stems from the fact that the preselection algorithm should select at least one sample of each microbial type in the agar plate and it is highly undesirable to miss a sample altogether.

The first set of clustering features is a simple vector of the RGB channel values of each pixel - thus each pixel in the image gets clustered. The values in the channels were scaled to the range from zero to one. Since this solution is probably the easiest approach for the image clustering we consider it to be a baseline to overcome with the other two approaches. A diagram of the experiment is in fig. 2.

The second approach (diagram in fig. 3) is different from the naive pixel-wise clustering. We used a colony detection algorithm presented in the article [8] to obtain a set of all colonies in an image with their respective segmentation masks. The aforementioned article uses machine learning segmentation algorithms to obtain a binary mask of all the colonies. The used segmentation model is not flawless so we used the ground truth segmentation masks instead. This way the proposed solution is not bound to the performance of the segmentation model but rather its core principle is challenged. At first we calculate the average color of agar surrounding the colonies with the use of segmentation mask. We then calculate the color properties such as its average color in RGB space and standard deviation of the RGB pixel values weighted by alpha values which are determined as a Manhattan distance between pixels and average color of the surrounding agar. These features represent only the properties of the colony alone. Note that the pixel values were transformed to the range between zero and one.

We suspect that the properties of the nearby surrounding area can be used as significant features too. This stems from the fact that various visual phenomena such as haemolysis or color changes on a chromogenic agar types can commonly be seen on the agar plates. The example of haeomolysis and chromogenic color is in the fig. 4. The example image of haemolysis was slightly adjusted for clearer visibility of the phenomenon. We thus also calculated the average color of the colony background obtained with a similar manner as the

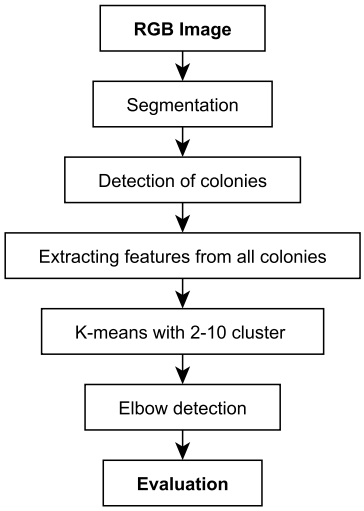


Fig. 3: Diagram of the manual features clustering experiment

average color of the colony with the difference being that the weights were alphas being subtracted from one.

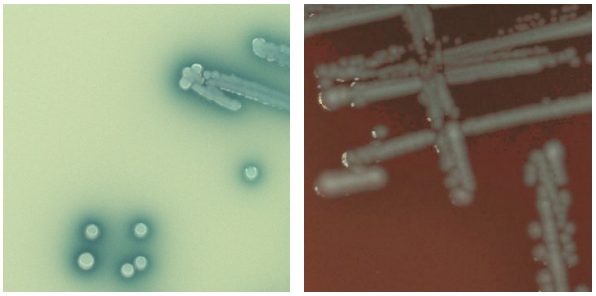


Fig. 4: Examples of the chromogenic agar reaction (on the left) and haemolysis on the blood agar (on the right)

Finally we designed an autoencoder of a U-Net [15] like architecture. We adjusted the output so that its shape matches the input shape. We used a binary cross entropy loss, learning rate of $1e^{-4}$ with a batch size of 16 images and ADAM optimizer. As the input dataset we used the publicly available BRUKERCOLONY dataset presented in [8]. We trained the model with the use of the Tensorflow framework in the Python environment, hardware accelerated with the NVIDIA 4090 GPU. For the purposes of clustering we used the autoencoder model as a feature extractor. The features are taken from the output of the concatenation layer which concatenates the output from the last up-sampling layer and the outermost skip connection. The feature tensor is thus represented as sixteen stacked matrices in the shape of the input image. The clustering is performed in the same manner as the naive RGB clustering (diagram in the fig. 5).

Since we have access to the ground truth annotations we use the V-measure metric for the evaluation of the proposed solutions in a pixel-wise manner. It can therefore be easily applied to the naive RGB clustering and the autoencoder

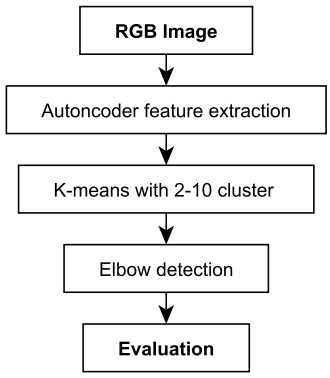


Fig. 5: Diagram of the autoencoder features clustering experiment

clustering but not to the manually extracted features clustering. We tackle this issue by filling a contour of each colony by the values of its given cluster. Example of the cluster infill is in fig. 6.

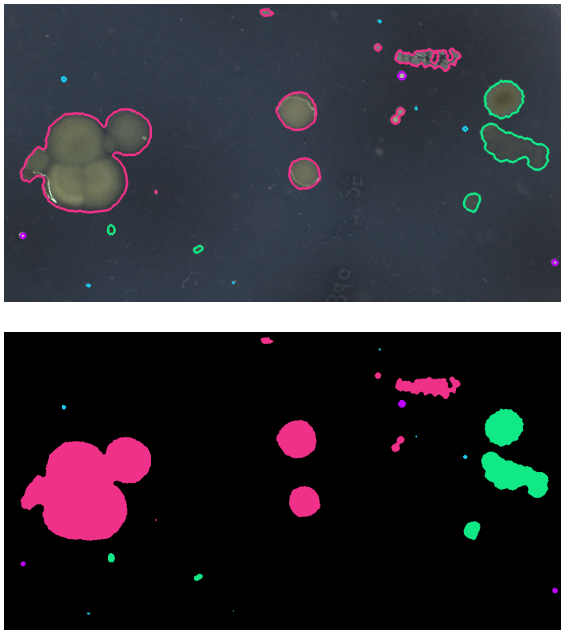


Fig. 6: Example of the contour infill needed for the evaluation

We evaluated four different clustering experiments in total. The first was performed with the naive RGB features and the second and third were performed with the manually extracted features with and without the influence of the surrounding area included. The clustering features in the fourth experiment are obtained with the use of the custom trained autoencoder.

III. Results

We used the V-measure metric for evaluating the results from all of the solutions. The V-measure is an extrinsic metric used for measuring the performance of clustering algorithms. It can be described as a trade-off between homogeneity and

completeness. Homogeneity measures how many samples in the output clusters have the same class labels and completeness then measures how many samples of the same class label are grouped into the same clusters.

$$V = \frac{2 \cdot Homogeneity \cdot Completeness}{Homogeneity + Completeness} \quad (1)$$

The V-measure metric reliably shows the quality of the clustering output against the ground truth labels. As was stated before the proposed solutions are developed for the preselection algorithm of an automated picking robot and thus it is crucial not to miss a sample. This means that we put more emphasis on the homogeneity metric which we included in our results as well. We also included the completeness for the sake of easier comparison to other authors' solutions.

Experiment	V-measure	Homogeneity	Completeness
RGB	0.870	0.890	0.856
Manual basic features	0.909	0.951	0.873
Manual features with background	0.911	0.950	0.877
Autoencoder features	0.912	0.924	0.903

TABLE I: Agar customization experiments result

In the table I there are results of performance evaluation of all experiments. The highest performance in terms of V-measure and completeness was achieved by clustering using the autoencoder features. The highest homogeneity was gained by the clustering algorithm applied to features from colonies and their background. It is clear that features extracted from the surrounding agar make little to no difference. In general, the differences between the solutions, aside from RGB, are comparable. It should be noted that both the manual feature extraction and autoencoder solutions can be further expanded. Adding more features to the manual solution generally requires more domain expertise, but is typically less computationally intensive. In contrast, training and prediction of autoencoders requires more computational power and data, but the features can be extracted without requiring expert knowledge.

We also applied the algorithms to the real-life images of agar plates with microbial mixtures. Since there are no ground truth annotations of the images we assessed the performance only visually. In fig. 7 the are examples of clustering based on the features of colonies and the surrounding agar because the homogeneity of this solution was the highest and it is thus the most suitable for the application in the colony picking robot. In the top row there are examples in which the proposed solution is working properly. It is reliably able to distinguish colonies with different colors even when the colonies are of varying sizes. The figure in the middle shows that the algorithm also distinguishes between colonies with and without visual *halos* around them. In the bottom row several issues of the algorithm are shown. The left figure shows that a colony of mold was grouped with bacterial colonies because of similar color despite different structure. The middle figure shows that a colony of different color was grouped with a cluster of different colonies probably because it was the only outlier. In the fig. on the right visually very similar colonies were clustered into four different clusters.

IV. Conclusion

In this paper we proposed a new area of interest in the field of image processing of agar plates with microbes which to our knowledge has not yet been very explored. We presented three different approaches for clustering of microbial colonies in agar plate images. We used a naive approach for clustering the pixels where each pixel was represented with its RGB values. In the second approach we first detected the microbial colonies in the image and extracted image features from each colony separately. We included average colors of the colonies in the input vector for clustering and we also performed an additional experiment in which we also included features of the agar around the colonies. The third approach utilized features obtained by a custom trained autoencoder. For the evaluation we used the V-measure metric and closely related homogeneity and completeness metrics. The clustering algorithm is designed for a prototype of colony picking robot so more emphasis is put on the homogeneity metric rather than completeness. The best results were achieved by the approach with a manual feature extraction from the previously detected colonies and the surrounding agar with the V-measure 0.911 and homogeneity of 0.950. It must be noted that the results of the manual experiment with the agar features of agar not included have almost identical results. We also present examples of the approaches applied to real-life images. Due to the lack of annotations we assessed the results visually.

This paper is planned to be further expanded. The manually extracted features used for clustering can be significantly enlarged e.g. with features related to texture, shapes and sizes. Using other color spaces such as HSV also comes into question since the hue component separates information about color into a single feature and the value component captures information about brightness, which could be helpful in identifying colonies that have different opacity than others. Furthermore, results obtained using the HSV color space may be more interpretable for humans due to the perception-based nature of the HSV color space. Semi-automated approaches towards the feature extraction such as Gabor filter banks should also be considered. Other architectures of the autoencoder can also be utilized. And finally different clustering algorithms such as DBSCAN or hierarchical UPGMA can be examined as well.

Acknowledgment

This article has been made in close cooperation with Bruker Daltonics GmbH & Co. KG. The company provided annotated dataset of synthetic agar plate images with microbes and other useful resources and know-how.

References

[1] Q. Geissmann, Opencfu, a new free and open-source software to count cell colonies and other circular objects, PloS one 8 (2013) e54072. doi:10.1371/journal.pone.0054072.

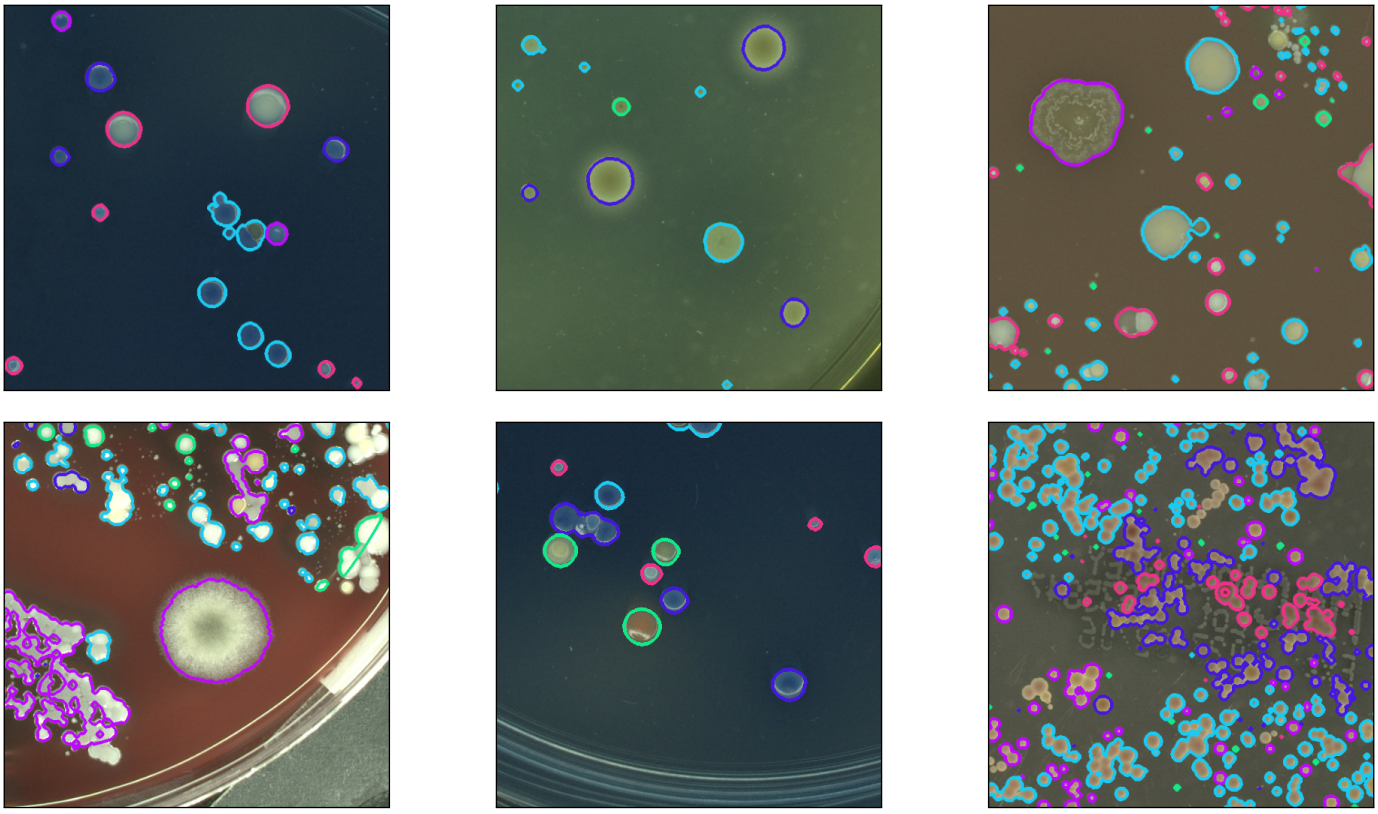


Fig. 7: Examples of clustering on the real-life agar plate images

[2] A. u. M. Khan, A. Torelli, I. Wolf, N. Gretz, Autocellseg: robust automatic colony forming unit (cfu)/cell analysis using adaptive image segmentation and easy-to-use post-editing techniques, Scientific Reports 8 (1) (2018) 7302. doi:10.1038/s41598-018-24916-9.
URL https://doi.org/10.1038/s41598-018-24916-9

[3] D. R. Stirling, M. J. Swain-Bowden, A. M. Lucas, A. E. Carpenter, B. A. Cimini, A. Goodman, Cellprofiler 4: improvements in speed, utility and usability, BMC Bioinformatics 22 (1) (2021) 433. doi:10.1186/s12859-021-04344-9.
URL https://doi.org/10.1186/s12859-021-04344-9

[4] D. Nie, E. A. Shank, V. Jojic, A deep framework for bacterial image segmentation and classification, in: Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics, BCB '15, Association for Computing Machinery, New York, NY, USA, 2015, p. 306–314. doi:10.1145/2808719.2808751.
URL https://doi.org/10.1145/2808719.2808751

[5] P. Andreini, S. Bonechi, M. Bianchini, A. Mecocci, F. Scarselli, A deep learning approach to bacterial colony segmentation, in: V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis, I. Maglogiannis (Eds.), Artificial Neural Networks and Machine Learning – ICANN 2018, Springer International Publishing, Cham, 2018, pp. 522–533.

[6] G. Zhu, B. Yan, M. Xing, C. Tian, Automated counting of bacterial colonies on agar plates based on images captured at near-infrared light, Journal of Microbiological Methods 153 (2018) 66–73. doi:https://doi.org/10.1016/j.mimet.2018.09.004.
URL https://www.sciencedirect.com/science/article/pii/S0167701218306146

[7] P. Andreini, S. Bonechi, M. Bianchini, A. Mecocci, F. Scarselli, Image generation by gan and style transfer for agar plate image segmentation, Computer Methods and Programs in Biomedicine 184 (2020) 105268. doi:https://doi.org/10.1016/j.cmpb.2019.105268.
URL https://www.sciencedirect.com/science/article/pii/S0169260719311216

[8] M. Cicatka, R. Burget, J. Karasek, Machine-learning approach to microbial colony localisation, in: 2022 45th International Conference on Telecommunications and Signal Processing (TSP), 2022, pp. 206–211. doi:10.1109/TSP55681.2022.9851236.

[9] T. Naets, M. Huijsmans, P. Smyth, L. Sorber, G. de Lannoy, A mask r-cnn approach to counting bacterial colony forming units in pharmaceutical development (2021). doi:10.48550/ARXIV.2103.05337.
URL https://arxiv.org/abs/2103.05337

[10] J. Pawłowski, S. Majchrowska, T. Golan, Generation of microbial colonies dataset with deep learning style transfer, Scientific Reports 12 (1) (2022) 5212. doi:10.1038/s41598-022-09264-z.
URL https://doi.org/10.1038/s41598-022-09264-z

[11] A. Ferrari, S. Lombardi, A. Signoroni, Bacterial colony counting with convolutional neural networks in digital microbiology imaging, Pattern Recognition 61 (2017) 629–640. doi:https://doi.org/10.1016/j.patcog.2016.07.016.
URL https://www.sciencedirect.com/science/article/pii/S0031320316301650

[12] S. A. Nagro, M. Kutbi, W. M. Eid, E. J. Alyamani, M. H. Abutarboush, M. A. Altammami, B. K. Sendy, Automatic identification of single bacterial colonies using deep and transfer learning, IEEE Access 10 (2022) 120181–120190.

[13] A. Rosenberg, J. Hirschberg, V-measure: A conditional entropy-based external cluster evaluation measure., 2007, pp. 410–420.

[14] V. Satopaa, J. Albrecht, D. Irwin, B. Raghavan, Finding a "kneedle" in a haystack: Detecting knee points in system behavior, in: 2011 31st International Conference on Distributed Computing Systems Workshops, 2011, pp. 166–171. doi:10.1109/ICDCSW.2011.20.

[15] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation (2015). doi:10.48550/ARXIV.1505.04597.
URL https://arxiv.org/abs/1505.04597

Improvement of microbial colony segmentation using synthetic data

1st Michal Cicatka
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
0000-0002-8119-6679

2nd Jan Lancos
Department of Intelligent Systems
Brno University of Technology
Brno, Czech Republic
xlanco00@stud.fit.vutbr.cz

3rd Radim Burget
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
0000-0003-1849-5390

***Abstract*—This paper presents an image processing pipeline for the generation of synthetic agar plate images. The generator superimposes patches of microbial colonies onto the images of empty agar plates. Data acquisition and processing of both the empty agar plate images and colony patches is described. A testing dataset was created which consists of real-life agar plate images with various microbes cultivated on various agars. The performance of the generator is evaluated by trained U-Net models with training and validation dataset enlarged by 33 %, 66 %, 100 %, 150 % and 200 % of artificially created data. The best performance (0.729 F1 score and 0.656 Jaccard index) was achieved by a model trained on 150 % of additional synthetic data. However, all models trained with more than 66 % of the artificially generated data showed comparable performance, with only small differences in F1 score and Jaccard index.**

***Index Terms*—agar plates, synthetic data generator, segmentation, microbial colonies**

I. Introduction

Agar plates with cultivated microbes still frequently occur in many laboratory workflows such as colony picking and colony counting. Especially colony picking is to this day regularly performed by skilled laboratory technicians manually. Since this task is both tedious and repetitive the idea of instrument automation arises. Available colony picking systems usually offer manual selection of the colonies. However, automated colony preselection has already proven itself to be of great use. The foundation of algorithms responsible for the colony localisation is often segmentation of microbes and agar on the plates [1]–[8]. Even though articles dealing with the segmentation of microbial colonies have been published, they usually focus on one type of agar and/or microbe. Moreover to our knowledge there are not any available annotated data which could be used for development and especially evaluation of such methods. Acquisition and annotation of biomedical data is an expensive and time-consuming task. An answer to this issue is generation of annotated synthetic data. Studies have been conducted specifically regarding generation of agar plate images [9], [10]. Their common major drawback is the aforementioned lack of input data and evaluation.

This article presents a pipeline for generation of synthetic agar plate images. The quality of the proposed solution is evaluated by the the performance of a trained deep learning segmentation model on real data. U-Net [11] was chosen for the evaluation as it is the gold standard in biomedical image segmentation.

When it comes to datasets that contain labeled agar plates with microbial colonies, the selection is rather limited. One of the more well-known datasets is the MicrobIA Haemolysis dataset [3] which includes 235 images of agar plates with microbes that are labeled through pixel wise segmentation masks. Another dataset, the Segments enumeration dataset, contains roughly 28 500 images of microbial patches, each with a classification label indicating the number of colonies present in the image. The AGAR dataset [12], which is suitable for colony counting applications, includes 18 000 plates and labels each colony with a bounding box. The recently published BRUKERCOLONY dataset [13] contains 547 hand-labeled images of agar plates, mostly with microbial monocultures on a red or white agar. This dataset will be used in the methodology of this article.

The rest of the paper is structured as follows. Section 2 describes the used materials and methods. Section 3 describes the results and section 4 concludes the subject.

II. Materials and methods

The main objective of this paper is to implement and evaluate a generator of annotated synthetic agar plate images. The annotation is represented as a segmentation map that consists of binary values since we try to segment all microbial material on an agar plate as one class. The proposed solution is evaluated by the performance of a trained U-Net segmentation model. U-Net architecture was chosen for this paper because it is a gold standard in the deep learning image segmentation. This fact makes it easier for other authors to compare their solutions with ours. The generator requires a base image represented by an image of an empty agar plate and content images represented by the images of so called colony patches. The colony patches must be annotated.

We used the publicly available BRUKERCOLONY dataset as the initial source of our data. The BRUKERCOLONY dataset contains 547 images of plates with mostly red blood agar and white Sabourad agar. Most of the agar plates are monocultures which means that usually only one type of microbe is present. We divided the dataset to a training (70 %), validation (20 %) and testing (10 %) part. The training

and validation data was used for training the U-Net baseline comparison model. Since the data of the BRUKERCOLONY dataset lacks variability we cultivated microbes on various agar plates. Those data were then scanned under six various lighting conditions using a prototype of colony picking robot and manually annotated. We enlarged the testing part of the BRUKERCOLONY dataset with the newly obtained data. 866 images were scanned of which 43 images were included in the testing dataset. We did not use agar plates which were contaminated with mold or contained little to no microbial material. This testing dataset is used for evaluation of all the trained models including the baseline. Examples of the training data are in fig. 1.

With the use of a prototype of an automated picking robot we scanned the several agar plates without any cultivated microbes under different lighting conditions. These images are used as base images for the generator. Examples of the agar plate images are in fig. 2.

We used the training part of the BRUKERCOLONY dataset for the acquisition of the colony patches used in the generator as content images. We adjusted the processing pipeline proposed in [13] to localise and extract isolated colonies with their corresponding annotation masks. These images were then visually inspected and only images which contained whole colonies were included in the dataset. In some cases when the colony patch contained more colonies the annotation mask was modified so that it contained only one colony of interest (fig. 3). For easier superimposition of the colony patches we added a transparency layer by calculating the alpha channel as a distance of each pixel channel from the average background color. The mean background is calculated with the use of the annotation mask:

$$PB = Mean(PI_{xy}) : \ PM_{xy} \neq 0 \tag{1}$$

$$\alpha_R = |I_{xy_R} - PB_R| \tag{2}$$

$$\alpha_G = |I_{xy_G} - PB_G| \tag{3}$$

$$\alpha_B = |I_{xy_B} - PB_B| \tag{4}$$

$$PI_{xy_\alpha} = 255 - (\alpha_R + \alpha_G + \alpha_B) \tag{5}$$

where PB denotes mean colony patch background, PI denotes the colony patch image, PM denotes the colony patch mask and α denotes the alpha channel of the colony patch. Many microbial colonies are translucent and the agar may be visible on the colony patches. This phenomenon is compensated with a color adjustment according to the equations:

$$CS = PIM - PB \tag{6}$$

$$API_{xy} = PI_{xy} - CS \cdot (1 - \frac{PI_{xy_\alpha}}{255}) \tag{7}$$

where CS denotes a color shift, PIM denotes mean color of the colony patch, PB denotes a mean colony batch background color, API denotes adjusted colony patch image and PI denotes a colony patch image.

We were able to sort the colony patches to groups of visually similar microbes because most of the microbial colonies come from monocultures.

In total, 20 417 colony patches were extracted and 1 610 empty agar plate images acquired.

We proposed a simple yet effective algorithm for the synthetic image generation (diagram in fig. 4). A random empty agar plate image is selected as the base image for the generator. The target coverage of the agar plate is set and several random cultures are then chosen. Patches from the culture groups are then selected so that the total area of the microbes is within a 5 % bounds of the target coverage. The chosen patches are then randomly rotated to increase the variability of the agar plates. The positions of the colony patches are optimized with a genetic algorithm based on the loss function:

$$Current = \sum_{n=1}^{N} M_{xy} : \ (M_{xy} \neq 0 \wedge |M_{xy}C| < r) \tag{8}$$

$$Total = \sum_{n=1}^{N} PM_{n_{xy}} : \ PM_{n_{xy}} \neq 0 \tag{9}$$

$$Fitness = \frac{Current}{Total} \tag{10}$$

where M denotes the currently assessed mask, PM denotes the mask of a colony patch, C denotes the center of the image and r denotes the radius of the Petri dish. The loss thus estimates the ratio between area of the currently assessed mask without pixels outside of the Petri dish radius and total area of the colony patches. Fitness increases when there are fewer overlapping colony patches and colony patches outside of the Petri dish. The optimization is ended once fitness is equal to one. The genetic algorithm optimization is performed on resized images to decrease the computational expense of the algorithm. This decreases the minimum size of a single colony patch. This was however empirically proved to be an insignificant factor.

Before the superimpositions take place the color of all the patches selected from a single culture is adjusted by the same margin to increase the variability of the data while keeping the cultural similarities. The color is changed in the HSV color space. To ensure the colony patches looks realistic and the manipulation of both the saturation and value channel can be done for entire cultures without risking an overflow in any of the affected colonies' pixels, both the minimum and maximum values for both channels are found prior. The following shift in saturation and value is then selected from a range reflecting these extremes so that no overflow can occur. The hue is changed by a random offset no larger than 90°.

Recolored colony patches are then superimposed onto the empty agar plate with a simple blending algorithm which weights each pixel of the colony patch by its alpha channel and adds it to the pixel of the agar plate. This leads to a smooth blending of the colony edges and a realistic synthetic image. This process can be described by the equation:

$$SIxy = (1 - PI_{ij_\alpha}) \cdot BI_{xy} + PI_{ij_\alpha} \cdot PI_{ij} \tag{11}$$

where SI denotes the synthetic image, BI denotes the baseline empty agar image and PI denotes the patch image.

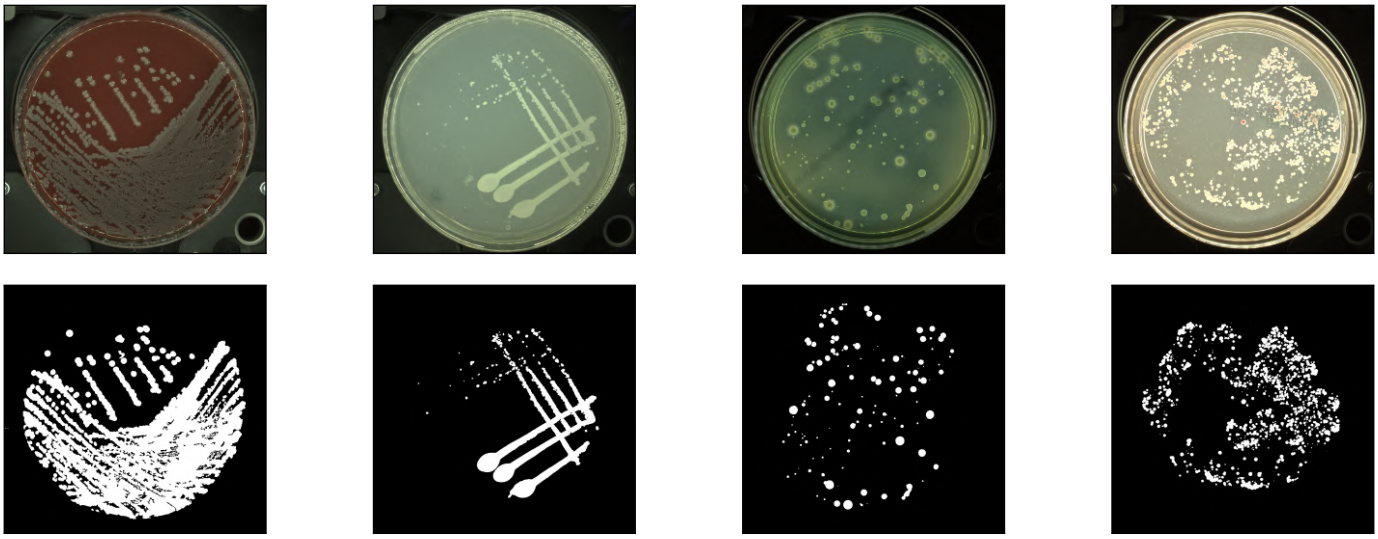


Fig. 1: Examples of two samples from the BRUKERCOLONY dataset (left) and two newly obtained samples (right). In the bottom row their annotation is shown.

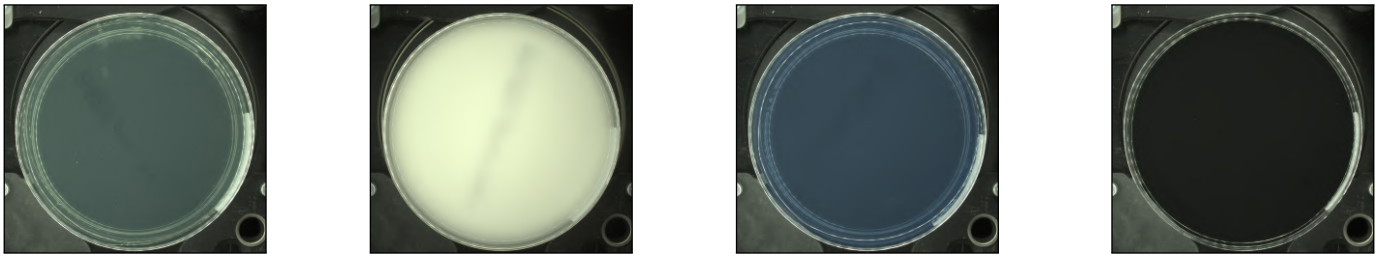


Fig. 2: Examples of base images for the generator



Fig. 3: Removal of an incomplete colony

In figures 5 there are examples of the generated images. In the top row there are examples of the synthetic images generated with the proposed solution. In the middle row there are corresponding identical images, the difference being the colony patches were superimposed without the blending algorithm. In the bottom row there are the annotation masks.

We used the generator to obtain 985 synthetic agar plate images. We enlarged the U-Net training validation dataset by 33 %, 66 %, 100 %, 150 % and 200 % with said synthetic data and rerun the training session. The training of the models was performed with the use of Tensorflow framework in Python environment hardware accelerated by NVIDIA RTX 4090 GPU. The dataset images have the size of 1900 to 1800 pixels which was too large to train such a network with available hardware. We therefore split each image into sixteen smaller images with the size of 512 to 512 pixels. The used hardware

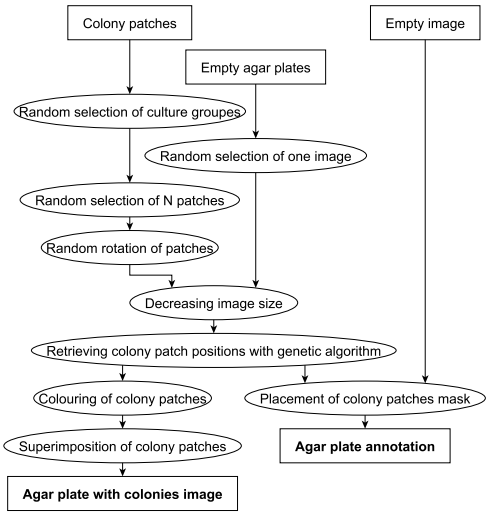


Fig. 4: Diagram of the image generation algorithm

allowed us to train all models with minibatches of 16 images. We used the ADAM optimization algorithm with the learning

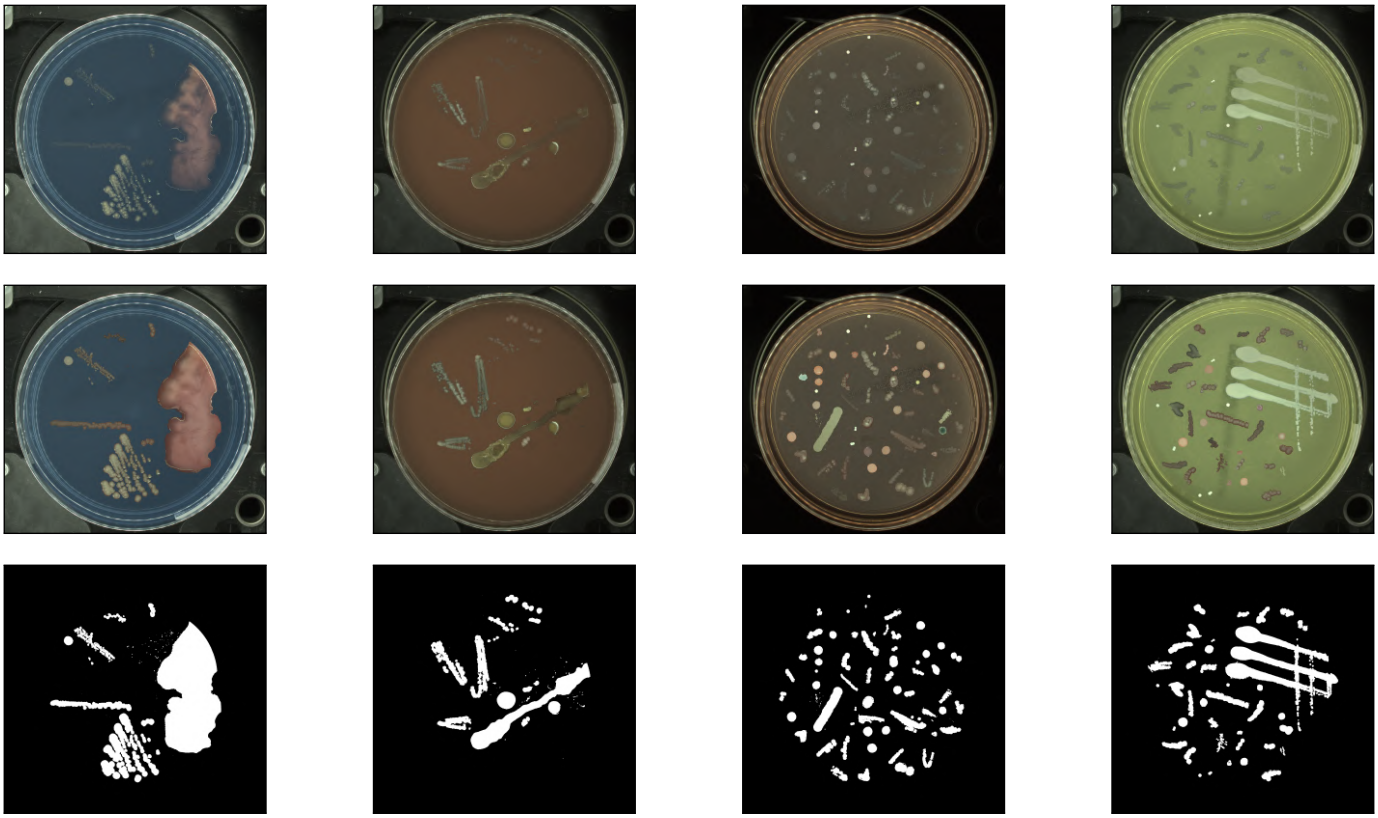


Fig. 5: Examples of the generated synthetic images

rate of $1e^{-4}$ and an exponential decay after 10 epochs. We used soft Dice loss as our training loss function:

$$SoftDice = \frac{2TP + 1}{2TP + FP + FN + 1} \tag{12}$$

$$Loss = 1 - SoftDice \tag{13}$$

where TP denotes true positive, FP denotes false positive and FN denotes false negative.

III. Results

We trained 6 neural networks in total with different amounts of synthetic data. The experiment results are in the table I. We used the F1 score (eq. 14) and the Jacard score (eq. 15) for segmentation evaluation since they are the two most well-known and used metrics. In the equations TP denotes true positive, FP denotes false positive and FN denotes false negative.

$$F1\ score = \frac{2TP}{2TP + FP + FN} \tag{14}$$

$$Jaccard\ index = \frac{TP}{TP + FP + FN} \tag{15}$$

Due to the image split to smaller images there were some samples without any true positive values. These images were discarded from evaluation with the trained models because division by zero could occur.

Experiment	F1 score	Jaccard index
Baseline dataset	0.518	0.453
Baseline dataset augmented by 33 %	0.677	0.606
Baseline dataset augmented by 66 %	0.719	0.647
Baseline dataset augmented by 100 %	0.726	0.656
Baseline dataset augmented by 150 %	**0.729**	**0.656**
Baseline dataset augmented by 200 %	0.717	0.645

TABLE I: Agar customization experiments result

We can see a clear trend of both the F1 score and the Jaccard index increasing when synthetic data is included in the training and validation dataset up until 150 % enlargement.

The model trained with data enlarged by 150 % achieved the best performance in terms of F1 score. The model trained with 200 % of synthetic data achieved a significantly lower performance. This fact is probably caused by there not being enough variable data of the colony patches, especially the large areas of microbial material which lead to model overfitting. The best model performance increased from 0.518 to 0.729 with the difference of 0.211 (increase by 40.7 %) in terms of F1 score and 0.453 to 0.656 with the difference of 0.203 (increase by 44.8 %) in terms of Jaccard index when compared to the baseline model. It is also worth mentioning that adding synthetic data, even in small amounts (33 %), improved the F1 score (0.677, up by 30.7 %) and Jaccard index (0.606, up by 33.8 %) and models trained with over 66 % of synthetic

data showed minimal differences in performance.

In the article [10] the team achieved a significantly higher Jaccard index of 0.936. However, it must be disclosed that they used a much less variable data from the microbIA dataset [3] and the validation was performed on samples of solely red blood agar. Other types were assessed visually due to the lack of annotated data.

Further comparison with the article [9] is very limited because the proposed agar plate generator was evaluated by object detection models since the targeted application was in the colony counting tasks rather then segmentation models. This also means that the generated agar plates mostly lacked large areas of microbes which are very common in laboratory workflows besides colony counting. Despite the use of advanced techniques such as style transfer the output data were mostly plates with white agar.

IV. Conclusion

In this article we presented an algorithm for generating synthetic agar plates. We gathered samples of empty agar plates which were then used as base images and a large dataset of colony patch images which were superimposed onto the base images by the generator. The performance of the generator was evaluated by segmentation U-Net models which were trained on real data with different amounts of synthetic data mixed into the training and validation dataset. Publicly available data were not variable enough for a conclusive evaluation of our proposed solution so we also scanned and annotated real-life plates with various types of agar plates and microbes. These images were used for the evaluation of the trained models. Performance of the best model trained with training and validation dataset enlarged by 150 % of synthetic data was increased by the difference of 0.211 (an increase of 40.7 %) F1 score and 0.203 (an increase of 44.8 %) Jaccard index in comparison to the trained baseline model. This suggests that the generation pipeline provides qualitative results. Nevertheless, it is highly plausible that other segmentation architectures could perform better (e.g. [13], [14]). Performance increase could also be seen if the models were trained on larger images so that a larger area of interest could be fed to the model.

The generation algorithm is easily reusable for other purposes. The possibility of configuring the target colony count can be very useful for colony counting experiments. Our sorted colony patch dataset also enables to generate annotated synthetic images with distinct microbes. This could be used e.g. in multiclass segmentation of microbes or various post-processing analyses.

The work is planned to be further improved. A common visual phenomena such as haemolysis or colouring of chromogenic agars exist which we wish to explore. The generator is also going to be tweaked and reused for generation of plates with multiple microbes each with its own value in the annotation mask. This dataset is going to be used for the exploration of clustering techniques applied to the problem of grouping colonies of the same microbes together.

Acknowledgment

This article has been made in close cooperation with Bruker Daltonics GmbH & Co. KG. The company provided us with a prototype of its colony picking robot and other useful resources such as data and know-how.

References

[1] Q. Geissmann, Opencfu, a new free and open-source software to count cell colonies and other circular objects, PloS one 8 (2013) e54072. doi:10.1371/journal.pone.0054072.

[2] D. Nie, E. A. Shank, V. Jojic, A deep framework for bacterial image segmentation and classification, in: Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics, BCB '15, Association for Computing Machinery, New York, NY, USA, 2015, p. 306–314. doi:10.1145/2808719.2808751.
URL https://doi.org/10.1145/2808719.2808751

[3] A. Ferrari, S. Lombardi, A. Signoroni, Bacterial colony counting with convolutional neural networks in digital microbiology imaging, Pattern Recognition 61 (2017) 629–640. doi:https://doi.org/10.1016/j.patcog.2016.07.016.
URL https://www.sciencedirect.com/science/article/pii/S0031320316301650

[4] G. Zhu, B. Yan, M. Xing, C. Tian, Automated counting of bacterial colonies on agar plates based on images captured at near-infrared light, Journal of Microbiological Methods 153 (2018) 66–73. doi:https://doi.org/10.1016/j.mimet.2018.09.004.
URL https://www.sciencedirect.com/science/article/pii/S0167701218306146

[5] P. Andreini, S. Bonechi, M. Bianchini, A. Mecocci, F. Scarselli, A deep learning approach to bacterial colony segmentation, in: V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis, I. Maglogiannis (Eds.), Artificial Neural Networks and Machine Learning – ICANN 2018, Springer International Publishing, Cham, 2018, pp. 522–533.

[6] A. u. M. Khan, A. Torelli, I. Wolf, N. Gretz, Autocellseg: robust automatic colony forming unit (cfu)/cell analysis using adaptive image segmentation and easy-to-use post-editing techniques, Scientific Reports 8 (1) (2018) 7302. doi:10.1038/s41598-018-24916-9.
URL https://doi.org/10.1038/s41598-018-24916-9

[7] T. Beznik, P. Smyth, G. de Lannoy, J. A. Lee, Deep learning to detect bacterial colonies for the production of vaccines (2020). doi:10.48550/ARXIV.2009.00926.
URL https://arxiv.org/abs/2009.00926

[8] T. Naets, M. Huijsmans, P. Smyth, L. Sorber, G. de Lannoy, A mask r-cnn approach to counting bacterial colony forming units in pharmaceutical development (2021). doi:10.48550/ARXIV.2103.05337.
URL https://arxiv.org/abs/2103.05337

[9] J. Pawłowski, S. Majchrowska, T. Golan, Generation of microbial colonies dataset with deep learning style transfer, Scientific Reports 12 (1) (2022) 5212. doi:10.1038/s41598-022-09264-z.
URL https://doi.org/10.1038/s41598-022-09264-z

[10] P. Andreini, S. Bonechi, M. Bianchini, A. Mecocci, F. Scarselli, Image generation by gan and style transfer for agar plate image segmentation, Computer Methods and Programs in Biomedicine 184 (2020) 105268. doi:https://doi.org/10.1016/j.cmpb.2019.105268.
URL https://www.sciencedirect.com/science/article/pii/S0169260719311216

[11] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation (2015). doi:10.48550/ARXIV.1505.04597.
URL https://arxiv.org/abs/1505.04597

[12] S. Majchrowska, J. Pawłowski, G. Guła, T. Bonus, A. Hanas, A. Loch, A. Pawlak, J. Roszkowiak, T. Golan, Z. Drulis-Kawa, Agar a microbial colony dataset for deep learning detection (2021). arXiv:2108.01234.

[13] M. Cicatka, R. Burget, J. Karasek, Machine-learning approach to microbial colony localisation, in: 2022 45th International Conference on Telecommunications and Signal Processing (TSP), 2022, pp. 206–211. doi:10.1109/TSP55681.2022.9851236.

[14] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, J. Liang, Unet++: A nested u-net architecture for medical image segmentation (2018). doi:10.48550/ARXIV.1807.10165.
URL https://arxiv.org/abs/1807.10165

Computational approaches to find potential phytochemical for ADHD treatment

Vaishali Pankaj[1]
[1]*Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology*
Brno, Czech Republic
234084@vut.cz

Inderjeet Bhogal[1]
[1]*Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology*
Brno, Czech Republic
234190@vut.cz

Sudeep Roy[1]
[1]*Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology*
Brno, Czech Republic
roy@vut.cz

***Abstract*— The presented study is focused on investigating the potential role of phytochemical compounds for the treatment of ADHD. Comparative analysis is carried out for compounds with the dopamine receptor in presence and absence of sodium cation (Na^{+}). Molecular docking studies suggest that the compound silibinin has higher binding affinity in presence of Na^{+}. This was further validated through molecular dynamics simulation and binding free energy calculations which shows that it binds strongly and has a more stable complex. Our study suggests that silibinin could be considered as a potential candidate for the development of drug against ADHD.**

Keywords—ADHD, phytochemicals, binding affinity, silibinin

I. Introduction

Attention deficit hyperactive disorder (ADHD) is a neurodevelopmental disorder (NDD) where the developmental changes lead to long-lasting behavioral and physiological deficits affecting 8.5% of children and 2.5% adults [1]. The symptoms observed are lack of attention, hyperactive and impulsive behavior [2]. There is no specific cause and cure for NDDs, but some evidence suggests that imbalance in the neurotransmitters such as dopamine and norepinephrine lead to the cause for neurological disorders [3]. These neurotransmitter receptors play a role in keeping the balance of behavioral and cognitive functions in the central nervous system. Treatments are available both pharmacological and non-pharmacological (behavioral therapy) but unfortunately most individuals are unable to tolerate drugs side effects such as nausea, abdominal pain, headache, decreased appetite, and insomnia. Therefore, for safety medication, natural plant-derived phytochemical compounds are used that have no side effects and a long-term effect [4].

The objective of the present study is to investigate the potential role of phytochemical compounds with crystal structure of dopamine receptor in presence and absence of sodium (Na). Computational approaches were used for the identification of responsible genes for the construction of protein-protein interaction (PPI) network. ADMET screening of neuroprotective phytochemical compounds was carried out based on their therapeutic properties. Molecular docking studies were carried out to generate favorable docked poses. The thermodynamic properties of docked complexes are further evaluated by molecular dynamics simulation. Binding free energy calculations for complexes were performed through MM-GBSA studies to provide significant correlation with experimentally determined data.

II. Material and Methods

A. *Identification of genes and PPI network construction*

The genes involved in ADHD were retrieved from different databases as NCBI [5], MalaCard [6], GeneCard [7] and DisGeNET [8] to identify common genes. The protein-protein interaction (PPI) network for common genes was constructed using STRING database [9].

B. *Ligand identification and ADMET screening*

The neuroprotective phytochemicals were selected as ligands from various literatures. The 3D structures of the phytochemicals were downloaded from PubChem database. ADMET screening was performed using ADMETLab 2.0 webserver tool [10] to study various druggable properties of phytochemical molecules for further screening.

C. *Protein and ligand preparation*

The crystal structure of dopamine receptor was retrieved from RSCB Protein Data Bank (PDB) database. Two structures were selected as 1) Structure of human D4 dopamine receptor in complex with nemonapride (PDB ID:5WIU); 2) Structure of sodium bound human D4 dopamine receptor in complex with nemonapride (PDB ID: 5WIV). Both the protein structures were prepared using Protein preparation workflow in Maestro 13.0 of Schrodinger suite 2022-4 [11]. Ligands were prepared using LigPrep with an OPLS4 force field in Maestro 13.0. It generates 32 possible stereochemical structures per ligand.

D. *Molecular Docking*

Induced-fit docking (IFD) was performed to understand the interactions between DRD4 protein with sodium-bound and without sodium structures. IFD uses rigid receptor docking (using Glide) to provide the best docking conformations. Ten poses were generated for each docked complex and the best scoring pose was chosen based on the docking scores.

E. *Molecular dynamics simulation*

MD simulation study was performed for best docking score complex to investigate the conformational changes during the docking. It was run for 100ns using Desmond. Simulation process comprises of three processes: System builder, Energy Minimization and Molecular dynamics via NPT ensemble run. The pressure is adjusted to 1 bar and temperature up to 300K. Simulation interaction diagram was generated for the calculation of thermodynamic properties.

F. *Binding free energy calculation*

The free energy of the binding of ligands to biological macromolecules were calculated using prime MM-GBSA (Molecular Mechanics/Generalized Born Surface Area) [12]. This method computes the difference between the free energy of protein, ligand and complex in solution. The prime Dg (MM-GBSA) is calculated by the following equation:

dGbind (MM-GBSA) = Gcomplex - Gprotein- Gligand

To better understand the contribution of each residue interaction with the ligands, the binding free energies were decomposed into individual residues by MM-GBSA decomposition method.

III. Results and Discussion

A. *Identification of genes and PPI network construction*

The number of identified genes is 251 (NCBI), 198 (MalaCard), 1920 (GeneCard) and 842 (DisGeNET). Among these, 74 were retrieved as common genes from all the databases. PPI network was constructed for these common genes using STRING, which showed that DRD4 gene has the strongest interaction network with other genes. Recent studies support that dopamine receptors have an active role in the pathogenesis of ADHD, also responsible for neuronal signaling in mesolimbic system of brain [13].

B. *ADMET screening of ligands*

Twenty phytochemical compounds obeyed the ADMET properties. Some of the properties that are within the acceptable range include Molecular weight, Hydrogen bond acceptor and donors, logP, CNS activity, Blood brain barrier (BBB), Madin-Darby canine kidney cells (MDCK) permeability, Human intestinal absorption (HIA), Caco-2 permeability, human oral absorption and carcinogenicity. Some of the top screened compounds are silibinin (*Silybum marianum)*, berberine (*Berberis asiatica)*, curcumin (*Curcuma longa)*, quercetin (*Allium cepa)*, naringenin (*Dalbergia parviflora)*, resveratrol (*Vitis vinifera)*, leonurine (*Leonurus sibiricus)* and genistein (*Glycine max)*.

C. *Molecular docking*

IFD studies revealed that silibinin compound has higher binding affinity in both presence and absence of Na with the protein. The docking scores are comparable along with the interacting amino acid residues forming hydrogen bond and pi-pi bonds (aromatic-aromatic interactions) (Table I). This suggests that sodium is important for binding of silibinin with the protein receptor (Fig:1).

TABLE I. IFD for standard and top dock score compound with DRD4 protein

Protein	**Compound**	**Docking Score (kcal/mol)**	**Glide energy (kcal/mol)**	**Interacting amino acid residues**
5WIU (without Na)	Standard (Nemonapride)	-9.390	-62.487	Asp115*, Leu187, Phe410#
	Silibinin	**-9.616**	**-71.992**	Ser196*, Arg186*, Glu95*, Phe411#
5WIV (with Na)	Standard (Nemonapride)	-10.407	-65.944	Asp115*, Ser196*, Leu187*, Phe410#
	Silibinin	**-11.482**	**-78.606**	Val87*, Ser196*, Ser197*, Leu187*, Cys185*, Arg186*, Phe411#

***Hydrogen bonds; #pi-pi bonds**

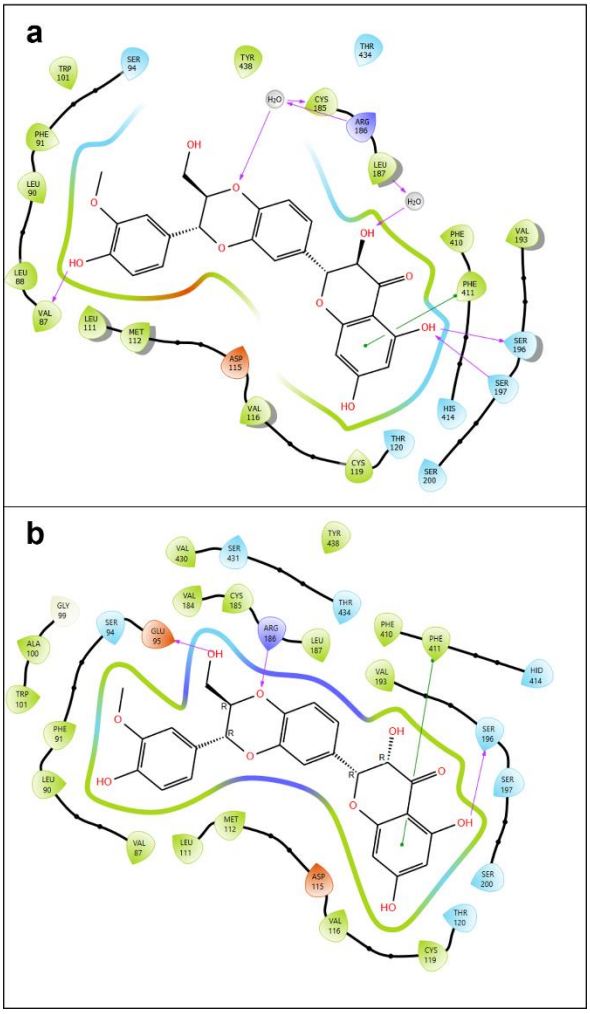


Fig.1. 2D ligand interaction a) silibinin with Na. b) silibinin without Na.

D. *Molecular dynamic simulations*

MD simulation for the docked complexes is performed to study the thermodynamic properties of the complex for 100ns simulation run. The root mean square deviation (RMSD) represents the stability of protein-ligand complex and indicates that simulation has equilibrated. RMSD values should be in the range of 1-4 Å. The results show that the values for silibinin is not within the permissible ranges (without sodium). This shows that the system is not equilibrated. But in case of simulations (with sodium), the system has equilibrated for both the standard and the silibinin (Fig: 2). Protein-ligand (PL) contacts show different types of interaction bonds formed by amino acid residues of protein with the ligand which are in the form of hydrogen bonds, hydrophobic bonds, ionic bonds and water bridges. “Fig. 3” shows the PL contact histogram along with its 2D ligand interaction diagram for silibinin with and without Na. Hydrogen bond forming residues are similar in both cases which are ASP115, CYS185, ARG186, LEU187, SER196, and HIS414; same as hydrophobic interactions formed are by PHE91, LEU11, MET112, CYS185, PHE410 and PHE411. Whereas SER200 forms multiple bonds as hydrogen, ionic and water bridges in presence of Na which is not case of absence of Na. It shows that due to the involvement of sodium, SER200 has multiple roles of biological importance in terms of silibinin binding with the target protein.

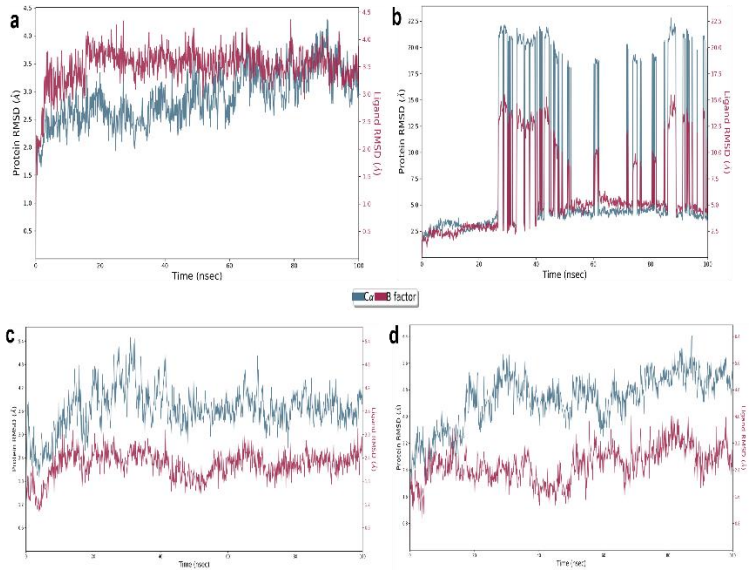


Fig. 2. Protein-ligand RMSD a) Standard without Na. b) silibinin without Na. c) Standard with Na. d) silibinin with Na.

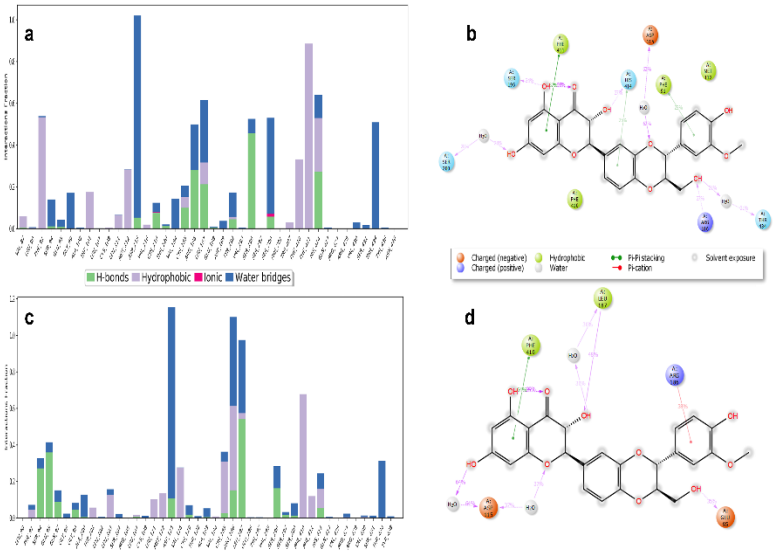


Fig.3. Protein-ligand interactions and ligand 2D interactions; (a,b) silibinin interaction in presence of Na. (c,d) silibinin interaction in absence of Na.

E. Binding free energy calculations

The results computed using MM-GBSA show that silibinin in presence of Na have higher binding affinity (-63.28 kcal/mol) in comparison without Na (-58.59 kcal/mol). For silibinin, the complex interactions like coulombic, van der waals and hydrogen bonds are most likely to be favored in the presence of Na with higher binding energies. Increased binding energy shows that silibinin compound in binding with protein in presence of Na bound is more energetically favorable and stable than other compounds (Table II). Per-residue decomposition results also supported the role and involvement of sodium which has -0.16 kcal/mol binding energy with the silibinin complex.

Table II. Binding free energy calculated using prime MM-GBSA

Protein	Compound	ΔG Bind (Kcal/mol)	ΔG Bind Coulomb (Kcal/mol)	ΔG Bind Hydrogen bond (Kcal/mol)	ΔG Bind van der Waals (Kcal/mol)
5WIU (without Na)	Standard	-41.60	-14.83	-1.22	-35.77
	Silibinin	-58.59	-18.81	-1.44	-44.40
5WIV (with Na)	Standard	-63.96	-6.80	-0.83	-47.98
	Silibinin	**-63.28**	**-19.97**	**-1.14**	**-49.07**

IV. CONCLUSION

The aim of the study was to find a potential phytochemical compound that could be considered as drug candidate for the treatment of ADHD. The target receptor selected through gene identification and PPI network construction was DRD4, a dopamine receptor. The crystal structure of dopamine receptor was availabe with and without sodium ion. Silibinin shows the potential binding capability with the protein in presence of sodium which is confirmed with several *in-silico* studies conducted. These studies led to the inflod of several amino acids and their potential role in the binding, which resulted in the favor that when sodium present with the protein, it provides a strong binding and stability of protein-ligand complex**.**The results are comparable with the nemonapride used as a standard under current investigation. The result needs to be validated further for future applications.

ACKNOWLEDGMENT

This research work has been carried out with the support of Department of biomedical engineering, Faculty of electrical engineering and communication, Brno university of Technology and Jjihomoravské centrum pro mezinárodní mobilitu (JCMM) which is a specialized non-profit interest association of legal entities operating in the South Moravian Region.

REFERENCES

[1] Danielson, M. L., Bitsko, R. H., Ghandour, R. M., Holbrook, J. R., Kogan, M. D., & Blumberg, S. J. (2018). Prevalence of parent-reported ADHD diagnosis and associated treatment among US children and adolescents, 2016. Journal of Clinical Child & Adolescent Psychology, 47(2), 199-212.

[2] Sundari, J., Amuthalakshmi, S., & Nalini, C. N. (2020). Review on analytical methods for quantification of ADHD drugs in human biological samples. Reviews in Analytical Chemistry, 39(1), 130-156.

[3] Rangel-Barajas, C., Coronel, I., & Florán, B. (2015). Dopamine receptors and neurodegeneration. Aging and disease, 6(5), 349.

[4] Ahn, J., Ahn, H. S., & Cheong, J. H. (2016). Natural product-derived treatments for attention-deficit/hyperactivity disorder: safety, efficacy, and therapeutic potential of combination therapy. Neural plasticity, 2016.

[5] Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., ... & Yaschenko, E. (2007).

Database resources of the national center for biotechnology information. Nucleic acids research, 36(suppl_1), D13-D21.

[6] Rappaport, N., Twik, M., Plaschkes, I., Nudel, R., Iny Stein, T., Levitt, J., Gershoni, M., Morrey, C.P., Safran, M. and Lancet, D., 2017. MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. Nucleic acids research, 45(D1), D877-D887.

[7] Safran, M., Dalah, I., Alexander, J., Rosen, N., Iny Stein, T., Shmoish, M., Nativ, N., Bahir, I., Doniger, T., Krug, H. and Sirota-Madi, A., 2010. GeneCards Version 3: the human gene integrator. Database, 2010.

[8] Piñero, J., Bravo, À., Queralt-Rosinach, N., Gutiérrez-Sacristán, A., Deu-Pons, J., Centeno, E., ... & Furlong, L. I. (2016). DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. Nucleic acids research, gkw943.

[9] Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., ... & Mering, C. V. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic acids research, 47(D1), D607-D613.

[10] Xiong, G., Wu, Z., Yi, J., Fu, L., Yang, Z., Hsieh, C., Yin, M., Zeng, X., Wu, C., Lu, A. and Chen, X., 2021. ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. Nucleic Acids Research, 49(W1), pp.W5-W14.

[11] Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. Journal of computer-aided molecular design, 27(3), 221-234.

[12] Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert opinion on drug discovery, 10(5), 449-461.

[13] Ptáček, R., Kuželová, H., & Stefano, G. B. (2011). Dopamine D4 receptor gene DRD4 and its association with psychiatric disorders. Medical science monitor: international medical journal of experimental and clinical research, 17(9), RA215.

CONGRATULATIONS

to all the Authors and Presenters at EEICT 2023!



The Leader in Semiconductor Test & Robotics

We bet you like semiconductors... Talk to us!
We're looking to expand our teams in Semiconductor Test and Robotics.

TERADYNE | LITEPOINT | UNIVERSAL ROBOTS | MiR



Engineer a Difference

Join our team! We are the test and automation experts behind the electronics you use every day. Our customers depend on us to get their product to market, fast. It's an exciting business and we're looking for talented people who share our passion and drive.

@Teradyneinc. linkedin.com/company/teradyne TERADYNE

Segmentation of optic disc and cup in retinal images using of deep learning approaches

Michal Nohel
Department of Biomedical Engineering
BUT, FEEC
Brno, Czech Republic
0000-0002-2679-2160

Radim Kolář
Department of Biomedical Engineering
BUT, FEEC
Brno, Czech Republic
0000-0002-0469-6397

***Abstract*—This paper presents a comparative analysis of optic disc and cup segmentation in retinal fundus images using two deep learning models: the classical U-net and its modified version, nnU-Net. The models were trained and tested on publicly available databases consisting of 1295 images for training and 555 images for testing. The results indicate that while nnU-Net demonstrated only slight improvements in disc segmentation on the test database, it significantly outperformed the U-net model in optical cup segmentation.**

***Index Terms*—optic disc, optic cup, segmentation, U-net, nnU-Net, machine learning, deep learning, fundus image**

I. Introduction

Retinal images play a crucial role in the diagnosis of various eye disorders. These digital images are acquired using a specialized camera called a fundus camera. The purpose of retina image analysis is to detect any abnormalities present on the retina, allowing the identification of diseases such as diabetic retinopathy, glaucoma, etc. As glaucoma progresses, an integral part of the diagnostic process is the accurate segmentation of the optic disc (OD, outer border) and the optic cup (OC, inner border) (see Fig.1). The cup-to-disc ratio is a critical parameter that is routinely assessed, with values of 0.3 or less indicating healthy eyes and higher values indicating potential abnormalities [1]. However, manual segmentation of the OD and OC is time-consuming and depends on the ophthalmologist's experiences; thus there is an urgent need to find automatic methods for segmenting these structures. [1]

In recent years, deep learning-based methods have gained widespread popularity in various fields, including medical imaging. However, these methods typically require a large amount of training data. Fortunately, there are several publicly available databases of retinal images with OD and OC annotations, such as [2] - [7]. It's important to note, however, that these databases may have various parameters, such as image size, the field of view, illumination, and ophthalmoscope type, which can affect the performance of deep learning algorithms.

For the purpose of this paper, machine learning models based on convolutional neural networks were trained on publicly available databases. Specifically, the original U-net [9] and its modification nnU-Net [10]. This paper presents a comparison of these models for segmenting OD and OC on different databases.

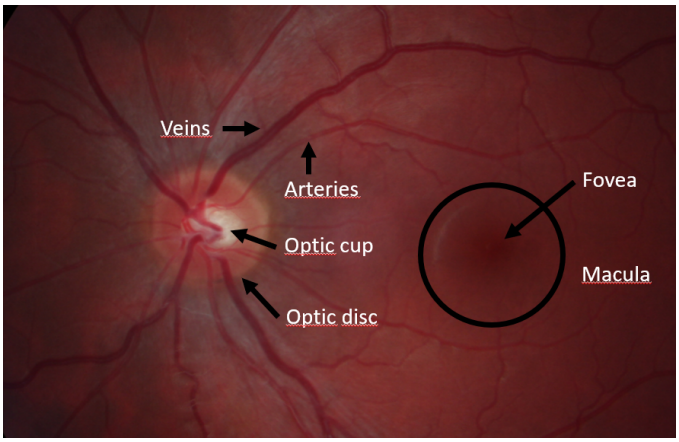


Fig. 1. An example of a retinal image from a fundus camera with basic retinal structures.

II. Materials and data

A. Data description

This paper addresses the segmentation of the optic disc and optic cup in retinal images (see Fig. 1). We used publicly available databases, namely Dristi-GS [2], HRF [3], REFUGE [4], RIGA [5], RIM-ONE [6], and UoA_DR [7]. Tab.I shows the parameters of the databases used in this study. To ensure consistency across the different databases, we performed preprocessing steps that included resampling all images to a uniform spatial resolution of 35 pixels per degree of field-of-view (FOV). As a result, we obtained a total of 2026 retinal images with the same spatial resolution, with the majority of images being from healthy subjects, and a smaller proportion of images being from patients with glaucoma or diabetic retinopathy. Each of these databases includes annotations for both the OD and OC.

From the entire available dataset, we selected only healthy subjects for subsequent preprocessing, training, and testing of convolutional neural networks. Our final dataset consisted of 1850 images. We divided our datasets into training and testing sets using two different approaches. For datasets that already had separated training and testing sets, we used those pre-defined splits. In cases where no such splits existed, we

TABLE I
PARAMETERS OF INDIVIDUAL DATABASES

Databases	Num. of img	FOV	Size of FOV[px]	no. of px per 1°
Dristi-GS	101	30°	1750×2048	68
HRF	45	60°	2336×3504	54
REFUGE-train	400	45°	2056×2124	47
REFUGE-test, valid	400,400	45°	1565×1565	35
RIGA	480	45°	1435×1435	32
RIM-ONE	169	20°	1424×1072	54
UoA_DR	200	45°	2056×2124	47

randomly selected 20% of the images from the database to form the testing set, while the remaining images were used for training. For consistency, we included any validation sets as a part of the training set. Thus, a total of 1295 images were used for training and 555 images for testing.

B. Data preprocessing

As illustrated in Fig.2, the images in our various databases exhibit significant variation in color and brightness. To remove these differences and ensure consistency across our dataset, we used preprocessing based on a modification of the approach in [8]. Specifically, we first performed brightness normalization using the formula:

$$I_{mod} = \frac{Img - G}{G} + 0.5 \tag{1}$$

where I_{mod} is the normalized image, Img is the original image and G is the blurred image obtained using a Gaussian filter with sigma = 50. Next, we converted the image to the HSV color space and applied adaptive equalization to the value component. Finally, we converted the image back to RGB color space. The effects of this normalization can be observed in Fig.3.

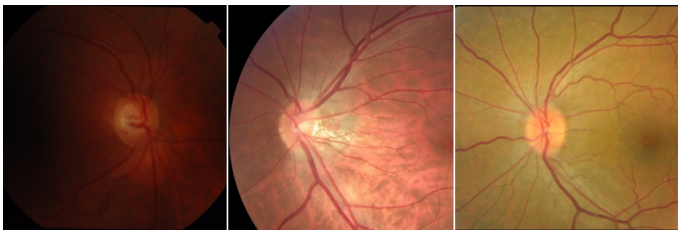


Fig. 2. Demonstration of brightness and color differences for different databases used. From the left image from the Dristi-GS, REFUGE, and RIGA databases.

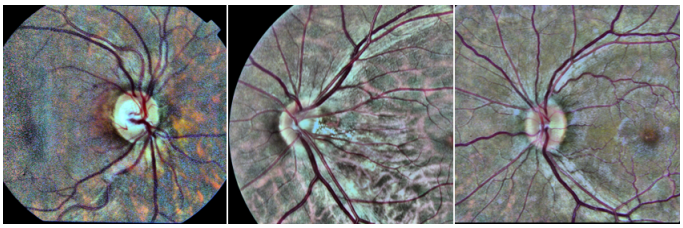


Fig. 3. This is an example of images with brightness modification from Fig.2.

III. MODEL IMPLEMENTATION

Here we describe two deep learning models, we used - the original U-net [9] and its modified version nnU-Net [10].

A. U-net description

The U-net architecture was a breakthrough for biomedical image segmentation because it allows the model to be trained even with a limited dataset, which is common in biomedical tasks. The U-net architecture consists of an encoder and a decoder and resembles the shape of the letter U. The encoder is where feature extraction occurs and is composed of three successive convolutional layers followed by a pooling layer. This sequence is repeated four times in this architecture. This is followed by the decoder, which is composed of an unpooling layer followed by convolutional layers. Between the encoder and the decoder, there are connections that carry information from a given encoder layer to the decoder, providing useful information. The decoder results in a segmented image with the same dimensions as the input image.

The problem with this network is that it requires the same size input every time, but after modification and resampling of the available databases, the images have different sizes. Therefore, the following procedure was used: In the first stage, optical disc detection was performed, which was published in [11]. This detection was performed using matched filtering. As the optical disc (OD) can be approximated by a 2D Gaussian function, a 2D Gaussian function was selected as the mask for matched filtering. The value of sigma, which was estimated from the average size of the OD, was equal to 30 since the average radius of the OD was 200 pixels. The result of the matched filtering obtained by the convolution between the image and the mask was used for subsequent detection, where the maximum was detected and identified as the position of the OD. To detect the OD during preprocessing, binary field-of-view (FOV) masks were generated and utilized. Since the OD center should not be at the edge of the FOV, the maxima in the image were detected by performing matched filtering only on the reduced FOV that was eroded by a disc structuring element with a neighborhood size of 40 pixels. Following the detection of the OD, a database of training images was created by cropping images to a size of 480×480 pixels, with the detected OD positioned at the center of each image.

A convolutional neural network inspired by the U-net architecture was trained on the created database of cropped optical disc images from scratch. The number of filters in each layer

was chosen as 8, 16, 32, 64, and 128. The initialization of the weights was performed using Xavier initialization. The size of the filters in each layer was set to 3×3. The ReLU activation function was used in the activation layer. Transposed convolution followed by the ReLU activation function was used for oversampling in the decoder part of the network. In the output of the network, the sigmoid activation function was applied followed by thresholding on the binary image.

The network was trained using a training database that consisted of 1295 images. To augment the data, random cropping was performed to achieve a size of 448×448 px, and random mirroring and rotation by multiples of 90 degrees were also utilized. The training was performed in Python using the PyTorch library. The learning parameters were optimized to achieve maximum accuracy, with the learning rate set at 0.001 and a decreasing learning rate scheduled every 7 epochs with a gamma parameter of 0.1. The total number of epochs used for training was 30, with a batch size of 12. The optimization algorithm used was Adam, and the error function was based on the Dice coefficient. The model was trained on a computer equipped with an NVIDIA GeForce GTX 1650 graphics card and an 11th Gen Intel(R) Core(TM) i5-11400 @2.60 GHz processor and 15.84 GB of RAM.

B. nnU-net description

nnU-Net is a deep learning architecture for medical image segmentation that is based on the U-Net framework. Originally developed for 3D volumetric medical data, nnU-Net can also be adapted to 2D image segmentation. It has been shown to achieve state-of-the-art results on multiple medical imaging datasets and is widely used in medical image analysis. What sets nnU-Net apart from other architectures is that it includes a framework that automatically designs the optimal U-Net architecture and parameters for a given training dataset, including batch size, patch size, criterion function, learning rate, and more. The framework also includes preprocessing of the training data and its execution.

Compared to the classical U-Net, nnU-Net has several advantages, including the ability to handle input images of different sizes, optimize the pre-processing step, and train the U-net model. This eliminates the need for disc detection and image cropping during training, which is a significant advantage. It is worth noting, however, that nnU-Net requires data in NIfTI format as input. Therefore, to proceed with the planning and preprocessing steps for the modified 2D fundus images, they had to be converted to NIfTI format. After the conversion, nnU-Net was utilized to automatically design the optimal architecture and learning parameters, with a learning rate set at 0.01 and gradually reduced during training. The batch size was set at 2, with a kernel size of 3x3, and a total of 1000 learning epochs were performed. To validate the model's performance, 5-fold cross-validation was utilized.

One potential limitation of nnU-Net is that it is only compatible with Linux systems, which may be a barrier for some users. Consequently, to train the model, the Metacenter was used, which had 12 cores, 32GB of RAM, and a dedicated graphics card.

IV. METRICS

The following metrics were used to evaluate optical disc and cup segmentation. The first metric is Dice, which can be obtained by using the formula:

$$Dice(P,T) \quad = \quad 2\frac{|P \cap T|}{|P| + |T|}, \tag{2}$$

where P represents predicted pixel value, T correct values – ground truth. The final dice score has no unit [–] and ranges between 0 and 1.

When applied to Boolean data, using the definition of true positive (TP), false positive (FP), and false negative (FN), it can be written as:

$$Dice \quad = \quad \frac{2TP}{2TP + FP + FN} \tag{3}$$

The second evaluation parameter used is the relative area error expressed as a percentage. This is the difference between the area of the ground truth (GT) contour and the contour predicted by the convolutional neural network, and it can be calculated using the following formula:

$$E_a \quad = \quad 100\frac{|S_{GT} - S_{pred}|}{|S_{GT}|}, \tag{4}$$

where E_a represent relative area error S_{GT} represents area of ground truth segmentation mask and S_{pred} represents area of predicted segmentation mask.

The third evaluation parameter used is the relative value of the mean of the difference distances between points x_i and y_i of the GT contour and points x_i and y_i of the predicted contour, normalized by the length of the GT contour and can be calculated using the following formula:

$$E_d \quad = \quad 100\frac{1}{M}\sum_{i=1}^{M}\frac{\sqrt{(x_i^{GT} - x_i^{pr})^2 + (y_i^{GT} - y_i^{pr})^2}}{R_i}, \tag{5}$$

where E_d represents the error of distance, M represents the number of contour points from which the error was calculated, x_i^{GT} y_i^{GT} represents the points of ground truth contour and x_i^{pr}, y_i^{pr} points of the predicted contour, R_i represent the average distance of the GT contour from the center of contour (centroid).

V. RESULTS AND DISCUSSION

To evaluate the trained models, we tested them on a test database consisting of 555 images. Fig.4 shows an example of the contour masks for optic disc and optic cup from the original images for ground truth, U-net segmentation, and nnU-Net segmentation. Overall, both methods performed well, but there were some small differences between them. The contour of the segmentation using classical U-net was occasionally less smooth and distorted by the presence of

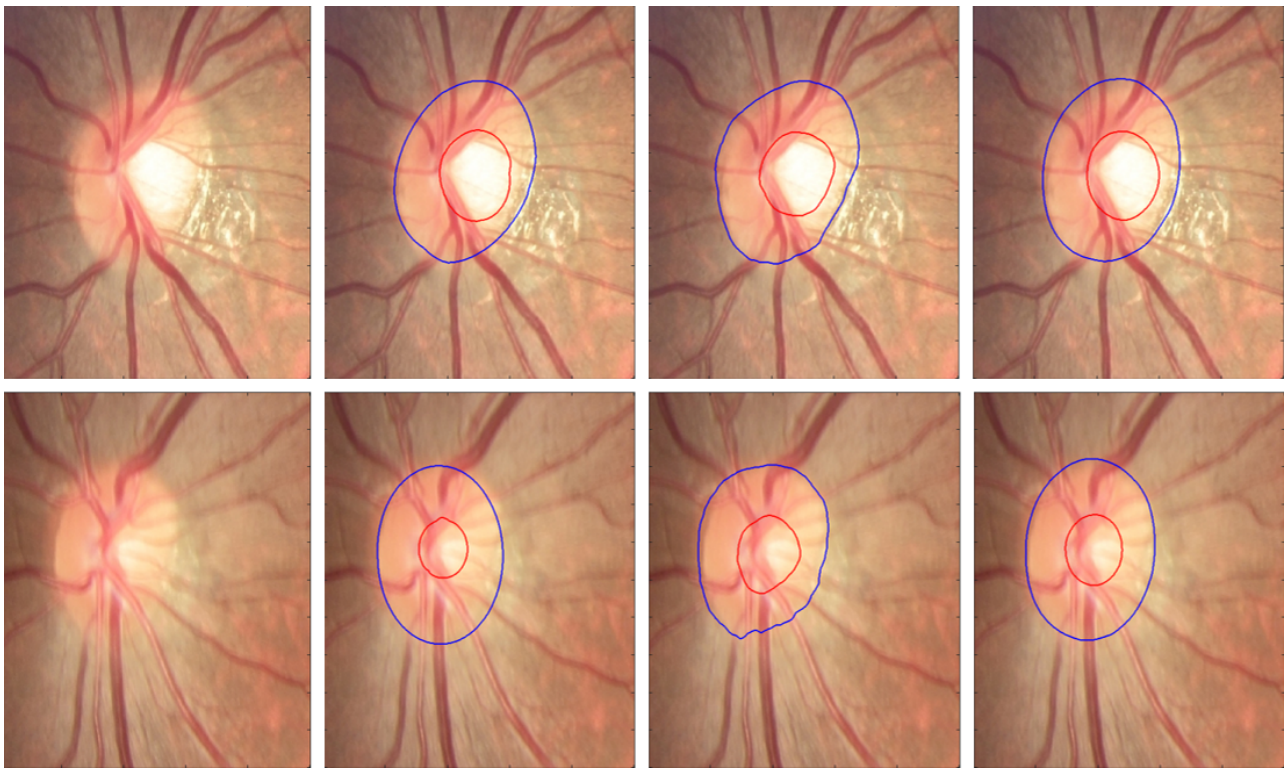


Fig. 4. Example of optic disc and cup segmentation results for two different images. The images from left to right represent the original image, the ground truth contour, the output of classical U-net, and the output from nnU-Net. The boundary of the optic disc is highlighted in blue, while the boundary of the optic cup is highlighted in red.

blood vessels. In contrast, the nnU-Net segmentation produced much smoother results.

To quantify the accuracy of the segmentation, we calculated evaluation metrics on the test data. Tab.II summarizes the segmentation results for optic disc segmentation, while Tab.III summarizes the evaluation for optical cup segmentation.

Tab.II summarizes the segmentation results for the optical disc, indicating that nnU-Net outperforms classical U-net in all metrics. Notably, the Dice coefficient is exceptionally high, with an average value of 0.947 for U-net and 0.959 for nnU-Net. Both methods also exhibit relatively low standard deviation of the Dice coefficient. The area error E_a was quite low for the OD segmentation for both methods. The statistics for the E_d length error are also much better for nnU-Net

Tab.III summarizes the segmentation results for the optic cup. Optic cup segmentation is a challenging task, even for experienced ophthalmologists, making it a difficult challenge for the models to perform well. The results indicate that nnU-Net is better suited for optical cup segmentation on the available data, as it achieves a higher average Dice coefficient with a smaller standard deviation. The other evaluation metrics also demonstrate that nnU-Net achieves a much higher success rate in comparison to classical U-net.

To examine the results more closely and identify outliers, we constructed box-and-whisker plots for the Dice coefficient results for both optic disc segmentation (see Fig.5) and cup segmentation (see Fig.6).

TABLE II
RESULTS FOR OPTIC DISC

Methods	U-net	nnU-net
Average Dice	0.947	0.959
Std. Dice	0.033	0.019
Average Ea [%]	8.067	6.120
Average absolute Ed [px]	5.950	4.768
Std. absolute Ed [px]	3.499	2.207
Average relative Ed [%]	5.459	4.187

TABLE III
RESULTS FOR OPTIC CUP

Methods	U-net	nnU-net
Average Dice	0.829	0.885
Std. Dice	0.126	0.068
Average Ea [%]	26.659	16.607
Average absolute Ed [px]	9.752	6.817
Std. absolute Ed [px	8.238	5.472
Average relative Ed [%]	18.116	12.659

Fig.5 demonstrates that the boxplot for nnU-Net contains fewer outliers and is more compact than the boxplot for classical U-net. Additionally, the boxplot for U-net also includes some images that were noticeably poorly segmented.

Fig.6 highlights the challenges in segmenting the optical cup and demonstrates that nnU-Net achieves better results. The boxplot for nnU-Net is more compact and contains fewer

outliers, while classical U-net segmentation failed to segment the optical cup in some images.

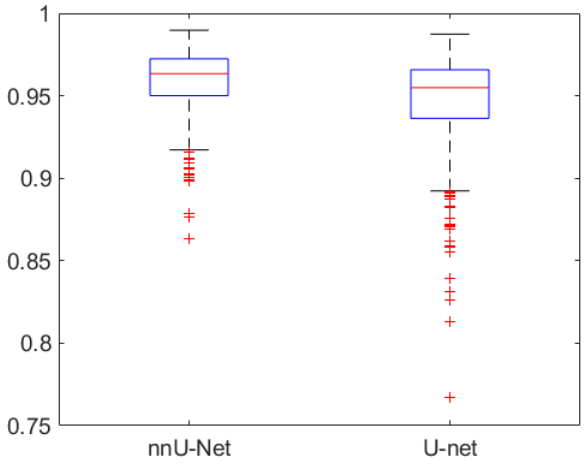


Fig. 5. Box-and-whisker plot of Dice coefficient for optic disc segmentation. On the left is the boxplot for nnU-Net and on the right is the boxplot for U-Net. The red line indicates the median value in the dataset, the rectangle box indicates the interquartile range, the black lines indicate minimum and maximum, and red marks are outliers in the dataset.

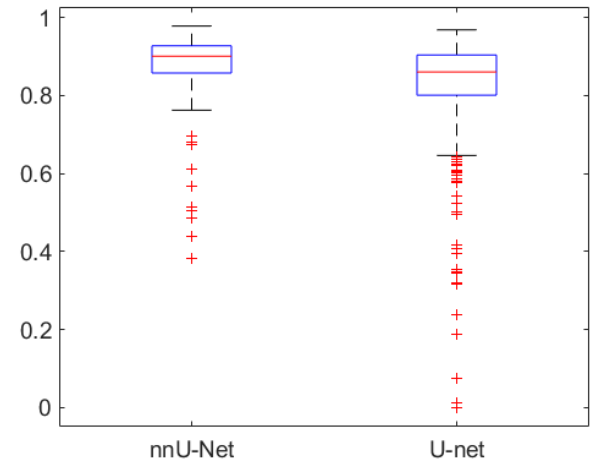


Fig. 6. Box-and-whisker plot of Dice coefficient for optic cup segmentation. On the left is the boxplot for nnU-Net and on the right is the boxplot for U-Net. The red line indicates the median value in the dataset, the rectangle box indicates the interquartile range, the black lines indicate minimum and maximum, and red marks are outliers in the dataset.

VI. Conclusion

This study aimed to compare the effectiveness of deep learning approaches for optical disc and cup segmentation. Two convolutional neural network models, the U-net and its modified version, nnU-Net, were trained on publicly available databases, which were adjusted for the same spatial resolution and normalized. The methodology for each approach was presented, and their advantages and disadvantages were summarized. The trained models were tested on a separate test database, and the results were evaluated. Overall, the nnU-Net model outperformed the U-net model on every metric for both optic disc and cup segmentation.

Acknowledgment

This work was supported by project no. 21-18578S from Czech Science Foundation.

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. Computational resources were supplied by the Ministry of Education, Youth and Sports of the Czech Republic under the Projects CESNET (Project No. LM2015042) and CERIT-Scientific Cloud (Project No. LM2015085) provided within the program Projects of Large Research, Development and Innovations Infrastructures.

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.

[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. Odstrcilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Cernosek, O. Svoboda, and E. Angelopoulou, "Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database", *IET Image Processing*, vol. 7, no. 4, pp. 373-383, 2013.

[4] 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.

[5] 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. doi:10.1117/12.2293584

[6] F. Fumero, S. Alayon, J. L. Sanchez, J. Sigut, and M. Gonzalez-Hernandez, "RIM-ONE: An open retinal image database for optic nerve evaluation", *2011 24th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 1-6, 2011.

[7] 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.

[8] R. Chrástek, M. Wolf, K. Donath, H. Niemann, D. Paulus, T. Hothorn, B. Lausen, R. Lämmer, C. Y. Mardin, and G. Michelson, "Automated segmentation of the optic nerve head for diagnosis of glaucoma", *Medical Image Analysis*, vol. 9, no. 4, pp. 297-314, 2005.

[9] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation", *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015*, pp. 234-241, 2015.

[10] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "NnU-Net: a self-configuring method for deep learning-based biomedical image segmentation", *Nature Methods*, vol. 18, no. 2, pp. 203-211, 2021.

[11] M. Nohel and R. Kolář, "Detection of optic disc in retinal fundus image", in *Proceedings I of the 28th Conference STUDENT EEICT 2022 General Papers*, 2022, pp. 239-242.

Comparing Hough transform to other automated methods of feature extraction and matching for document correction

1st Michal Gavenciak
Department of telecommunications
Brno University of Technology
Brno, Czech Republic
0000-0001-8835-1056

***Abstract*—In some specific cases of cognitive task assessment in children with developmental dysgraphia, experts need to apply advanced modelling and machine learning techniques on a combination of scanned image data and online drawing signals. Nevertheless, in the first step, they usually need to remove external imperfections in the scanned data (such as the rotation of the input image) so that they could be easily registered with a product recorded by a digitizing tablet. The aim of this work is to introduce a method that could be used for such a purpose. We briefly describe and compare a few techniques commonly used for feature point detection. Next, we describe an algorithm based on the Hough transform developed specifically for use with our dataset and compare its effectivity with more generalistic tools. Detected feature points are employed to estimate a homography matrix and to correct the scanned images. The proposed approach was successful in 89 % of cases.**

***Index Terms*—Developmental Dysgraphia, Hough transform, Feature detection, Handwriting analysis**

I. Introduction

The majority of research in the domain of automated developmental dysgraphia (DD) assessment is focusing on the analysis of on-online handwriting data, often using feature extraction methods in concert with machine learning, statistical analysis and mathematical modelling [1]–[3]. An alternative approach is to consider the use of image processing tools to directly evaluate the handwriting product, rather than the process [4]. To achieve the best results in training a model suited to recognize or evaluate certain alterations, the input data must be cleared of any external influences as much as possible; otherwise, a model might be trying to recognize these external influences, which are, however, unrelated to the actual data. One of such external flaws might be imperfections that occur during data acquisition, such as flawed image scans. In this article, we will explain the correction process of a database with imperfect scans, where most images are slightly rotated, and compare the success rate of multiple methods.

This study was supported by a project of the Technology Agency of the Czech Republic no. TL03000287 (Software for advanced diagnosis of graphomotor disabilities).

II. Methodology

This paper will describe the process of image correction for the specific purpose given by the imperfections present in the database. The cornerstone of the necessary correction is a homography matrix. Homography matrices are often used within the context of computer vision, where one or more cameras are looking at the same object. The homography matrix then describes the relationship between the two cameras and the transformation necessary to match their perspective. In general terms, the two images are scanned for feature points and their position. After correlating the feature points, a projection matrix is obtained which describes the transformation necessary to align the two images. We have compared multiple feature extracting algorithms, that despite some key differences, can be roughly described by the flowchart in Fig. 1.

Please note that the process nodes give a very generic description of each step. Every step of the algorithm will be described in more detail (for each of the feature extraction processes used) in the subsections of this chapter.

While in our case, we are not speaking about a situation where two different cameras are looking at the same object, it is a similar one. If we consider the scanned document as an object perceived by one observer and then the empty protocol template (see Fig. 2) as the same object perceived by a second observer, we can then determine the transformation (and the matrix by which it is described) that was applied on the scanned paper during handling. Applying the inverse transformation to the flawed document will serve as a correction.

A. Database

The database used as the basis for this article was gathered in elementary schools throughout the Czech Republic in grades ranging from 1st to 5th grade. All subjects were asked to fill out a protocol designed to evaluate DD. The protocol was completed using a Wacom Intuos Pro L (PHT-80) digitizing tablet in conjunction with a Wacom inking pen. The protocol was printed on a sheet of paper that was fitted to the surface of the tablet. The combination of a typical sheet of paper and the inking pen allowed for the usual visual and haptic feedback of

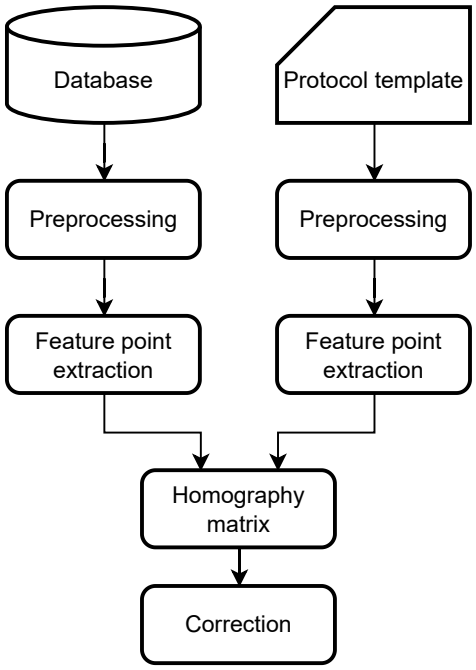


Fig. 1: Outline of the figure correction process.

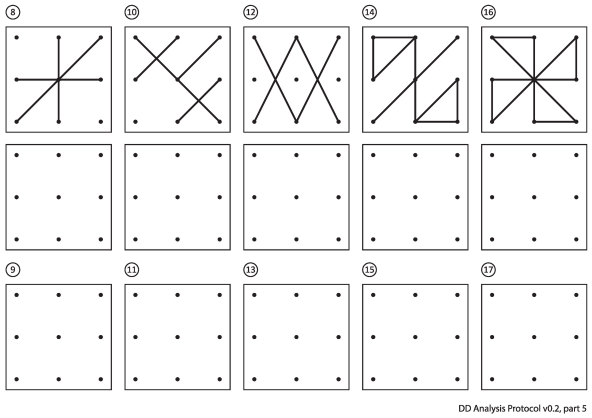


Fig. 2: The empty protocol to be filled out by a subject. The top row consists of exercises for the subjects to copy into the empty cells beneath, then once again with the template figure covered.

drawing/writing and also allowed us to inspect the product of the handwriting in real-time and also later in a more analytic fashion.

Part of the protocol which cannot be completely evaluated using only the on-line handwriting data was scanned using generic office-grade scanner machines, which resulted in 416 image files, one for each subject. Most of the scanned files show signs of imperfect scanning, mostly slight rotation up to 1° and variance in dimensions. All scanned files are matched in resolution (3507×2480 px) with an average size of 2.05 MB.

The protocol the subjects filled out can be seen in Fig. 2. An example of a scanned protocol can be seen in Fig. 3. Parents of all children signed an informed consent form. The study was approved by the Ethics Committee of Masaryk University. Moreover, we followed the Ethical Principles of Psychologists and Code of Conduct released by the American Psychological Association (see https://www.apa.org/ethics/code/).

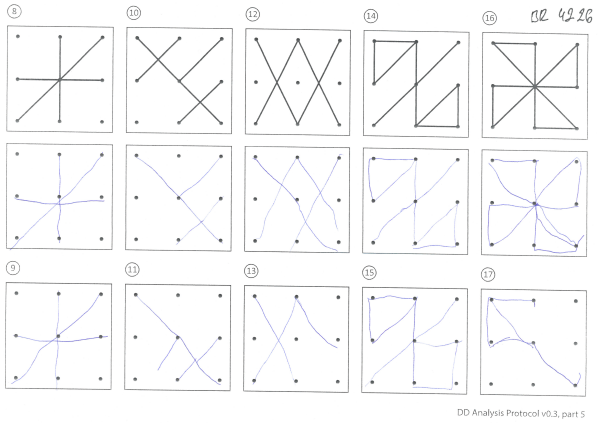


Fig. 3: Scanned protocol with a visible rotation of approximately 1°.

B. Feature point extraction and matching

Related to the task of image correction, features are generally those parts of the image that are in some way distinct from other parts of the pictures (e.g. textures, colours) or are the points of abrupt change, for example between two sections of different colours or a change of direction of an edge.

This section will briefly describe the different methods of feature extraction and matching tested in this paper.

1) Harris corner detector: Harris corner detector (HCT) is an algorithm designed to detect corners in an image [5]. In broad terms, it works by taking a derivative of the image in both the horizontal and vertical plane, creating a structure tensor M. Then the Harris value R is calculated, using:

$$R = \det(M) - k(\text{trace}(M))^2, \tag{1}$$

where k is a changeable parameter in HCT, which is generally set in the range of [0.04; 0.06]. We have tried values within this range with a step of 0.002.

2) Hough transform: Hough transform (HT) is an image processing technique used to isolate specific shapes within an image [6]. Typically, HT can recognise shapes that can be parametrically described, e.g. lines or circles. Given the structure of the analyzed image(s), with an abundance of such shapes (each individual exercise is framed in a square field with thick borders, which should be detectable as lines and exercises are also indexed with a number in a circle of a set diameter), HT seems to be a good candidate in either setting.

Considering the uniform spacing between the circled exercise indices, we concluded that a variant of the Hough transform for detecting circles has a higher potential for success. A version of the Hough transform specifically tuned to find circles in an image is often referred to as the hough circle detector or the circle hough transform (CHT). As mentioned for HT, CHT is searching for a circle as a shape that satisfies the equation:

$$(x-a)^2 + (y-b)^2 = r^2, \tag{2}$$

where a and b are the positional coordinates of the centre and r is the radius. Since we know the radius of the circles we are searching (36 px), the computational difficulty is vastly reduced.

Algorithmically CHT first converts the image to greyscale (in our case we can directly convert to an inverted binary image for better edge detection), and then detects edges (for example with a Canny Edge detector). An accumulator matrix in the dimensions of the picture is created. For each edge pixel, the accumulator matrix is then raised by one on every position (pixel) that fulfils the properties given by equation (2). This can be visualised as drawing thin circles of the diameter r around every edge pixel. Spots where many of such circles cross can be suspected as a centre of a circle. See Fig. 4 for a visualisation. If a picture contains only one circle, its centre is then given by the coordinates of the maximum value in the accumulator. As we expected to find more than one circle, we have applied a 0.79 relative threshold to the accumulator, which was an empirically determined value. Located circles can be seen in Fig. 5.

After applying this process to both the protocol template and a corrected scan, we presume to find 10 circles in each. To construct the homography matrix we select the corner circles (forming a rectangle, in Fig. 2, they would be located in the circles around the indices 8, 9, 16, and 17) of each image. These points across the figures are then considered as paired feature points for the process of homography estimation.

3) SURF: SURF (Speeded Up Robust Features) is a state-of-the-art computer vision algorithm often used in robotics and automation, among others, to extract features from multiple images to correlate two or more different cameras, but also for object recognition and object tracking [7]. As this method is more advanced than the two previously mentioned, it is beyond the scope of this article to explain it in detail. In short, SURF detects points of high variance of the brightness gradient using a hessian matrix. Then, SURF extracts descriptors of such identified spots, which are vectors of numbers that describe the area around key point (orientation, fluctuation of brightness).

4) Feature matching and homography estimation: For conventional feature matching, we have tried using two different algorithms. K nearest neighbours with k in the range of [5; 20] and Correlation matching with window sizes in the range of [5;20]. Homography was then estimated using RANSAC homography estimator [8].

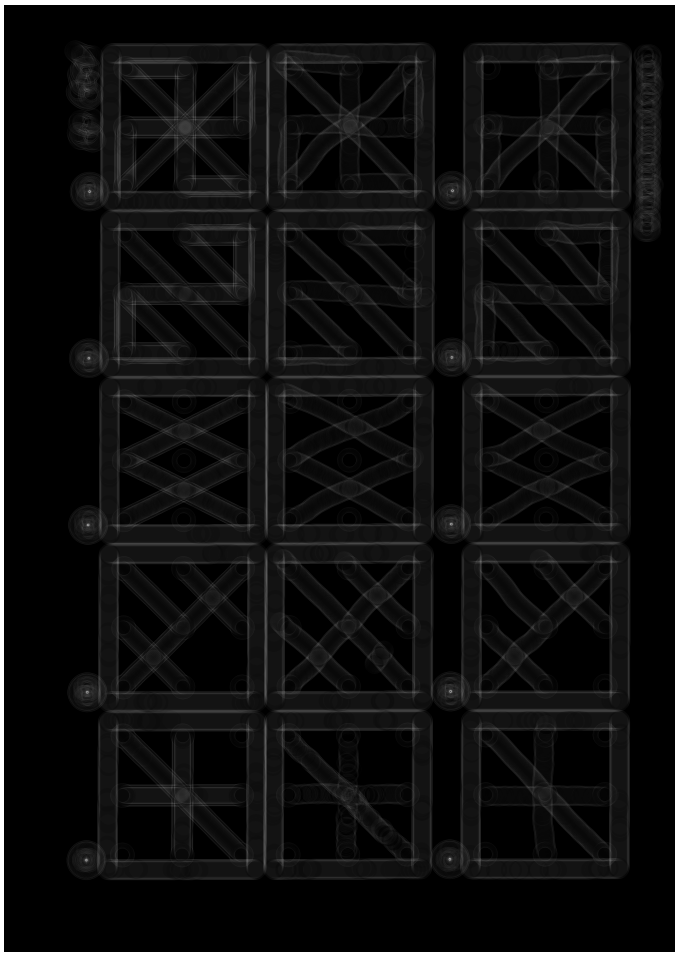


Fig. 4: Visualisation of the accumulator during CHT. The brightness of a pixel correlates with the weight of the accumulator cell. Higher values mean a higher likelihood of a circle.

C. Evaluation metric

While evaluating a corrected image is sufficiently possible with the naked eye, checking the whole dataset for each algorithm and parameter setting would be very time-consuming. We have decided to employ the image subtraction method to evaluate whether a correction was successful enough. Since what we are expecting as an output from the correction algorithms is compensation of translation, rotation, and scale, we expect that an ideal correction results in two perfectly aligned images. Subtracting one image from the other, pixel-wise should result in a blank image. In a situation where the alignment is not pixel perfect, some pixels will be left with a non-zero value. Counting all these pixels (in binarised image format) can then be interpreted as a very simplistic metric of the similarity of the images.

To use this as an evaluation metric, we first binarise the output and template image and take a part of the image, that is expected to have the same content in both images, i.e. the first line of template exercises (see Fig. 2, which should be

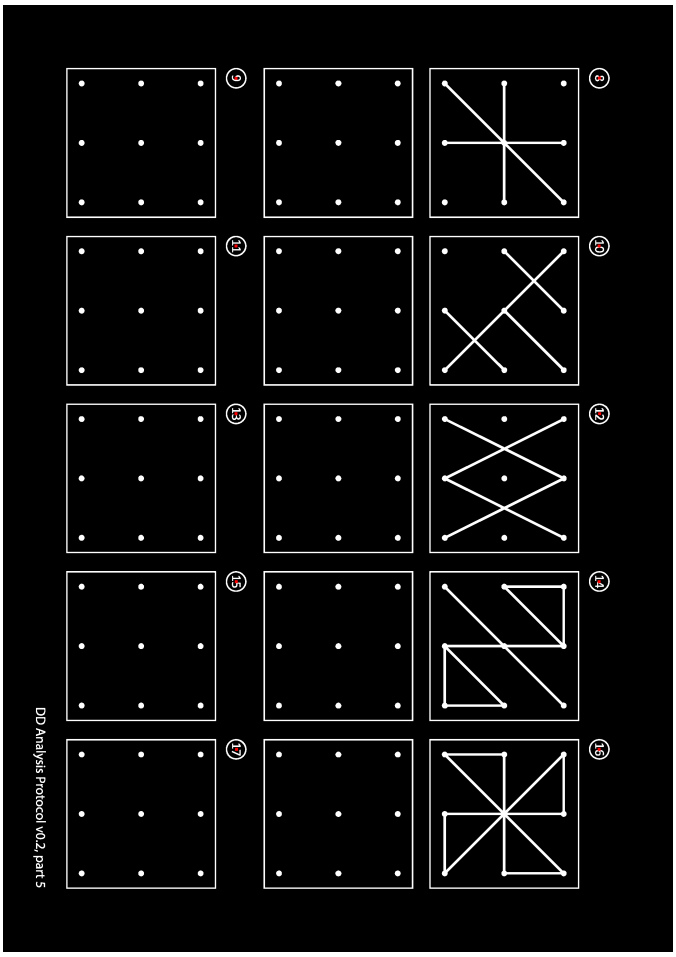


Fig. 5: All detected circles in the protocol template (red).

absent of any input from the subjects or evaluator notes). We then invert both images, so that data content is represented with ones and empty space with zeros (assuming typical image representation of black as zeros and white as ones) and since the image matrix is very sparse, we dilate both images with two iterations using a 3-by-3 kernel. Subtracting the output image from the template image then typically produces an image with some non-zero pixels left over. Their amount (ΔP) is inversely correlated to the success of the correction.

Then, by manually assessing several corrected images both with successful correction (showing no or minimal rotation and translation) and unsuccessful correction we iteratively determined a threshold for the ΔP value, which we then used as a success criterium for evaluation of the algorithms.

III. Results

Hough transform was able to successfully correct 89 % of the database (failed with 45 images). It was by far the most successful algorithm for feature detection. All the other feature detection algorithms in combination with the tested feature matchers were unable to consistently correct a substantial amount of images. Despite the large range of tested parameter values, we were not able to find a setting combination that would reach at least a 10 % correction success rate. Examples of failed corrections can be seen in Fig. 6 and 7.

Taking a deeper look at the process of feature detection using the conventional methods, the number of detected feature points was usually in the order of 10^3, when using commonly recommended settings. Tuning the parameters of the algorithms to be less sensitive usually resulted in too few features being detected and thus the inability to determine the homography matrix. Setting the parameters to the lowest possible sensitivity for most of the images having at least 10 feature points detected did not result in an increase in the success rate.

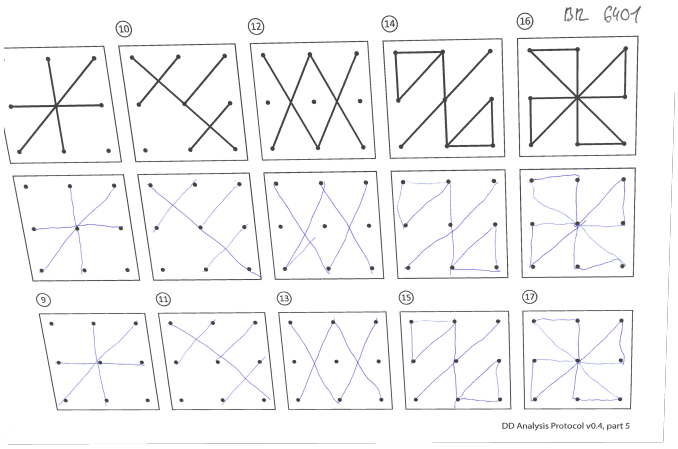


Fig. 6: An example of failed correction, which shows an obvious projective transformation along with a translation to the side, which resulted in some of the image being cropped. However, the original protocol is still recognisable.

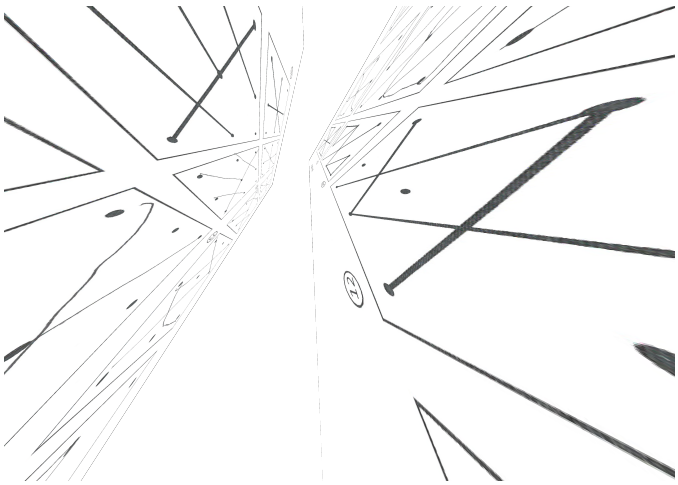


Fig. 7: An example of failed correction with a high error, which shows signs of perspective warping.

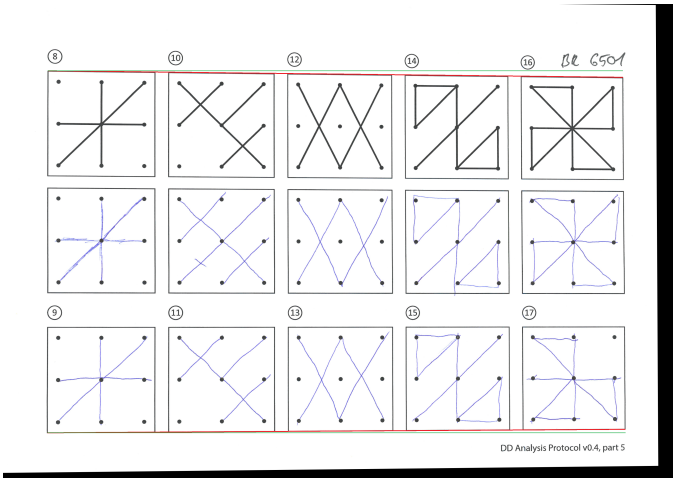


Fig. 8: Scan with a failed correction with a minimal error. While this particular scan was originally rotated, it did not show signs of a projective transformation, which must have been added by the correction process. Green lines show the expected (horizontal) outline of the exercise boxes, red shows the actual outline. The black frame showing around the bottom and right edge is an artifact of the correction, specifically, the translation of the image.

IV. Discussion and conclusion

With such a high failure rate of the conventional algorithms, it is necessary to state as a limitation of this study, that there is a possibility, that we were unable to find the optimal combination of settings to achieve an acceptable result.

The amount of detected feature points using the conventional algorithms might point to the fact, that the information in the images is too sparse. Since these detectors are mostly used for processing photographies or other photorealistic media, we suspect, that since most of our input is white (or black, for inversed binary image), the algorithms are prone to over- or under-detection of features with high volatility between different inputs.

It might be interesting to note, that it was possible to find a setting for most of the algorithms, where their correction error was small for individual images in the dataset. The example in Fig. 8 was distorted by the algorithm, however, the error is hard to see with the naked eye and might suggest, that this setting might have been close to a much more successful one.

We have also tested a variant of the algorithm with only rigid transformations taken into account, which reduced the complexity of the whole process. While rigid transformations did not produce as erroneous output as the homography approach when they failed, the approach did not produce a larger amount of successfully corrected images. In fact, both rigid transforms and projection transforms were successful in correcting the same images. A possible interpretation of this outcome is that it is usually the feature-detection algorithm that fails and not the transformation itself.

In conclusion, we have reached the original goal of image correction with a reasonable success rate, with an algorithm specifically tailored to our usecase. This points in the direction, that while generalistic approaches that are very successful in their own domain might not be as successful in slightly different jobs and it is still convenient and necessary to create usecase specific algorithms, which will reach much higher success rates, albeit in a very narrow, specialized area.

Acknowledgments

The author would like to thank his supervisor Ing. Jiří Mekyska, Ph.D. for his input and guidance.

References

[1] K. Šafárová, J. Mekyska, and V. Zvončák, "Developmental dysgraphia: A new approach to diagnosis," *The International Journal of Assessment and Evaluation*, vol. 28, no. 1, pp. 143–160, 2021. [Online]. Available: https://doi.org/10.18848/2327-7920/cgp/v28i01/143-160

[2] Z. Galaz, J. Mucha, V. Zvoncak, J. Mekyska, Z. Smekal, K. Safarova, A. Ondrackova, T. Urbanek, J. M. Havigerova, J. Bednarova, and M. Faundez-Zanuy, "Advanced parametrization of graphomotor difficulties in school-aged children," *IEEE Access*, vol. 8, pp. 112 883–112 897, 2020. [Online]. Available: https://doi.org/10.1109/access.2020.3003214

[3] T. Asselborn, M. Chapatte, and P. Dillenbourg, "Extending the spectrum of dysgraphia: A data driven strategy to estimate handwriting quality," *Scientific Reports*, vol. 10, no. 1, p. 3140, Feb 2020. [Online]. Available: https://doi.org/10.1038/s41598-020-60011-8

[4] Z. Galaz, P. Drotar, J. Mekyska, M. Gazda, J. Mucha, V. Zvoncak, Z. Smekal, M. Faundez-Zanuy, R. Castrillon, J. R. Orozco-Arroyave, S. Rapcsak, T. Kincses, L. Brabenec, and I. Rektorova, "Comparison of CNN-learned vs. handcrafted features for detection of Parkinson's disease dysgraphia in a multilingual dataset," *Frontiers in Neuroinformatics*, vol. 16, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fninf.2022.877139

[5] C. G. Harris and M. J. Stephens, "A combined corner and edge detector," in *Alvey Vision Conference*, 1988.

[6] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Commun. ACM*, vol. 15, no. 1, p. 11–15, jan 1972. [Online]. Available: https://doi.org/10.1145/361237.361242

[7] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Computer Vision – ECCV 2006*, A. Leonardis, H. Bischof, and A. Pinz, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 404–417.

[8] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, p. 381–395, jun 1981. [Online]. Available: https://doi.org/10.1145/358669.358692

DNA Methylation Detection With Pacific Biosciences sequencing and Oxford Nanopore sequencing Technology

Mohammad Umair
Department of Biomedical Engineering,
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
253834@vutbr.cz

Karel Sedlář
Department of Biomedical Engineering,
Faculty of Electrical Engineering and Communication
Brno University of Technology,
Brno, Czech republic
sedlar@vut.cz

***Abstract*—DNA methylation plays a crucial role in epigenetic modification which regulates gene expression and genome stability. Application of Pacific Biosciences sequencing technology (PacBio) and Oxford Nanopore sequencing technology (ONT) has shown promising approaches for detecting single base modification for DNA methylation. In this article, we analysed the base modification detection on ONT sequences and then compared it with PacBio analysed data. We found that both Nanopore and PacBio methods were effective but showed variations in their result. We analyzed the methylation pattern in *Clostridium beijerinckii* NRRL B-598 and found that the ONT data in combination with tombo tool can help detect modification at low coverage sites while avoiding false reading at these areas. Our study provide insight into these two approaches of DNA methylation detection, which will help in guiding the choices of appropriate method for specific research purposes.**

Keywords— DNA Methylation, Clostridium beijerinckii, ONT, PacBio, 6mA

I. INTRODUCTION

DNA methylation is a significant epigenetic change which plays an important role in gene expression. DNA Methylation occurs when a methyl group added is to nucleotide bases of DNA, which can change the shape, stability, and interactions of DNA with proteins as well as its chromosomal structure [1]. In contrast to DNA sequence, genomic methylation patterns are largely reprogrammed during embryogenesis, with changes accruing during mitosis of the cell cycle [2]. Depending on the cellular environment, methylation can be both steady and unpredictable and is frequently linked to transcriptional repression. It is essential for growth, cellular differentiation, and cellular identity preservation. Numerous illnesses, including cancer and neurodegenerative ailments, have been related to changes in DNA methylation patterns and its study has the ability to shed new light on a variety of biological processes and advance our knowledge of illness pathogenesis and treatment [3-6].

DNA methylation of bacteria occurs at various sequence contexts and plays an important functional role in gene regulation, host defence system and cell cycle etc. Methylation is regarded as a critical epigenetic regulator in bacteria because these species lack essential components like nucleosomes and histone, which modify DNA in eukaryotic organisms [7]. Methylation in bacteria are majorly caused by methyltransferase enzyme which are frequently found paired with restriction enzymes and genes encoding the restriction enzyme and the methyltransferase are often located next to each other and form a unit called a restriction-modification (RM) system [8]. The process of methylation helps in protecting the host genome from the restriction activity by the corresponding restriction enzyme. The efficiency of recognition of RM system can be increased with detection of methylation modification as the presence or absence of specific RM systems can be correlated with the methylation patterns of the DNA sequences While methylation studies have mainly focused on eukaryotes, detection of methylated bacterial bases has been a long-standing challenge. Detection of 5mC modification through Bisulfite genome sequencing was mostly used in eukaryotes and their use in detecting bacterial methylation has been limited as the majority of bacterial DNA methylation occurred in 6mA form. The recent advancement in sequencing technology, including Pacific Biosciences 'Single molecule real time (SMRT) and Oxford Nanopore technologies' nanopore approaches, are capable of detecting methylation signals and have allowed characterization of prokaryotic genome-wide epigenetic regulation [9]. Among the 3rd gen methylation detection technologies, DeepSignal, Deepmod, SMALR and Tombo are all powerful solutions. Tombo is particularly suited for nanopore data analysis as it uses a statistical approach to detect and quantify methylation events in nanopore sequencing data and uses annotated raw sequence file for detection [10]. It can also recognise other base changes such as 5-hydroxymethylcytosine (5hmC) and 4-methylcytosine, which is not easily detected by other sequencing technologies. PacBio SMRT sequencing tools use polymerase that can incorporate fluorescently labelled nucleotides and detect their integration. This technology has proven to be highly precise in determining methylation patterns, both 5-methylcytosine (5mC) and 6-methyladenine (6mA). Detection of modification is key to learning the mechanism of underlying

diseases and further identifying potential therapeutic targets [11, 12].

Pacific Biosciences (PacBio) sequencing and Oxford nanopore Technology (ONT) are two emerging technologies in Long-read sequencing technologies and are being used for high-resolution DNA methylation analysis and have shown high potential in DNA methylation detection [3, 4]. However, their performances have not been fully compared and evaluated. In this study, we will analyse the modification detection coverage and sensitivity of the ONT for DNA methylation against PacBio in *Clostridium beijerinckii. C. beijerinckii* is an oxygen tolerant, spore-forming bacterium that has high hydrogen production and acetone-butanol fermentation abilities. We aim to analyse the data and assess the similarities and contrasts between the two sequencing technologies to help researchers choose the most appropriate method for their specific research questions.

II. METHODS

A. Sequencing

C. beijerinckii NRRL B-598 was selected as an ideal reference strain for the comparison of methylation calling using various sequencing technologies as the strain contains numerous restriction modification systems with several active methylases [13]. The strain was already sequenced genome DNA of *C. beijerinckii* NRRL B-598 was sequenced using PacBio (PacBio RSII; GATC Biotech, Inc., Constance, Germany) in the past to reveal methylation of a genome wide-scale [14]. We took the original raw single molecule real-time (SMRT) sequencing data and re-analysed them. The detection of 6mA modification was done from SMRT sequencing data used for genome assembly to study the DNA methylation a genome-wide scale [10]. The data was re-analysed for 6mA methylation detection by base modification detection tool of SMRTlink v7.0.1.66975 and output was given in the GFF file format. Moreover, we did the resequencing of the genome with ONT sequencing. Genomic DNA was extracted using MagAttract HMW DNA Kit (Qiagene, NL). The sequencing library for Oxford Nanopore sequencing was prepared using Ligation sequencing 1D Kit (Oxford Nanopore Technologies, UK). The sequencing was performed using the R9.4.1 flowcell and the MinION platform (Oxford Nanopore Technologies).

B. Nanopore methylation calling

Tombo[15] is suite of tools developed by ONT, which is based on Mann-Whitney U statistics and is used to detect modification from nanopore sequencing data [16]. It provides tools for analysis and visualization of sequence data. The nanopore reads which are electrical signals present in fast5 form were first basecalled using the Guppy basecaller v6.4.2v6.4.2 [17]. In this process the nucleotide base is added to the electrical signals and the output was saved as fastq format. Annotation and alignment was then performed on it using Tombo v1.5.1 with the latest reference genome assembly of *C. beijerinckii* NRRL B-598 available under the accession number CP011966.3 in the GenBank database. The raw fast5s were annotated using the basecalled fastq sequences which saves annotated information of nucleotide bases in to the fast5 files. Then it was aligned to the reference sequence using the re-squiggling algorithm which refines the assignment from a squiggle to a reference sequence after basecalling and alignment of the tombo tool and in the process, it creates a Hdf5 file which saves information related to statistics related to genome base pairing result. After the alignment, methylation detection was performed using the alternate base detection method of tombo for detecting N6-methyladenosine(6mA) in the *C. beijerinckii* genome. The flowchart below represents the successive steps of the pipeline, see Fig. 1. Tombo has a dampened feature output option which estimates the low coverage sites modification to avoid false reading at low coverage areas.

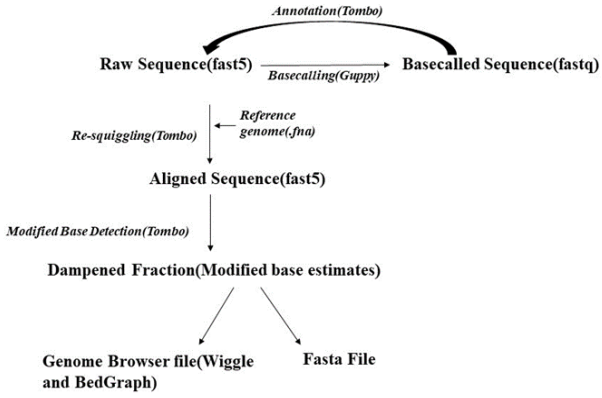


Fig. 1. Pipeline for Modification detection in the ONT sequence

III. RESULT AND DISCUSSION

A. Prediction of 6mA modification by Tombo

Prediction of 6mA modified base in ONT data was done using specific alternate base detection method. This method helps in identifying bases where signal matches the expected level for an alternate base better than the canonical expected level. This method computes statistic similar to a log likelihood ratio (LLR) but dynamically scaled to be more robust to outlier signal levels. Which is calculated using following equation:

$$ScaleDiff = NormSignal - \frac{Canonical\ Mean + Alt\ Mean}{2}$$

$$MeanDiffs = |CanonicalMean - AltMean|$$

$$OutlierRobustLLR = \frac{e^{\frac{ScaleDiff^2}{S_f.\sigma}}.\ LLR}{\sigma^2.MeanDiffs^{S_p}.S_{f2}}$$

Here "Canonical mean" refers to the average of canonical base signal while "Alt Mean" refers to the average of the alternate base signal. The equation is calculating the scale difference between mean difference of canonical base signals and alternate base signals, while also accounting for the outlier robust LLR, which is calculated by dividing the difference between the two means by the standard deviation of the mean difference. The LLR is scaled by the normal likelihood function

with the same variance and mean that is situated midway between the standard and alternate expected signal levels to produce this measure. In order to give sequence contexts with bigger disparities between the canonical and alternate expected signal levels, which naturally provide more power to discern the canonical and alternate base signal levels, more weight, three extra scaling factors are introduced to this function. For establishing single read threshold values, these parameters are also specified so that numbers are roughly on the same range as a log likelihood ratio. S_{f2}, S_f, and S_p are Scale factors. In order to determine the scaling factors, it is necessary to maximize the possibility of observing the signal intensity at the modified base location given the reference signal model and the conditional chance of observing the signal intensity given the existence of the modification.

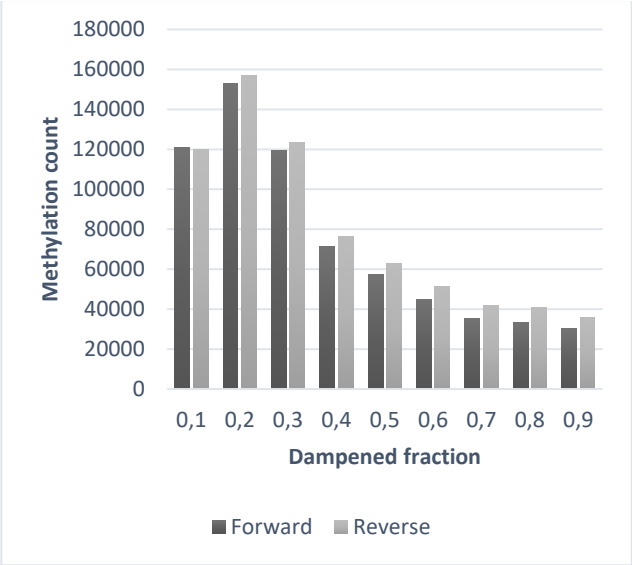


Fig. 2. Methylation counts from nanopore sequence in both forward and reverse strand plotting against their respective thresholds

The statistic is computed for each modified base with each read provided. (e.g., each adenine for 6mA detection). The results of the modification are generated in two types of text-based format: Genome Browser files (wiggle and Bedgraph) and FASTA files. Genome Browser files consist of separated files for both forward and reverse strands. Wig format output included base position and its estimated modified fraction (dampened fraction) which is the estimated fraction of significantly modified reads. The Nanopore data showed 1,030,462 and 1,023,661 base modifications from reserve and forward strand respectively with varying degree of damping fraction probability ranging from 0.1 to 0.9, see Fig. 2. A dampened fraction refers to a measure of the degree to which a signal has been suppressed/dampened or modified during sequencing. The value is determined by analysing the signal differences between bases that have been changed and those that have not. Higher value of dampened fraction may indicate a greater degree of base modification. Thus, its interpretation may vary depending on the specific analysis being carried out, though it is essential to remember that this dampened fraction output is intended to include poor coverage areas in downstream analysis without producing false positives.

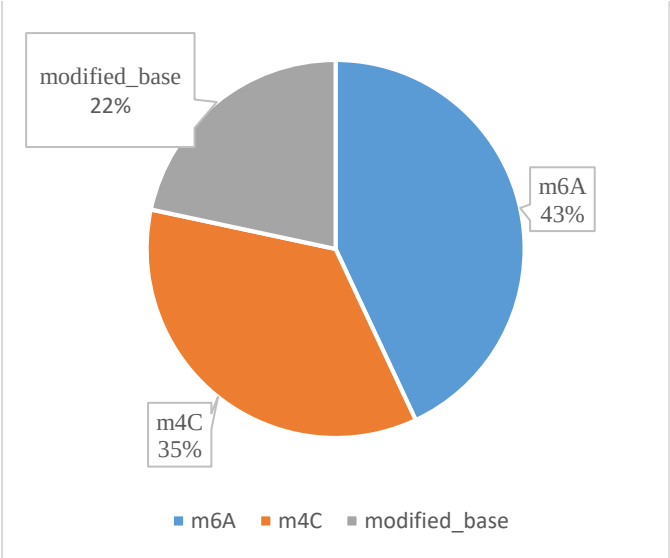


Fig. 3. Methylation type count from PacBio data

B. *Comparison of 6mA detection using PacBio data*

Both ONT and PacBio data were assembled on the same reference sequence genome. So we verified the detected methylation from ONT with Pacbio data. The PacBio data showed Methylation in 6394 bases out of which 2237 were 4mC modification, 2752 were m6A modifications and 1405 were other base modification, see Fig. 3, with varying degrees of scores. To check the methylation detection between two technologies gff file from SMRT and wiggle file from tombo was parsed to see the common methylated base sites. While only the 6mA modification was detected using tombo so only the 6mA modifications were filtered out from SMRT gff file and was parsed with the tombo wiggle file from both forward and reverse strand. After parsing the data we found that ~1300 6mA sites from PacBio data were found in the Nanopore modified base from both the strands combined, see Fig. 4.

Both the methods were effectively able to detect the 6mA methylation modification but the huge difference in number is due to the dampened fraction method of tombo tool that can be used with relatively low coverage and provides detailed quantitative information about methylation at each and can be valuable in studies that's need more precise measurements of methylation. Whereas SMRT analysis simply provides binary information about the presence and absence of methylation at each base.

On the basis of quality of sequencing data, the values of dampened fraction can be adjusted. In general, higher value of dampened fraction means result in a more conservative methylation calls i.e. a higher dampened fraction value indicates more robust signal and reduces the likelihood of false positive calls with potential of calling few false positive and lower values can result in more liberal methylation call with more false positive calls [18].

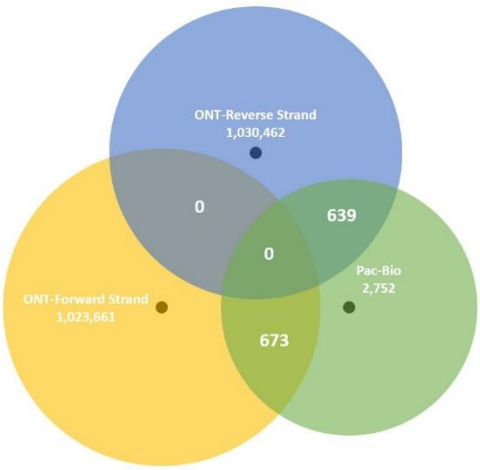


Fig. 4. Common 6mA methylated counts from Pac Bio and ONT dataset

IV. CONCLUSION

The analysis of DNA methylation modifications has become increasingly important for understanding the regulation of gene expression and the mechanisms underlying cellular processes in several organisms. In this research paper, we compared the methylation analysis of 6mA modification between PacBio and Nanopore data of *Clostridium beijerinckii*. Our findings revealed that both PacBio and Nanopore sequencing technologies are capable of accurately detecting 6mA modifications in the genome of *C. beijerinckii*. However, there were some differences in the performance of the two technologies, with ONT demonstrating a higher number of 6mA modifications in due to the ability of tombo in detecting varying degree of methylation using dampened fraction method. While both methods have their advantage and limitations, a direct comparison is complicated as both the methods use different sequencing technology and algorithm. Overall, our study provides valuable insights into these two sequencing technologies for the analysis of DNA methylation modifications, and highlights the need for careful consideration when selecting a sequencing technology as method to use depends on various factors, such as the type of sample, the required accuracy and coverage, and the available resources for methylation analysis studies.

ACKNOWLEDGMENT

This work was supported by the grant project GACR 23-05845S.

REFERENCE

1. Wang, S. and W. Wu, *DNA methylation alterations in human cancers*, in *Epigenetics in Human Disease*. 2018, Elsevier. p. 109-139.
2. Villicaña, S. and J.T. Bell, *Genetic impacts on DNA methylation: research findings and future perspectives.* Genome biology, 2021. **22**(1): p. 127.
3. Flusberg, B., et al., *Direct detection of DNA methylation and mutagenic damage through single-molecule, real-time (SMRTTM) DNA sequencing.* Cancer Research, 2010. **70**(8_Supplement): p. 1154-1154.
4. Clark, T.A., et al., *Characterization of DNA methyltransferase specificities using single-molecule, real-time DNA sequencing.* Nucleic acids research, 2012. **40**(4): p. e29-e29.
5. Smith, Z.D. and A. Meissner, *DNA methylation: roles in mammalian development.* Nature Reviews Genetics, 2013. **14**(3): p. 204-220.
6. Phillips, T., *The role of methylation in gene expression.* Nature Education, 2008. **1**(1): p. 116.
7. Tourancheau, A., et al., *Discovering multiple types of DNA methylation from bacteria and microbiome using nanopore sequencing.* Nature methods, 2021. **18**(5): p. 491-498.
8. Furuta, Y. and I. Kobayashi, *Restriction-modification systems as mobile epigenetic elements*, in *Bacterial integrative mobile genetic elements*. 2013, CRC Press. p. 85-103.
9. Seong, H.J., S.-W. Han, and W.J. Sul, *Prokaryotic DNA methylation and its functional roles.* Journal of Microbiology, 2021. **59**: p. 242-248.
10. Liu, Y., et al., *DNA methylation-calling tools for Oxford Nanopore sequencing: a survey and human epigenome-wide evaluation.* Genome biology, 2021. **22**(1): p. 1-33.
11. Jones, P.A., J.-P.J. Issa, and S. Baylin, *Targeting the cancer epigenome for therapy.* Nature Reviews Genetics, 2016. **17**(10): p. 630-641.
12. Gao, M., et al., *Direct therapeutic targeting of immune checkpoint PD-1 in pancreatic cancer.* British journal of cancer, 2019. **120**(1): p. 88-96.
13. Kolek, J., et al., *Dam and Dcm methylations prevent gene transfer into Clostridium pasteurianum NRRL B-598: development of methods for electrotransformation, conjugation, and sonoporation.* Biotechnology for biofuels, 2016. **9**(1): p. 1-14.
14. Sedlar, K., et al., *Reclassification of non-type strain Clostridium pasteurianum NRRL B-598 as Clostridium beijerinckii NRRL B-598.* Journal of biotechnology, 2017. **244**: p. 1-3.
15. *Tombo*. Available from: https://nanoporetech.github.io/tombo/index.html.
.
16. Stoiber, M., et al., *De novo identification of DNA modifications enabled by genome-guided nanopore signal processing.* BioRxiv, 2016: p. 094672.
17. *Guppy*. Available from: https://md-task.readthedocs.io/en/latest/corr.html.

Accuracy of cycling power meters

Lukáš Hrbotický
Department of Biomedical Engineering
Faculty of Electrical Engineering and Communication, Brno University of Technology
Brno, the Czech Republic
hrboticky@vut.cz

***Abstract*—This article talking about problematics of measuring the wattage of cycling on a road bike. It describes ways in which it is possible to measure the wattage on the bike and explains functional principle of power meters. As practical part of this paper the measurement is performed according to the proposed test protocol. Based on the analysis and processing of the obtained data, the accuracy of individual types of meter performance is discussed.**

Keywords—cycling, power meter, Assioma, Stages, Tacx

I. Introduction

The aim of this paper is to provide a clear and comprehensive picture of the possibilities of measuring rider wattage in road cycling. Thanks to the development of electronics, their miniaturization and accessibility, power meters are becoming affordable devices and are no longer a matter of professional riders only but are starting to appear among hobby and amateur riders as well. In the first section, a brief introduction to the subject matter is given to give the reader an overview, followed by a review of measurement methods. As this is a very interesting market, companies are also coming up with interesting approaches to measurement methods. These are listed and described in the next section, including their physical principles. The last section deals with the process of actually measuring the data, processing the data and then evaluating it.

II. Development of power meters

The history of power meters is roughly linked to the history of the development of heart rate measurement for sports use, but it was not until the beginning of this millennium that power meters became commercially widespread, mainly due to the gradual miniaturization of electronics. Just as Polar is associated with heart rate monitors, the history of power meters is made by the German company SRM and the American company CycleOps. [5]

SRM produced the first prototype as early as 1986, and in the same year filed a patent application for measuring cycling power through a transducer sleeve. Two years later, the company introduced a cycling computer capable of recording training data, and it was the wattage value that began to be key to training. The following year they introduced prototypes where they measured power in the pedal and then on the crank itself. [7]

In 1991, SRM's power meters are fitted to the entire German national team during their training camp in Switzerland. The following year, the company installs its meter on the cranks of the renowned Shimano company and the power meter begins to spread rapidly among professional riders. Not only the measuring accuracy improves, but also the electronics around it (battery life, memory for data recording...) [7].

During the 2004 Tour de France, for the first time, the actual wattage of a particular rider was displayed live along with his heart rate, speed and cadence. In 2006, the company adapted the wireless transmission for compatibility with the ANT+ communication protocol, which is increasingly used in the sports industry for data transmission between the sensor and the recording unit. In 2010, Stages enters the market and, in collaboration with the professional cycling team SKY, begins to develop a power meter that measures power on cranks. In 2013, Garmin introduces the Vector series, which is a power meter integrated into the pedals. At the time of its introduction, it was something of a "revolution" in cycling power measurement. In 2018, in collaboration with LOOK, SRM introduced a pedal-mounted power meter, and the following year they introduce the first pedal power meter for mountain bikes (the SPD system). In 2021, Garmin follows suit with the Rally line of power meters, which is compatible with several pedal systems. [7][20][16]

III. Principle of power measurement

From the previous subsection it was possible to understand that there are several ways of measuring wattage. It depends on the design and implementation by the manufacturer. However, the key component is always the strain gauge, which is a passive electrical component that changes its electrical resistance due to its deformation.

A. Strain gauge

A strain gauge is a device used to measure the mechanical stress on an object. Invented in 1938 by Edward Simmons and Arthur Ruge, it is the most common implementation of a strain gauge; a flexible insulating pad (resin or paper), which provides electrical insulation for the sensor and mediates the transfer of information from the surface of the body to the sensor itself, is glued to the object being measured and the sensor itself is on top of it. Based on the change in resistance of the circuit (usually measured using a Wheatstone bridge), knowledge of the physical properties of the material to be measured and the use of appropriate algorithms (used knowledge of Hooke's law, i.e. the

relationship of the material deformation to the applied force), one can calculate up to a resultant quantity representing the wattage of the cyclist. [4][18] In the elementary phase, the power calculation is based on the product of torque and pedaling cadence. [1] The exact procedure for converting the strain gauge signal into the resulting watt value is retained by the manufacturers as their know-how.

1) Metal: In the case of metal strain gauges, either wire or foil is used as the actual sensor. This self-sensor changes its length and thus its resistance when the object is deformed, which is then used for further conversion to the final wattage. The advantage of metal strain gauges is their low temperature dependence, high stability and accuracy. [14]

2) Semiconductor: The principle of operation of semiconductor strain gauges is different from metal strain gauges. This is because it takes advantage of the change in the resistivity of a semiconductor component due to the piezoelectric effect, i.e. the change in electrical resistance depending on the deformation of the semiconductor crystal. The advantage of semiconductor strain gauges is their high sensitivity. [14][15][17]

a) Piezoelectric effect: The piezoelectric effect is defined as a linear electromechanical interaction between a mechanical and an electrical state (in a crystalline material), where an electrical charge accumulates in response to an applied mechanical stress. There is also an inverse piezoelectric effect, which leads to deformation of the crystal (electrostriction) due to the electric voltage. The electrostriction is tiny, on the order of nanometres at voltages of a few hundred volts. [9] Physically, these are crystals with ionic bonding, where the deformation shifts ions of opposite charges in the crystal lattice so that the electrical centres of gravity of the negative and positive ions, which agree in the undeformed crystal, move away from each other. Thus, an electric charge appears on certain surfaces of the crystal and this induces an internal electric field in the crystal. [8]

B. Location

There are many places to place a strain gauge to measure wattage and many ways to calculate it, and manufacturers are trying to come up with original and increasingly accurate approaches. As the market has evolved, several solutions have been found to be the most suitable.

In this paper, the power meters installed on the bicycle of a cyclist and for outdoor use are still discussed. In practice, laboratory power meters integrated into an ergometer are also used, primarily for stress testing. [19]

1) Crank arm: The sensor for measurement is located on the left, right or both cranks. It therefore allows measurements for each lower limb separately and to evaluate, for example, left-right leg balance. The measuring unit is placed on the inside of the crank so that it does not interfere with the aerodynamics outside. On the left crank it is placed in the centre of the crank and on the right crank it is in the space between the crank and the small chainring. The mechanical stress of the crank itself is measured. This is probably the most widely used measurement method today, as Shimano integrates it into their cranks at the factory if desired. [2]

2) Crank spider: The measurement is made on the gear carrier. This is where the converters are bolted to the right crank structure. The principle already implies that such a solution only measures the total power, you cannot get separate values for the left and right lower leg. This is the solution that SRM started the history of power meters with and nowadays this solution is mainly used only for track bikes (the power meter does not interfere with changing the chainrings, which is a common activity for track cyclists). In road cycling it has been displaced by competing, more affordable, products. [2]

3) Pedal: The measuring device is integrated into the body of the pedal, often on its axis. It measures the mechanical stress in the pedal axis, which is screwed into the crank on one side and housed in the pedal bearing on the other. This is the most advantageous solution in terms of installation, as changing pedals is much easier and quicker than changing cranks, and compatibility is easier to ensure with pedals. In the event of mechanical damage to the pedal, only the pedal "body" can be replaced and the more expensive electronics are retained. [2]

4) Rear hub: Measurements are taken in the axis of the rear wheel. A solution was developed by PowerTap, but the solution has not met with much success. You need to ensure the nut is compatible with the cassette installed, the number of holes for spokes, and also the hub needs to be spun out on the rim. In professional cycling, swapping pairs of wheels is the order of the day and in racing, only the flat wheel is swapped in the event of a puncture, which is not beneficial in the case of a rear wheel with a power meter. Unless the rider gets a spare wheel with a power meter again, he loses information about his performance. This method of measurement is not very accurate mainly because there is a "loss" of power between the rider (his foot) and the rear head due to bearing friction and material twisting. [3]

5) Indirect measurement: The name implies that the power is not calculated from the mechanical stress in the material measured by the strain gauge, but by other physical quantities. This solution was developed by PowerPod, whose measuring device is placed under the stem and uses wind speed, wheel speed and slope, among other things, to calculate power. [10] In terms of installation, this is an efficient solution. In terms of accuracy, it is very important to correctly determine the rolling resistance of the wheel. If this is determined correctly, it can be as accurate as a strain gauge power meter. [25] However, it cannot take into account, for example, the aerodynamic behaviour of cycling. However, its positioning on the bike distorts the aerodynamics of the bike and makes it very susceptible to damage if the bike is dropped or handled. The question arises whether a particular user needs such a solution and whether it would not be appropriate to choose some of the more widespread and reliable measurement principles.

IV. Measurement procedure

For the purpose of this measurement, a research question was set: Do the wattage meters used show the same values during testing and is there no difference between them?

Three devices were used for the measurements; Tacx NEO 2T smart trainer, Stages Gen 3 Power LR power meter installed on Shimano Ultegra FC-R8000 cranks, and Favero Assioma DUO power meter integrated into the LOOK KEO compatible pedals.

Both of the above two power meters were installed on one road bike and this was further mounted to the trainer (Fig. 1). The power meters were paired via ANT+ protocol with a Garmin Forerunner 935 sports watch to record activity. The trainer was paired via Bluetooth protocol with a laptop from where the load on the trainer was controlled based on the test protocol. Before starting the measurements, initial calibration of all devices was performed according to the manufacturer's instructions.



Fig. 1. Measuring set-up

The actual measurement was performed with ten runs according to the test protocol. The protocol was designed to test all riding styles; power and frequency riding, slow and speed riding and standing riding. As each cyclist has a different seating position and pedaling style, this will be a reliable finding.

However, the measurements for this thesis were taken with only one participant (male, 27 years old, 194 cm, 80 kg) as it is not possible to provide the conditions for testing a larger group of individuals due to the unavailability of the necessary materials.

The test subject was instructed to maintain the same posture and riding style throughout all measurements to maximize the meaningfulness of the data obtained. The actual pedaling cadence sensed by the trainer was displayed on the laptop monitor and here the test person had the opportunity to check it to maintain the value given by the testing protocol.

During the measurements, it was found that the trainer had a problem with the transition from the resting phase to the phase simulating downhill ride. The trainer once increased the load unreasonably in excess of 500 W. The choice of a lighter (53 × 13) gear proved to be a partial solution to this problem, but even in this case the power output remained above the specified value of 250 W and oscillated greatly when maintaining a cadence of 110 rpm.

V. Data acquisition

A. *Data export*

Normally, recorded data is viewed either in the mobile app or in the Garmin Connect web-based training log. Here, data is presented in the form of graphs and statistical values for the general public. The data can be exported to TCX format for data processing. Although this is a non-standard file extension, MS Excel spreadsheets can process it. However, the data exported in this way does not contain all the recorded values that can be seen in the training log. In order to get all the recorded data, the device needs to be connected to a computer via USB interface as a high-capacity storage device. Here is the possibility to copy the raw activity data in FIT format. This is a specific file format, which requires the GoldenCheetah program to open. This can display the raw data and further use it as needed.

If the power meter is connected via Bluetooth (the case of a trainer connected to a laptop), the device is not able to receive additional information about pedaling dynamics. For this reason, all data analysis was limited to the overall rider performance only. This occurs because Garmin primarily uses the ANT+ protocol to connect external sensors and is only just introducing support for the Bluetooth protocol among its products. It is only a matter of time before the implementation of Bluetooth support is at a stage where it can also handle the transmission of ride dynamics metrics.

B. *Data processing*

All data processing was performed in the MATLAB software environment version R2019a. As the original signal is noisy, it was filtered before any subsequent analysis. The filtering was done with a low-pass filter and a cutoff frequency of 0.1 Hz was determined by optimization. This modified signal was subjected first to visual analysis and then to numerical analysis.

After an initial review of all measurements, the first and third measurements had to be excluded from the subsequent analysis, as the trainer did not transmit power information for several minutes. For this reason, the subsequent analysis is performed on only eight measurements.

By visual analysis of the measured signals, it was decided not to evaluate the part of the test protocol representing the downhill speed ride. The trainer was not capable of rapid load changes and hence the interval is not made up of riding at near constant power. Three load intervals are used for the analysis (power riding in the saddle, power standing riding and endurance riding).

C. *Statistical analysis*

Data corresponding to a given measuring instrument and a specific interval from the test protocol were combined into one data set and this was used for all subsequent analysis. This resulted in nine data sets (three measuring instruments times three test intervals).

1) Visual analysis: In addition to the eye-metric method of analysing all measured signals, box plots were used for visual representation. Boxplots are one way of graphically visualizing differences between numerical data in descriptive statistics without assuming a normal distribution (they are nonparametric). The middle, boxed, part of the graph (shown in blue below) represents the 3rd quartile (upper bound) and the 1st quartile (lower bound). The horizontal red line represents the median. The black boundaries represent the maximum and minimum, respectively, of the data set. Any outliers are represented by red crosses. [6]

2) Test analysis: The Lilliefors normality test was used to test for a normal distribution of the data, which is appropriate for larger sample sizes ($n > 80$). [13] The Lilliefors normality test is one of many so-called goodness-of-fit tests for random samples with a normal distribution. The test is based on the principles of the Kolmogorov-Smirnov test. The test was conducted at a significance level of $\alpha = 0.05$ and evaluated that all but one data set has no normal distribution of values. That one data set with a normal distribution of data is the endurance interval measured by the pedal power meter. Based on the above findings, a non-parametric test for dependent measurements was chosen for the subsequent analysis. [12] Friedman's test for repeated measurements proved to be a suitable nonparametric test for multiple samples. [13] In all cases, the test was performed at a significance level of $\alpha = 0.05$ and always for a trio of measuring instruments within the same test interval. A sign post-hoc test was additionally performed when necessary.

VI. Evaluation

A. Visual analysis

Looking at all the graphical representations of the entire measured signal (Fig. 2), it can be observed in all measurements that the cranks overestimate the power (compared to the other two power meters) and the trainer underestimates the power. This phenomenon is most pronounced during low-load relaxation driving. By visual analysis it is possible to determine when the differences between the power meters used are minimal; during low cadence power riding.

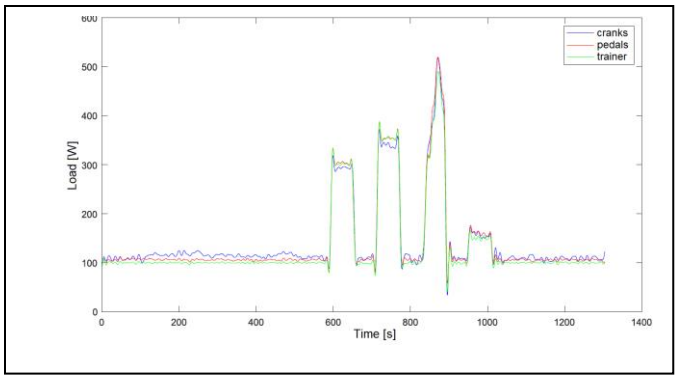


Fig. 2. Measurement nr. 1

The underestimation by the trainer can be justified by the fact that it is located "furthest" from the power source (the cyclist's foot). Thus, power is lost through bearing and chain friction and material flex (especially the frame at the pedal center).

However, if we focus only on the three intervals of interest, we can confirm the above conjecture (of overestimation by the cranks and underestimation by the trainer) for the endurance section from the box plots (Fig. 3, Fig. 4, Fig. 5). Conversely, power riding (both in and out of the saddle) appears to be the riding style where the meters give the most identical values. Power riding in the saddle appears eye-metrically to be the best for accuracy of measurement by different measuring devices (median values are closest to each other and intervals between the 1st and 3rd quartiles are lowest).

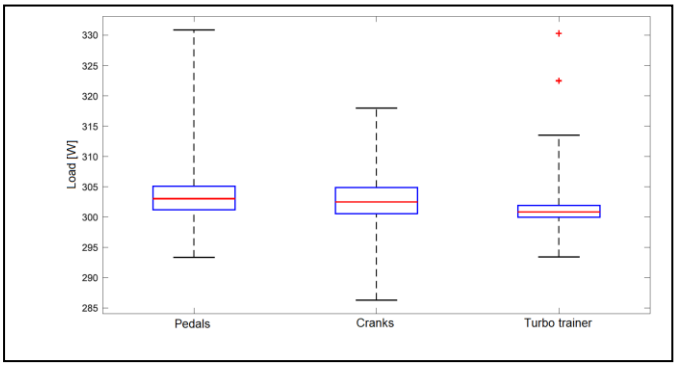


Fig. 3. Box plot for a section in the saddle

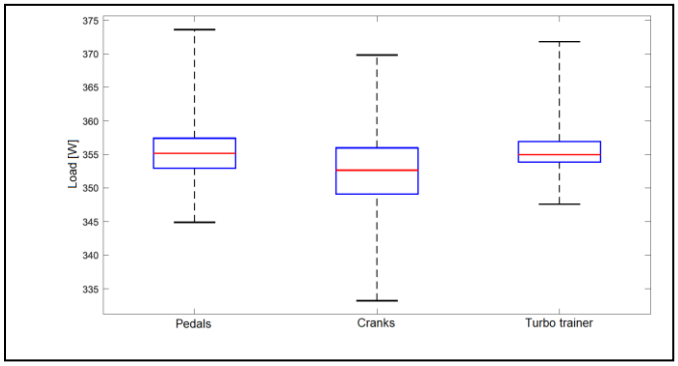


Fig. 4. Box plot for a standing section

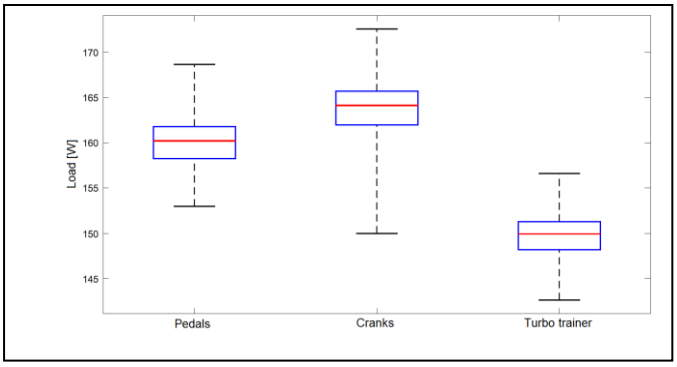


Fig. 5. Box plot for an endurance section

B. *Friedman test analysis*

To obtain data from which to draw conclusions, it was necessary to perform statistical tests and have numerical results. The Friedman test for repeated measures was performed. This was carried out at a significance level of $\alpha = 0.05$. Its results showed that within a given test interval there were statistically significant differences between all three measuring instruments. In order to reveal between which two measuring instruments there is or is not a difference, it was necessary to perform a post-hoc test. [11]

The sign post-hoc test revealed that there is only a statistically insignificant difference between the pedal power meter and the trainer during power riding from the saddle. In all other cases it showed a statistically significant difference.

P-values of Friedman's test compared to the significance level α are key. If the p-value does not exceed the significance level, the null hypothesis H0 that there is no statistically significant difference between the tested data sets is rejected. This happens with 95% confidence because we are working with a significance level of $\alpha = 0.05$.

VII. Conclusion

This article has provided a clear and comprehensive picture of the possibilities of measuring rider wattage in road racing cycling. In the first part, a brief introduction to the issues addressed was given, followed by a review of measurement methods. In the last part, the measurement procedure, subsequent data processing and evaluation were described.

For the evaluation of which of the three performance measures used is the most accurate, the statistical tests performed are not of much help, as they always showed a statistically significant difference between the data sets, except in one case. If we follow the theory described in the text that the power value measured by the trainer may be lower because of its "distance" from the power source, we will not take it into account when evaluating accuracy. Thus, we can determine the pedal power meter to be the most accurate (of the three used) based on its lower standard deviation compared to the crank power meter and narrower interquartile range.

The answer to the research question is no, i.e., the power meters used show different values during testing and there is a difference - both statistical and practical - between them.

References

[1] BINI, Rodrigo, Fernando DIEFENTHAELER a Felipe CARPES. Determining force and power in cycling: A review of methods and instruments for pedal force and crank torque measurements . INTERNATIONAL SPORTMED JOURNAL. 15. 96-112.

[2] BOUILLOD, Anthony, Julien PINOT, Georges SOTO-ROMERO, William BERTUCCI a Frederic GRAPPE. Validity, Sensitivity, Reproducibility, and Robustness of the PowerTap, Stages, and Garmin Vector Power Meters in Comparison With the SRM Device. International Journal of Sports Physiology and Performance [online]. 2017, 12(8), 1023-1030 [cit. 22. 5. 2022]. ISSN 1555-0265. doi:10.1123/ijspp.2016-0436

[3] BUREŠ, Petr. Dny PowerTapu jsou zřejmě sečteny, Sram ho potichu vyřazuje z nabídky [online]. 2021 [cit. 28. 12. 2021]. URL: <https://mtbs.cz/clanek/dny-powertapu-jsou-zrejme-secteny-sramho-potichu-vyrazuje-z-nabidky/kategorie/doplnky-na-kolo>

[4] KASHIRI, Navvab, Jorn MALZAHN a Nikos G. TSAGARAKIS. On the Sensor Design of Torque Controlled Actuators: A Comparison Study of Strain Gauge and Encoder-Based Principles. IEEE Robotics and Automation Letters [online]. 2017, 2(2), 1186-1194 [cit. 29. 12. 2021]. ISSN 2377-3766. doi:10.1109/LRA.2017.2662744

[5] KLEINER, Robert. Měřiče Wattového výkonu – zahájení seriálu [online]. 2011 [cit. 23. 12. 2021]. URL: <https://mtbs.cz/clanek/merice-wattoveho-vykonu-zahajeni-serialu/kategorie/trenink-a-fitness>

[6] Krabicový diagram (Box plot) [online]. [cit. 4. 8. 2022]. URL: <https://portal.matematickabiologie.cz/index.php?pg=analyza-a-hodnoceni-biologickych-dat--statisticke-modelovani--pruzkumova-analyza-jednorozmernych-dat--diagnosticke-grafy--krabicovy-diagrambox-plot>

[7] Our story of success [online]. [cit. 24. 12. 2021]. URL: <http://www.srm.de/company/history/>

[8] Piezoelektrické jevy [online]. [cit. 13. 5. 2022]. URL: <http://fyzika.fs.cvut.cz/subjects/fzmt/lectures/FZMT_11.pdf>

[9] POHANKA, Miroslav. Overview of Piezoelectric Biosensors, Immunosensors and DNA Sensors and Their Applications. Materials [online]. Materials [cit. 13. 5. 2022]. ISSN 1996-1944. doi:10.3390/ma11030448

[10] PowerPod In-Depth Review [online]. [cit. 26. 12. 2021]. URL: <https://www.dcrainmaker.com/2016/03/powerpod-depth-review.html>

[11] RONZHINA, Marina. Statistická analýza dat: Neparametrické testy pro jeden a více výběrů. 2022 [cit. 7. 8. 2022].

[12] RONZHINA, Marina. Statistika - dotaz [e-mail]. Message to: hrboticky@vut.cz. 4. 8. 2022 17:35 [cit. 4. 8. 2022].

[13] RUSZ, Jan. Experimental Data Analysis: Correlations, normality of data testing, parametric vs. non-parametric tests [online]. 2018 [cit. 27. 7. 2022]. URL: <http://sami.fel.cvut.cz/aed/EDA_Lecture4.pdf>

[14] Snímače mechanického namáhaní - odporové tenzometry [online]. [cit. 30. 12. 2021]. URL: <https://eluc.kr-olomoucky.cz/verejne/lekce/1588>

[15] Snímače mechanického namáhaní – piezoelektrický jev [online]. [cit. 30. 12. 2021]. URL: <https://eluc.kr-olomoucky.cz/verejne/lekce/1590>

[16] The Stages Story [online]. [cit. 25. 12. 2021]. URL: <https://stagescycling.com/en_us/discover/our-story>

[17] VOJÁČEK, Antonín. Polovodičové tenzometry - princip, provedení, použití, historie [online]. 2006 [cit. 31. 12. 2021]. URL: <https://automatizace.hw.cz/clanek/2006111601>

[18] WAGNEROVÁ, Renata a Jiří TŮMA. Základy automatizace [online]. Ostrava: Vysoká škola báňská - Technická univerzita, [2008] [cit. 31. 12. 2021]. ISBN 978-80-248-1523-7. URL: <http://www.elearn.vsb.cz/archivcd/FS/Zaut/Skripta_text.pdf>

[19] World-Class Smart Indoor Bike Trainers | Wattbike [online]. [cit. 15. 5. 2022]. URL: <https://wattbike.com/>

[20] ZELENÝ, Marek. Pedály Garmin Vector, revoluce v měření cyklistického výkonu [online]. 2013 [cit. 24. 12. 2021]. URL: <https://iluxus.cz/2013/12/13/pedaly-garmin-vector-revoluce-v-mereni-cyklistickeho-vykonu/>

A simplified 2D model of reentry in cardiac tissue using the FHN model.

1st Martin Hemzal
dept. of Power Electrical and Electronic Engineering
Brno University of Technology
Brno, Czechia
Martin.Hemzal@vut.cz

***Abstract*—Spatial inhomogeneities in cardiac tissue, such as fibroblast clumps, scar tissue, or ischemic tissue, are well known to be sources of reentrant activity in the muscle that can induce certain cardiac arrhythmias. In this paper, two models of cardiac tissue were investigated. First, a homogeneous cardiac tissue model is introduced, in which the normal action potential distribution is shown. The second model with spatial inhomogeneity (regional ischemia, scar tissue, fibroblast clumps) that induces the reentrant activity of the cardiac tissue. The distribution of the action potential was solved in the cardiac tissue using a monodomain model with a modified FitzHugh-Nagumo (FHN) ionic model. To simulate the inhomogeneity in the tissue, one parameter of the FHN model that accounts for the excitability of the cell was modified, as well as the diffusion coefficient in the reaction-diffusion equation. This modification led to the reentrant activity in the second model. This simplified model can be used to study arrhythmogenic waves induced by spatial inhomogeneities in cardiac tissue.**

***Index Terms*—FEM, FitzHugh-Nagumo, monodomain model, reentry**

I. Introduction

Cardiovascular diseases are the leading cause of death in the world. Death is often caused by variations in the normal rhythm of the heartbeat (cardiac arrhythmias). The most dangerous arrhythmias are ventricular tachycardia and ventricular fibrillation. Although atrial fibrillation is not a life-threatening state by itself, it can significantly decrease quality of life or even lead to other life-threatening states [1].

Biophysical mathematical models of cardiac tissue are powerful tools for investigating the electrical activity in cardiac tissue. Several mechanisms of cardiac fibrillation were described on the basis of numerical results. Modeling has also been used to study the spatiotemporal distribution of exciton waves in the 2D and 3D models of tissue. One of the underlying mechanisms that can cause variation in heart tissue (arrhythmias) may be reentrant activity in cardiac tissue. One of the reasons why reentrant activity can occur in cardiac tissue is the modified electrophysiological properties of a certain region of the tissue that can be caused by scars, fibroblast clumps, or local ischemia [1] [2].

This article focuses on the effect of inhomogeneity in cardiac tissue that can induce arrhythmogenic reentrant waves in the model. Two models of cardiac tissue were modeled. Firstly, the homogeneous sheet of tissue on which the normal action potential is shown, and then the model with reentrant activity.

II. Material and methods

A. Mathematical model

To solve for the action potential in cardiac tissue, the reduced monodomain model (1) based on the reaction-diffusion equation was used and it was coupled to the suitable ODE that describes cell dynamics (2).

$$\frac{\partial u}{\partial t} = \nabla \cdot (D\nabla u) + I_{ion}(u,v) + I_{st}, \tag{1}$$

$$\frac{\partial v}{\partial t} = f(u,v). \tag{2}$$

Where u is the action potential (V), D ($\mathrm{m^2\,s^{-1}}$) is the diffusion coefficient, I_{ion} ($\mathrm{A\,cm^{-1}}$) is the ionic current through various types of ionic channels, I_{st} ($\mathrm{A\,cm^{-1}}$) is the external current stimulus that initiates wave propagation of the activation, variable u, and v is the gating variable that controls the state of the ionic channels. To obtain an approximate solution of (1) and (2) the equations were discretized by finite difference with step Δt

$$\frac{u^n - u}{\Delta t} = \nabla \cdot (D\nabla u) + I_{ion}(u,v) + I_{st}, \tag{3}$$

$$\frac{v^n - v}{\Delta t} = f(u,v). \tag{4}$$

Multiplying these equations by the test function ψ equations (5) and (6) were obtained

$$u^n\psi - u\psi = \Delta t[\nabla \cdot (D\nabla u)]\psi + \Delta t I_{ion}(u,v)\psi + \Delta t I_{st}\psi, \tag{5}$$

$$v^n\psi - v\psi = \Delta t f(u,v)\psi. \tag{6}$$

Then by integrating by parts and applying the Green's divergence theory the weak solutions (7) and (8) of the differential equations (1) and (2) were obtained

$$\int_\Omega u^n\psi d\Omega = \int_\Omega D\Delta t\nabla u\nabla\psi d\Omega + \int_\Omega u\psi d\Omega + \int_\Omega \Delta t I_{ion}(u,v)\psi d\Omega + \int_\Omega \Delta t I_{st}\psi d\Omega, \tag{7}$$

$$\int_\Omega v^n d\Omega = \int_\Omega \Delta t f(u,v)\psi d\Omega + \int_\Omega v\psi d\Omega. \tag{8}$$

B. Ionic model

Different ionic models are used to compute the action potential distribution in cardiac tissue. Many of those models use a large number of parameters to model the ion channels. These parameters enable one to change the properties of the simulated ionic channels. But consequently, they are more computationally expensive. For simulations of reentry in the 2D model of cardiac tissue, the FitzHugh-Nagumo (FHN) model was chosen for its simplicity. The ionic currents are described by the function (9), and the gating variable is described by the function (10). The main advantage of a such simplified model over a realistic one lies in computational efficiency. The model was appropriately scaled both in time and in transmembrane potential [3] [4].

$$I_{ion}(u,v) = kc_1(u-B)\left[a-\frac{(u-B)}{A}\right]\left[1-\frac{(u-B)}{A}\right]-kc_2v \tag{9}$$

$$f(u,v) = ke\left[\frac{u-B}{A}-v\right] \tag{10}$$

Where v (-) is the recovery variable, a (-) is related to the excitation threshold, e (-) is related to the excitability of cells and controls the duration of the action potential, A (V) is the amplitude of the action potential, B (V) is the resting membrane potential and c_1 (-), c_2 (-) and k (ms^{-1}) are membrane-specific parameters. The parameters of the ionic model are shown in Table I.

TABLE I: Parameters Used in Simulations [3] [4]

Parameter	Cardiac tissue	Inhomogeneous tissue
a	0.13	0.13
B	-0.085 V	-0.085 V
A	0.1 V	0.07 V
e	0.013	0.0008
c_1	2.6	2.6
c_2	1	1
k	1000 ms^{-1}	1000 ms^{-1}
D	0.008 $\mathrm{m}^2\mathrm{s}^{-1}$	0.002 $\mathrm{m}^2\mathrm{s}^{-1}$

C. Geometry of the model

In order to calculate the spatial distribution of the action potential in cardiac tissue, two simple geometries were considered. For the homogeneous case, a rectangle with a dimension of 10 cm was used. The computational region is denoted by Ω and the boundary of the region by Γ. To solve the model the reentry model, a spatial inhomogeneity was introduced into the geometry by a rectangle with dimensions of 4 cm. The region of inhomogeneity is denoted by Ω_2, see Fig. 1.

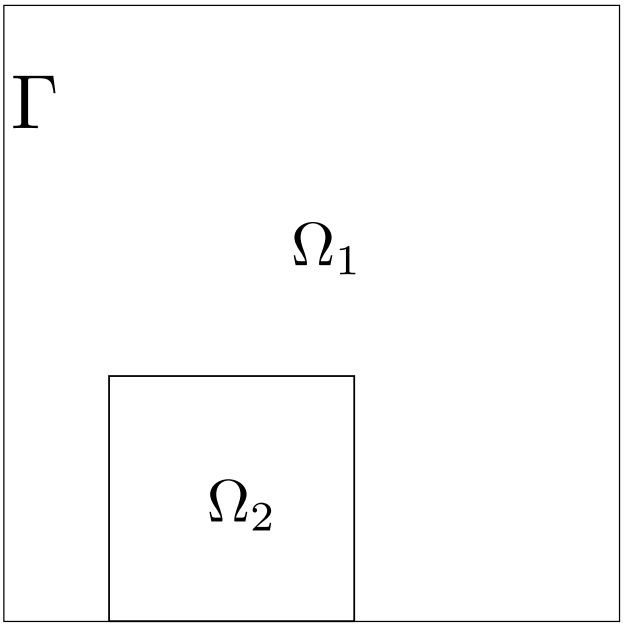


Fig. 1: Geometry of the reentry model.

D. Implementation of the model

The model was implemented using Freefem [5]. The explicit forward Euler integration scheme was used for both models with a time step $\Delta t = 0.05$ ms. The boundary conditions on all outer edges of both models were set to the Neuman boundary condition (11). The initial conditions of $u = -0.085$ and $v = 0$ were set.

$$\frac{\partial u}{\partial n} = 0 \ on \ \Gamma. \tag{11}$$

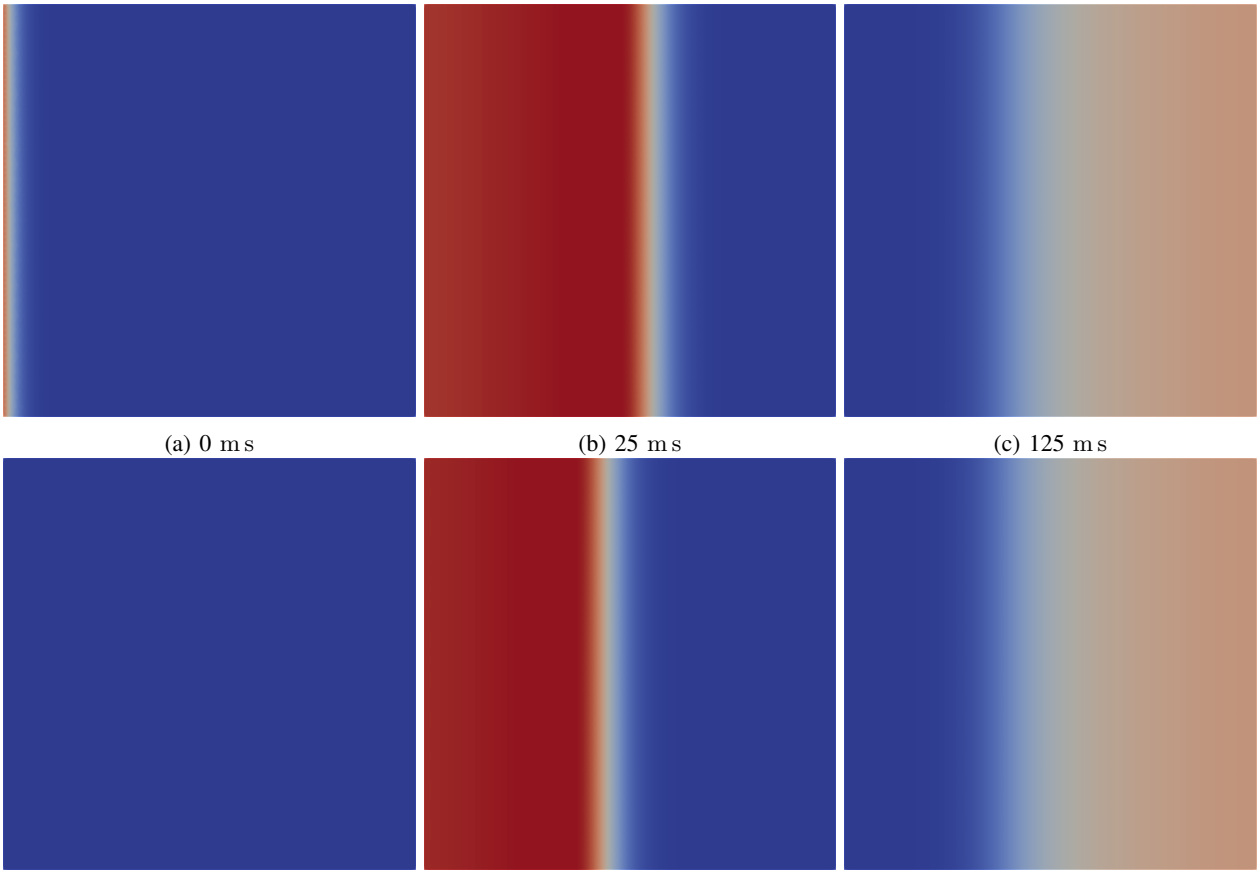
For pacing the tissue, two external stimuli of 400 $\mu\mathrm{A\,cm}^{-2}$ and 3 ms were applied to the left part of the Ω_1 region with a basic cycle period of 920 ms. In the reentry model, the Ω_2 region was introduced to model inhomogeneity in the tissue. The two parameters e and D were modified to represent inhomogeneous tissue.

III. Results

Simulations were carried out in simplified geometries to evaluate the distribution of the action potential in the 2D model of cardiac tissue. First, the results are discussed on a homogeneous sheet of cardiac tissue with a basic pacing cycle period of 920 ms. Then the reentry model is discussed.

A. Homogenous model

The resulting spatial distribution of the action potential is shown in Fig. 2. An external stimulus I_{st} was applied to the left part of the tissue at 0 ms. The tissue starts to depolarize and after 50 ms the whole tissue is activated. Then the action potential plateaus and at 100 ms and starts to repolarize. After 180 ms, the tissue is completely repolarized. At 920 ms the second stimulus is applied to the left part of the tissue. The action potential in the center of the homogeneous tissue is shown in Fig. 3.



(a) 0 ms (b) 25 ms (c) 125 ms

(d) 180 ms (e) 940 ms (f) 1045 ms

Fig. 2: Simulation of the action potential in homogenous tissue. The blue color represents repolarized tissue and the red color represents depolarized tissue. External current applied to the leftmost column of the tissue at t = 0 ms (a). Depolarization of the tissue after initiation of the APD at = 25 ms (b). Repolarization of the tissue after 125 ms (c). The tissue is completely repolarized after 180 ms (d). Depolarization at 940 ms after the second stimulus. Repolarization after the second stimulus at 1045 ms (f).

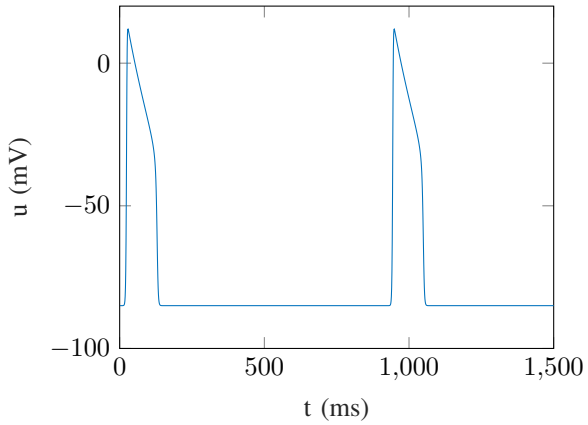


Fig. 3: Action potential at the center of the homogeneous case.

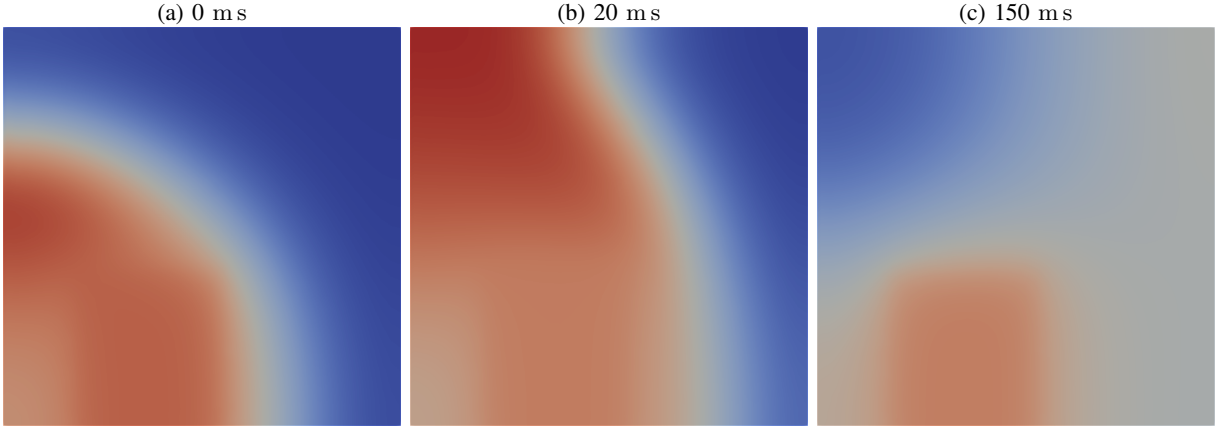
B. Inhomegeneity in the tissue

The distribution of the action potential in the model with the inhomogeneity is shown in Fig. 4. The external stimulus is applied to the left part of the tissue at 0 ms. The tissue starts to depolarize, but when the action potential propagates through the tissue, it enters the inhomogeneity, where it progresses slower due to the lower diffusion coefficient. The wave travels faster at the region's top and starts to curl around the area Ω_2, see Fig. 4b. Then the tissue starts to repolarize, but due to reduced excitability, the region Ω_2 remains depolarized; see Fig. 4c. This state of polarization of the tissue leads to spontaneous activation of the tissue, thus reentry, see Fig. 4d. This leads to sustained reentry waves in the model. Three extra wave cycles occur due to the fact that the region Ω_2 of the tissue has a prolonged repolarization time and behaves as a source of tissue reactivation.

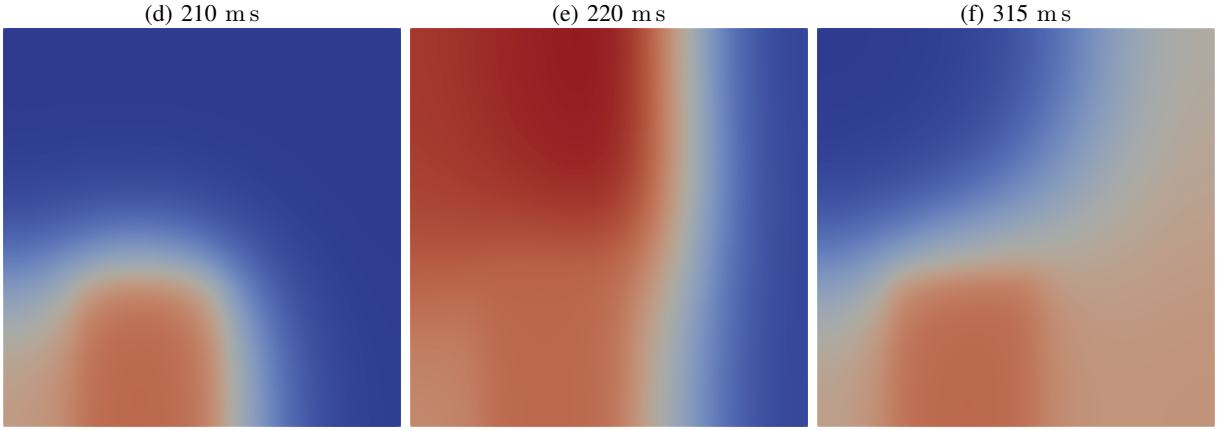
The action potential at three different locations is shown in Fig. 5. The action potential at the location of the external stimulus is unable to completely repolarize, as can be seen in Fig. 5a. The tissue does not have time to reach the resting value of the action potential (-85 mV). The action potential of the right part of the tissue is shown in Fig. 5b. The right part of the modelled tissue is able to completely repolarize to a resting action potential of -85 mV.



(a) 0 ms (b) 20 ms (c) 150 ms



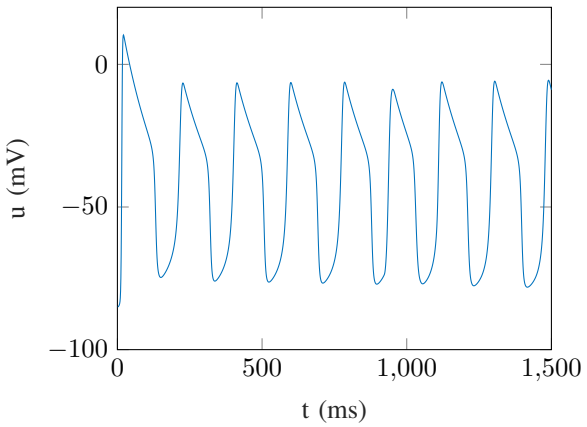
(d) 210 ms (e) 220 ms (f) 315 ms



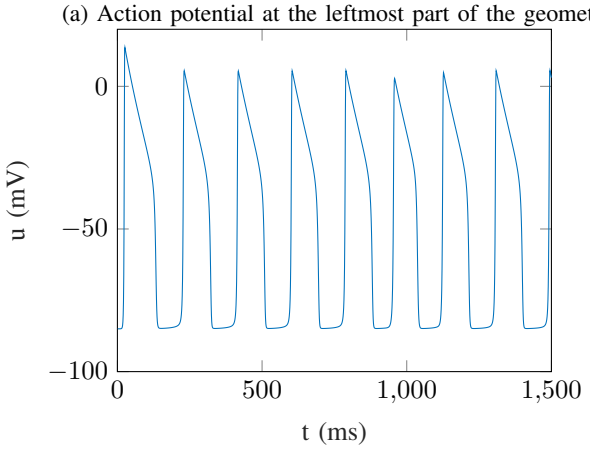
(g) 920 ms (h) 945 ms (i) 1030 ms

Fig. 4: Simulation of the spatial distribution of action potential in heart tissue with inhomogeneity introduced to the model. The blue color represents repolarized tissue while the red color represents depolarized tissue. External stimulus is applied to the left part of the model at t=0 ms (a). The tissue starts to depolarize (b). After 150 ms (c) the tissue starts to repolarize. Initiation of the first reentry at 210 ms (d). The tissue starts to depolarize (e). Repolarization of the tissue after first reentry (f). Application of the second stimulus at 920 ms (g). Depolarization of the tissue at 945 ms (h). Final repolarization of the tissue at the end of the simulation (i).

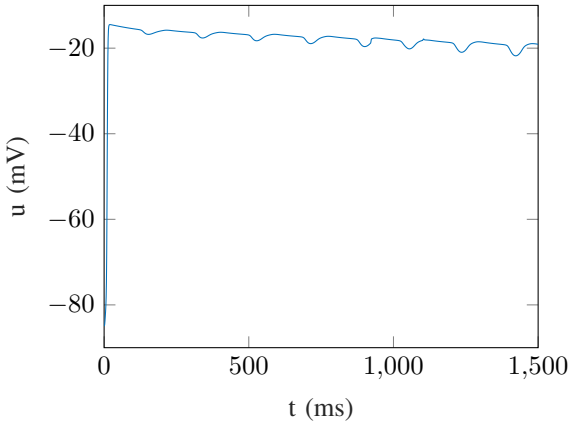
The action potential of the modified region Ω_2 is shown in Fig. 5c. The tissue is depolarized at the initial external stimulus at 0 ms. After the first external stimuli, the tissue in the region Ω_2 depolarizes and due to the increased parameter e, it does not have time to repolarize simultaneously with the tissue in the region Ω_1. This leads to the reactivation of the tissue without external stimuli, and thus to reentry. From the initial pacing of 920 ms, which corresponds to the activation frequency of 1.08 Hz the activation frequency increased to 5.35 Hz.



(a) Action potential at the leftmost part of the geometry.



(b) Action potential at the rightmost part of the geometry.



(c) Action potential at the region Ω_2.

Fig. 5: Action potential in reentry model at measured points.

IV. CONCLUSION

The spatial distribution of the action potential was studied in both a homogeneous tissue sheet and tissue with inhomogeneity incorporated using the monodomain model and the FHN ionic model. It was found that by introducing the inhomogeneous region Ω_2 with modified properties of excitability and diffusion coefficient, sustained reentry can be modeled by FHN. The reduced parameter e and the pacing frequency were found to play a crucial role in the initiation of reentry, but mainly in sustained reentry. If excitability is not reduced sufficiently, the reentry still occurs but after a certain amount of time, the tissue repolarizes completely until the next external stimuli. The geometry presented in this paper is a very simplified case of the 2D model of tissue and aims only to show the possibility of modeling reentry with a simple and efficient FHN ionic model. In the future, it would be convenient to use more realistic heart geometries and different ionic models and compare the results with the ECG obtained from fibrillating states. Another way of using this model is to model the distribution of action potentials after electroporation is used in the heart to treat arrhythmias and to compare the numerical results with the experimental results. These could be the action potential activation maps obtained from experiments that we are working on with the FNUSA-ICRC.

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-23-8430). The author acknowledges Ms. Veronika Novotná as my research supervisor for her valuable comments and help while writing this paper.

REFERENCES

[1] Xu, Aoxiang and Guevara, Michael R., "Two forms of spiral-wave reentry in an ionic model of ischemic ventricular myocardium," 1998 Chaos: An Interdisciplinary Journal of Nonlinear Science, vol. 8, pp. 157-174, 1998, doi:10.1063/1.166286.

[2] Martyn P. Nash, Alexander V. Panfilov, Electromechanical model of excitable tissue to study reentrant cardiac arrhythmias, Progress in Biophysics and Molecular Biology, Volume 85, Issues 2–3, 2004, Pages 501-522, ISSN 0079-6107, https://doi.org/10.1016/j.pbiomolbio.2004.01.016.

[3] S. Dokos, S. L. Cloherty and N. H. Lovell, "Computational Model of Atrial Electrical Activation and Propagation," 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 2007, pp. 908-911, doi: 10.1109/IEMBS.2007.4352438.

[4] Sovilj Siniša, Magjarević Ratko, Abed Amr Al, Lovell Nigel H. and Dokos Socrates. "Simplified 2D Bidomain Model of Whole Heart Electrical Activity and ECG Generation" Measurement Science Review 14, no.3 (2014): 136-143. https://doi.org/10.2478/msr-2014-0018

[5] HECHT, Frédéric. New development in FreeFem++. Journal of numerical mathematics, 2012, vol. 20, no 3-4, p. 251-266.

Short review of current EVs

1st Martin Šedina
Department of Electrotechnology
Faculty of electrical engineering and communication
Brno university of technology
Brno, Czech Republic
xsedin00@vutbr.cz

2nd Tomáš Kazda
Department of Electrotechnology
Faculty of electrical engineering and communication
Brno university of technology
Brno, Czech Republic
kazda@vutbr.cz

***Abstract*—This article discusses recent advances of EVs. It explores the level of electrification of the vehicle. It also explores cathode materials for electric vehicles, future battery technologies such as solid-state and lithium-sulfur batteries, and BMS, as well as various charging methods available for EVs.**

***Index Terms*—electromobility, Li-ion batteries, battery pack, BMS, EV charging**

I. Introduction

One of the things that ordinary people contribute to pollution is personal transport. The most common vehicles found on the road today burn some form of fossil fuel and this, combined with the fact that most vehicles are occupied by only one person, contributes greatly to the development of global warming. For this reason, and to protect the environment, there are emission limits for passenger vehicles. Because of these limits, manufacturers have then started to offer alternative propulsion in the form of electricity or hydrogen. These drives then allow local emission-free operation. [1]

The first drive for passenger vehicles was electric power, but over time, due to cost reductions, simplification of ICE vehicle operation and significantly longer range, electric propulsion was abandoned. The modern wave of electrified vehicles began with the Toyota Prius in 2000 [2]. After 2010, a wide range of fully electric vehicle models became available on the market. Fig.1 shows that 9.5 % of total car sales in 2022 were fully electric. [4] [3]

There are currently many rumours and confusion in the media about electromobility among the general public. In this respect, awareness of the principles of electric vehicles needs to be spread.

II. Level of electrification

In an effort to reduce emissions from the internal combustion engine, manufacturers have moved to higher or lower levels of vehicle electrification. Vehicles that do not require an active combustion engine to move can be considered electric vehicles.

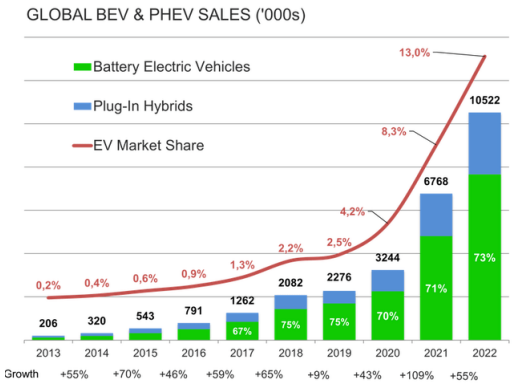


Fig. 1. Globa sales of BEV and PHEV 2013-2022 [4]

Hybrid vehicle (HEV)

Hybrid vehicles use a combination of an internal combustion engine and a small electric motor. The vehicle is capable of pure electric driving, but only at low speeds and for short distances. The biggest pioneer in this drive is Toyota. The other variant that has recently started to develop is a concept where the internal combustion engine serves as a generator and the electric motor takes care of the car's propulsion. Typical size of battery is about 1.5 kWh. [5]

Plug-in hybrid vehicle (PHEV)

This is the highest level of electrification of an internal combustion engine vehicle. In this configuration, the vehicle contains a significantly larger battery compared to the HEV. Due to the relatively large battery and the higher performance of the electric motors, operation on electricity only is possible. As the name suggests, the batteries of these vehicles can be charged from a charging stand, typically PHEVs have an on-board AC charger with an output of 7.2 kWh or less, but some models also have DC charging. The PHEV battery typically has 13-20 kWh. Often a vehicle configuration is used where an internal combustion engine is used to drive one axle and an electric motor drives the other, which can replace a 4x4 drive system. [5]

Electric vehicle (BEV)

These vehicles are powered purely by electricity from a battery. The battery is usually much larger than in the case of PHEVs. There are vehicles with capacities from 20 kWh to over 100 kWh on the market. Slower AC charging or fast DC charging is used to recharge the batteries. Vehicles also use regenerative charging to recharge batteries in use; some manufacturers offer such high levels of regeneration that there is no need to use the brake pedal at all, this regenerative mode is called One-pedal driving. [5]

Fuel cell electric vehicle (FCEV)

The difference with BEVs and FCEVs is that they carry with them energy in the form of compressed hydrogen, which they convert into electricity in a fuel cell, that mean that this type of vehicle is hybrid which using FCE instead ICE (internal combustion engine). At the same time, the vehicle contains a small battery to cover power spikes and to store the energy from the recuperation. The disadvantages are the complexity of storing and distributing the hydrogen, the almost non-existent network of filling stations, and the need for frequent maintenance of the air purification system. [5]

III. Batteries

All cars in previous section have common electric power train and battery. Difference is in size, construction and even battery technology. Most EVs today use Li-ion batteries, but there are still some HEVs on the market that use NiMH batteries. In this section will be focus on batteries for BEV. [5]

A. Current technology

Research on Li-ion batteries dates back to the 1970s, when British chemist S.Whittingham started researching them and created a functional but unstable cell. His research was then followed by J.B. Goodeoungh, who managed to achieve twice the potential. The last improvement was made by A.Yoshiro, who managed to increase safety and stability. They were then awarded the Nobel Prize in 2019 for their work. [6]
There are several cathode materials used in Li-ion batteries. Cathode material is made from combination of lithium and other metals. There is much more cathnode material in comparison with anode materials, which is typicali graphite and lithium titanate. The most used materials for automotive are LFP ($LiFePO_4$), NMC ($LiNi_{1-x-y}Mn_xCo_yO_2$), NCA ($LiNi_{1-x-y}Co_xAl_yO_2$) and LMO ($LiMn_2O_4$) in combination with graphite. Fig.2 shows the properties of each cathode material. [12]

The NCA battery is used in the NCA8155 ($LiNi_{0.8}Co_{0.15}Al_{0.05}O_2$) configuration and is currently used by Tesla in their vehicles. The battery has the highest energy density and very high theoretical capacity. [7] [8]

- **Potential vs. Li/Li$^+$**- 3.7 V
- **Teoretical capacity**- 279 mAhg^{-1}

The LFP battery is primarily used in electric vehicles used in China and for energy storage. They have the advantage of being free of heavy metals, long life and high reliability. [7] [8]

- **Potential vs. Li/Li$^+$**- 3.45 V
- **Teoretical capacity**- 170 mAhg^{-1}

The NMC battery is the most widely used battery type for electric vehicles. The current generation uses the NMC611 ($LiNi_{0.6}Mn_{0.2}Co_{0.2}O_2$) configuration and in 2022 has transitioned to the NMC811 ($LiNi_{0.8}Mn_{0.1}Co_{0.1}O_2$) configuration. The battery has the high energy density, very high theoretical capacity and high stability. [7] [8]

- **Potential vs. Li/Li$^+$**- 3.7 V
- **Teoretical capacity**- 280 mAhg^{-1}

LMO ($LiMn_2O_4$) is a material that was developed as a cheaper replacement for the first more powerful LCO ($LiCoO_2$) batteries. These batteries were used by e.g. Citroen Zero. These batteries have a low environmental impact and good stability. [7] [8]

- **Potential vs. Li/Li$^+$**- 4.1 V
- **Teoretical capacity**- 148 mAhg^{-1}

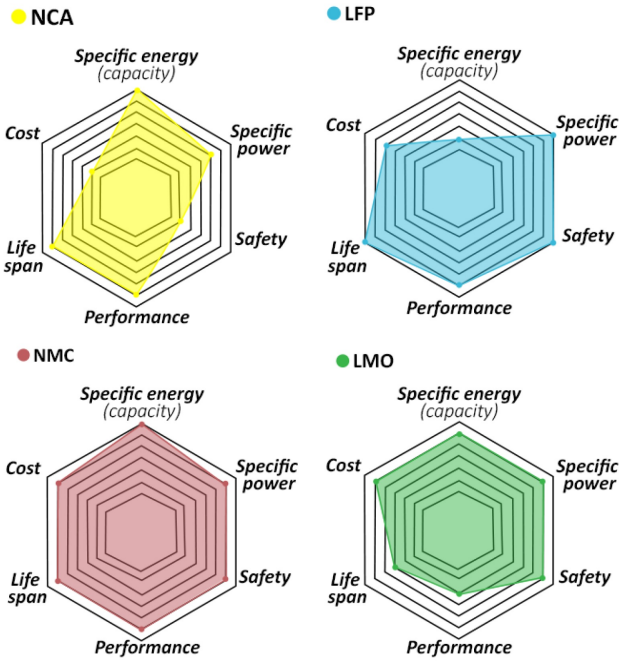


Fig. 2. Comparison of properties of individual cathode materials [8]

B. Future technology

The current research focuses in several directions. One is improving Li-ion technology, another is trying to find a better technology to replace Li-ion batteries, and another is the development of a cheaper equivalent to today's Li-ion batteries.

Nextgen Li-ion

Research on Li-ion battery enhancements focuses on solid-state electrolyte research and the use of silicon as an anode material. The use of solid electrolyte should improve the safety of batteries (no longer containing volatile solvents) and prevent internal short circuits caused by dendrites. The use of silicon as a composite with graphite or in pure form should increase the capacity of the battery because silicon has 10 times higher theoretical capacity (372 $mAhg^{-1}$ vs 3579 $mAhg^{-1}$). The biggest probem with silicon is its volume change durinch cycling (up to 300 %), which leads to a rapid capacity drop. [9]

Li-S batteries

The most researched post lithium-ion battery. It has been in research since the middle of the last century. The voltage between sulfur and Li/Li^{+} is between 1.7-2.5 V. Voltage depending on the modification of sulfur. Due to the theoretical capacity of S_8 1672 $mAhg^{-1}$ and pure lithium 3860 $mAhg^{-1}$, higher energy density than Li-ion batteries can be achieved. Another advantage is the possibility of 100 % discharge, whereas Li-ion only has 80 %. The biggest problem is the low lifetime and the associated large capacity degradation. Polysulphides are formed in the battery which react with the lithium, causing the shuttle effect. The use of a solid-state electrolyte to reduce this effect is currently being tested. [10] [11] [12]

Na-ion batteries

Research on Na-ion batteries started together with research on Li-ion batteries, but there are still some issues related to the use of sodium that has to be solved. A new anode material needs to be developed since graphite is not suitable for the larger sodium ions. Although Na-ion batteries have a lower capacity compared to Li-ion batteries, their undeniable advantage is the low cost of sodium, which will lead to a reduction in the price of batteries for less demanding applications. [12]

C. Construction of battery pack

The battery for an electric vehicle is one of the most expensive and most important part of the vehicle. The battery cells are built into modules and these modules then make up the battery pack.

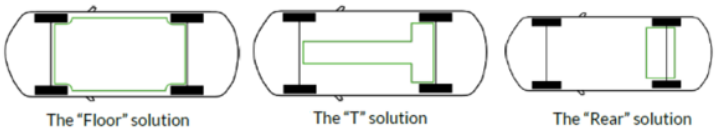


Fig. 3. Battery placement in BEV [13]

Regarding the integration of the battery pack into the car, different shapes and locations are used Fig.3. The fitting can also be related to whether the vehicle is designed directly as an electric vehicle or if it is fitting to an ICE vehicle. Pure electric vehicles most often use a ”skateboard” design, where the battery is placed between the wheels in the crew compartment and is incorporated as a rigid part of the cab. Another advantage is that the floor in the crew compartment is completely flat, which increases overall comfort. [5]
In vehicles converted from internal combustion engine vehicles, the most common use is a T-shaped battery, located under the rear seats and in the exhaust tunnel. This carries the disadvantage of a smaller battery and a step in the middle of the car floor. However, this battery shape is starting to disappear, as manufacturers are counting on the possibility of electric propulsion and the installation of a large battery in new generations of vehicles from the beginning. [5]
Another option is a battery in the rear of the vehicle (rear seat and trunk area). This option is most often used for HEV and PHEV variants, where the location of the battery reduces the usable capacity of the vehicle trunk. [5]

Each vehicle manufacturer designs the battery pack a little differently, but in general, all packs today contain battery modules, communication and power line harnesses and a battery management system (BMS). Additionally, the modules are equipped with direct heating or cooling.

The current trend is then to skip the modules and install the battery cells directly into the pack (cell-to-pack technology). This solution improves the weight-to-capacity ratio of the pack and reduces costs. [14] [15]

As mentioned above, the cells are placed in the battery modules. The shape and type of cell used then affects both the energy density in the module and the parameters that need to be regulated and controlled. In order to achieve the needed voltage and capacity, a series-parallel connection of the batteries is used.
The first type of cells are cylindrical cells. These cells have the advantage of zero volume change and good thermal management. The disadvantage may then be the worst utilization of the module space. [16]
The next type is prismatic cells. These are cuboid-shaped cells that are manufactured in various required dimensions. The cell suffers little volume change and its shape allows good module filling. The disadvantage may be a lower weight-to-capacity ratio than for pouch cells, but this is redeemed by better temperature control. [16]
The last type is the pouch cell. These are the only cells that do not use a rigid package but are encapsulated in a rigid foil. The advantage of these batteries is the highest weight-to-capacity ratio and the best use of module space. The disadvantage is then the most demanding thermal management, where the demand increases with the length of the battery. [16]

D. Battery management system

The BMS unit is a very important component of any battery. It is a device that serves as a kind of supervisory and

control system that oversees the individual modules or cells and keeps them in good condition.

There are several architectures of BMS units. The first is Centralized, so it consists of a single unit that controls the entire battery. The next is Modular, which uses a master unit as the controller and slave units to collect data from the modules. The last one is then Distributed, it also uses a master-slave hierarchy, but some of the operations are done by the slave units. Factors for selection tend to be system complexity, low cost, and the smallest amount of space needed in the battery pack. [5]

BMS has the following functions [5]:

- Monitoring and balancing module voltages and serial battery connections in moduls;
  - Active balancing is done by moving some of the energy from the battery with the highest potential to the battery with the lowest potential;
  - Passive balancing is done by discharging battery with the highest potential and converting it into heat [17];
- Control, management and charging;
- Correct connection to the charging station, when it performs the so-called ”handshake” when the charging instruction comes from the BMS;
- Monitoring the insulation voltage so as to avoid shock ;
- Temperature monitoring and regulation;
- Monitoring of the battery state of charge (SOC) and total battery capacity versus nominal (SOH);
- Communication with the rest of the vehicle.

IV. Charging

Charging is one of the most important phases of using an electric car. Charging process and components involved in it are specified by the international standard IEC 61851. [18]

Charging is then divided into several levels, depending on the type of current and power that can be used [18]

- **Level 1**-This is AC charging where the vehicle is directly plugged into an electrical outlet without any auxiliary equipment. The problem is the lack of any protection for the operator, as there is no intelligence involved in the charging and the whole process is only protected by the house circuit breaker. This method is banned in most parts of the world.
- **Level 2-** This is an AC charging from a normal socket using a cable containing an IC-CPD (In-Cable Control and Protection Device) unit that provides communication with the vehicle. The household charging is limited by the power of the socket to 3.7 kW (16 A).
- **Level 3**- This is the last variant of AC charging, which is offered through a home charger (wallbox) or a charging stand at the charging station. In Europe, the charging power of AC charging is limited to a maximum of 22 kW, typically using a three-phase power supply. However, most vehicles cannot use 22 kW charging because they are limited by the power of the on-board charger, which is typically 7.6 and 11 kW.
- **Level 4**- This is an ultra-fast DC charging, which bypasses the on-board charger so that the battery can be charged directly. These are purely public charging stations and are not intended for installation at the house. There are currently charging stations on the market with a maximum output of 550 kW at 800 V, which still have a power reserve for the currently available batteries.

A. On-Board charger

As mentioned in the previous section, the first three levels of charging are done using AC current, but this cannot go directly into the battery, but must first be converted by the on-board charger, which acts as an AC/DC converter and controls the charging. The on-board charger uses the CCCV (Constant current, constant voltage) method for charging. The power output is then limited by the number of phases it can use for charging, up to a maximum of 22 kW. [18]

B. Charging port

There is currently no uniform standard for the use of charging connectors. Table I shows the used charging connectors depending on the region and Fig.4 depicts the design of some representatives. All AC connectors are designed to use all three phases and are built for the same performance. However, this does not apply to connectors used for DC charging. Here the power limits are quite different. The European CCS-2 is built for 500 kW (to be increased), but for the example Chaoji (the latest version of the CHAdeMO connector) is certified for charging up to 900 kW. [19]

TABLE I
CONNECTOR TYPES USED IN THE WORLD [20]

	North America	Japan	China	Europe
AC	Type 1 Tesla	Type 1	GB/T (AC)	Type 2
DC	CCS 1 Tesla	CHAdeMO Tesla Chaoji	GB/T (DC) Tesla Chaoji	CCS 2



Fig. 4. Example of charging connectors [18]

As can be seen in Fig.4, one of the advantages of the CCS-1, CCS-2 and Tesla connectors is that they are also compatible with the AC variant of the connectors. This means that there is only one connector in the car for both AC and DC charging. Other standards require a separate connector for AC and DC, which increases space and wiring requirements. [18]

C. Vehicle-to-Everything (V2X)

V2X (Vehicle-to-Everything) technology is a communication system that allows vehicles to connect with their environment, including other vehicles, infrastructure, to optimize power consumption and charging (shows Fig.5). For its operation need bidirectional on-board charger and include next standards. V2G (Vehicle-to-Grid) technology enables electric vehicles to send power back to the grid, providing an efficient way to manage energy demand. V4G (Vehicle-for-Grid) technology uses electric vehicles as a storage solution for excess energy, which can then be used to power homes or businesses during peaks. V2V (Vehicle-to-Vehicle) technology is designed to enable cars to charge other cars from their battery. V2H (Vehicle-to-Home) technology is another system that enables electric vehicles to power homes during power outages or in places where is not possible to be connected to the grid. V2L (Vehicle-to-Load) technology is an emerging system that allows electric vehicles to power specific devices or equipment, such as tools or appliances. [21] [22]

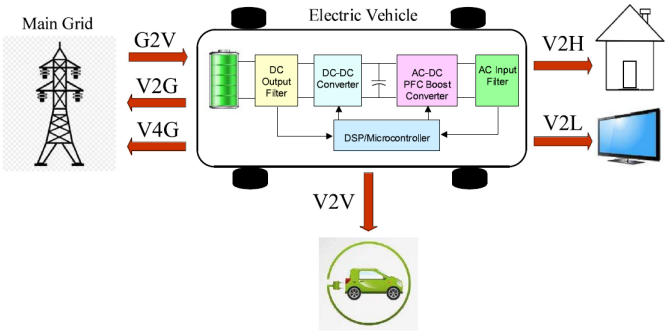


Fig. 5. Scheme of V2X technology [21]

V. Conclusion

The emission standards and environmental protection drive electrification of vehicles. Currently, BEVs use Li-ion batteries with various chemistry. There are some technologies under development that can make BEVs cheaper, with longer range, safer, and with faster charging. This can lead to an even faster transition to electric cars. One of the most critical parts is the BEV's battery pack, which contains battery cells, BMS, and other parts needed for the vehicle's safe operation. There are several methods to charge an electric vehicle with a wide range of charging power. There is a problem with the charging port because there is a lack of standardization, and almost each region uses different charging ports, which can slow the spreading of BEVs. On the other hand, there are technologies like V2X that can offer a promising solution for renewable energy sources and balancing the grid. This is a capability that ICE vehicles cannot offer.

Acknowledgment

This work was supported by the specific graduate research of the Brno University of Technology No. FEKT-S-23-8286.

References

[1] "Worth the Watt: A Brief History of the Electric Car, 1830 to Present", *Car and Driver*, 02018. .
[2] "History of the Toyota Prius", *Toyota UK Magazine*, 02015. .
[3] J. V. Mierlo, M. Berecibar, M. E. Baghdadi, C. D. Cauwer, M. Messagie, T. Coosemans, V. Jacobs, and O. Hegazy, "Beyond the State of the Art of Electric Vehicles: A Fact-Based Paper of the Current and Prospective Electric Vehicle Technologies", *MDPI*, vol. 2021, p. 27.
[4] "EV Company News For The Month Of February 2023", *Seeking Alpha*, 02023. .
[5] "Intelligence in Mobile Battery Applications (D5.1 Desk Research Data Analysis IMBA – Release 1)", *ALBATTS*, 02020. .
[6] "The History of the Lithium-Ion Battery", *ThermoFisher scientific*, 02019. .
[7] N. Nitta, F. Wu, J. T. Lee, and G. Yushin, "Li-ion battery materials: present and future", *Materials Today*, vol. 18, no. 5, pp. 252-264, 2015.
[8] R. M. Salgado, F. Danzi, J. E. Oliveira, A. El-Azab, P. P. Camanho, and M. H. Braga, "The Latest Trends in Electric Vehicles Batteries", *Molecules*, vol. 26, no. 11, 2021.
[9] J. Warner, *The Handbook of Lithium-Ion Battery Pack Design*. Elsevier, 2015.
[10] A. Gröbmeyer, J. Becking, P. M. Bieker, M. Winter, and M. C. Stan, "A Facile Preparation of S 8 /C Composite Cathode for Lithium-Sulfur Cells: Influence of Intrinsic and Extrinsic Cathode Properties on the Electrochemical Performance", *Energy Technology*, vol. 7, no. 4, Mar. 2019.
[11] M. Johnson-Groh, "Advances in lithium sulfur batteries pave the way for their commercial use", *Scilight*, vol. 2019, no. 32, Aug. 2019.
[12] *RD and Technological Perspectives for the Battery Sector (Desk research and data analysis for sub-sector IMBA - Release 2)*. 2021.
[13] R. Bisschop, O. Willstrand, F. Amon, and M. Rosengren, "Fire Safety of Lithium-Ion Batteries in Road Vehicles", *Research gate*, vol. 2019, p. 107, 2019.
[14] X. -G. Yang, T. Liu, and C. -Y. Wang, "Thermally modulated lithium iron phosphate batteries for mass-market electric vehicles", *Nature Energy*, vol. 6, no. 2, pp. 176-185, 2021.
[15] M. F. R. Zwicker, M. Moghadam, W. Zhang, and C. V. Nielsen, "Automotive battery pack manufacturing – a review of battery to tab joining", *Journal of Advanced Joining Processes*, vol. 1, 2020.
[16] P. Nur Halimah, S. Rahardian, and B. A. Budiman, "Battery Cells for Electric Vehicles", *International Journal of Sustainable Transportation Technology*, vol. 2, no. 2, pp. 54-57, Oct. 2019.
[17] W. Kurpiel, P. Deja, B. Polnik, M. Skóra, B. Miedziński, M. Habrych, G. Debita, M. Zamłyńska, and P. Falkowski-Gilski, "Performance of Passive and Active Balancing Systems of Lithium Batteries in Onerous Mine Environment", *Energies*, vol. 14, no. 22, 2021.
[18] A. Bahrami, "EV Charging Definitions, Modes, Levels, Communication Protocols and Applied Standards", p. 58.
[19] *European EV Charging Infrastructure Masterplan*. 2022.
[20] "CHARGING BASICS 102: Electric Vehicle Charging Levels, Modes and Types Explained — North America Vs. Europe Charging cables and plug types", *E-Mobility Simplified*, 02019. .
[21] S. Islam, A. Iqbal, M. Marzband, I. Khan, and A. M. A. B. Al-Wahedi, "State-of-the-art vehicle-to-everything mode of operation of electric vehicles and its future perspectives", *Renewable and Sustainable Energy Reviews*, vol. 166, 2022.
[22] J. Y. Yong, V. K. Ramachandaramurthy, K. M. Tan, and N. Mithulananthan, "A review on the state-of-the-art technologies of electric vehicle, its impacts and prospects", *Renewable and Sustainable Energy Reviews*, vol. 49, pp. 365-385, 2015.

Experimental breeding station for edible insects

Adam Dvorský
Department of Microelectronics
Brno university of technology
Brno, Czech Republic
xdvors12@vutbr.cz

Martin Adámek
Department of Microelectronics
Brno university of technology
Brno, Czech Republic
adamek@vutbr.cz

***Abstract*—The task of this work is to introduce the reader to the possibilities and conditions of breeding edible insects in the Czech Republic, with a focus on a selected species, namely the mealybug (Tenebrio molitor). To propose the optimization of breeding conditions for different stages of breeding, with the possibility of accelerating breeding. Based on these established conditions, a breeding station was designed and implemented, which was controlled by an ESP32 microcontroller from Espressif. The article describes the measurement and regulation of breeding conditions with a focus on temperature, humidity, dosing of food and liquids. The penultimate chapter mentions the experimental measurement of regulated conditions with the evaluation of the measured data. The conclusion is an evaluation of the solution and its other possibilities.**

Keywords—Edible insects; breeding; Mealworm; Tenebrio molitor; ESP32; regulation of conditions; sensors

I. INTRODUCTION (*HEADING 1*)

As the population increases, many aspects that affect human life become strained. At the beginning of this year (2023), the population exceeded 8 million inhabitants. With this trend, existential problems arise or only worsen. One can mention the problem with housing, this topic is already very problematic in some parts of the world today. The population tends to cluster in areas with a growing economy, work, etc. This further creates localized needs and shortages. The emergence of demand for new sustainable sources of food is therefore in place. A common breeding animal, such as cattle, has requirements for a large space for its successful life, requirements for feeding, a source of water. Along with the income side, it is also necessary to mention the expenditure side, when there are high levels of waste. Meat yield depending on these factors is not ideal. Over time, this resource will not provide food. In order to supplement the missing foods in the future market in the necessary nutritional breadth, new sources of food are being sought. Much attention is directed to edible insects.

At the same time, in many cultures around the world, the consumption of a variety of food sources, from the large animals that are commonly used for consumption, to the smallest, is the order of the day. Edible insects are among them.

Selected insect species, for example, contain a significant proportion of protein compared to conventional sources of poisoning, which are an integral part of the daily diet. Edible insects, compared to the cattle, abound in many benefits. The energy costs of breeding are small. Furthermore, we can mention the large difference in the produced nutrient materials in terms of breeding area, food requirements, and the biggest difference can be found in the produced waste materials. The European Union has already taken steps to legalize the breeding of edible insects for human consumption.

Edible insects are a large percentage of pests. His demands for life are not very high. By regulating some conditions for life, it is possible to slow down or speed up the existence cycle. For this reason it was necessary to design and create an experimental breeding station in which various parameters can be regulated.

II. DESIGN

As wos written, edible insects are often pests. His normal living conditions are not demanding. The Mealworm (lat. Tenebrio molitor) was chosen as the primary insect for the work. Normal conditions for his life are standard room values. Temperature in the range of 18-20 °C, humidity 30-60 %. According to sources, the temperature should not exceed 35 °C, when there is a high death rate of breeding. By adjusting these two mentioned parameters, it is possible to change the time, the so-called life cycle. Which is the time between egg hatching, worm stage (they are used for protein), maturing into a beetle, and then laying more eggs. By changing the composition of the diet, the nutritional share of the obtained food is also changed. Only the adult worm stage is allowed for human consumption, Figure 1.

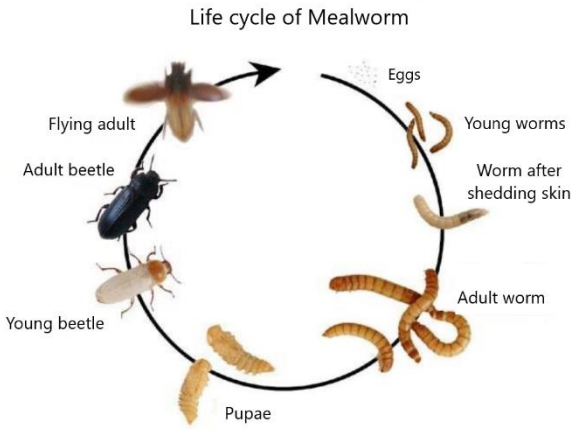


Fig 1. Life cycle of Mealworm

Based on these requirements, an experimental facility is designed, shown in Figure 2. To ensure a separate breeding area, the breeding is placed in a closed experimental box. The breeding itself is placed in a food crate 60 x 40 x 20 cm (standard E2). Placed in this farm is chosen based on the possibility of easy handling. This crate is easily removable, interchangeable with other breeding, and allows for easy cleaning and disinfection. The overall box is made of a transparent material, Plexiglas, for easy visual inspection of the breeding and individual components of the device. The box includes inlet and outlet openings with a fan, which brings fresh air into the rearing area and discharges waste gases to the space behind the experimental box, where it is ideal

to place a hood to remove gas from the room. It is because of the spread of diseases. There can be several of these experimental breeds in the room, and in the case of an ongoing disease, not all breeds are infected. The device enables regulation of the temperature of the breeding substrate using a heating pad with a power of 14 W with a maximum temperature reached in the room of 42 °C. This temperature is sufficient for the needs of the chosen breed. The heating is realized by a terrarium pad intended for heating by infrared radiation due to the unevenness of the crates. This pad is insulated and easily washable for possible disinfection of the breeding area. The box also enables humidity regulation by supplying water to the breeding area and its spontaneous evaporation under the influence of increased temperature. With the help of humidity and temperature sensors located in the box, the current value determined as feedback for regulation is sensed. The temperature sensors for the breeding crate are located in the holding rails.

Any insect needs a suitable temperature and humidity for its proper development. As a result of too low humidity, the breeding product can lead to death. Conversely, excessive humidity can lead to degradation of the substrate, namely the formation of mold, which also leads to the death of the breeding. For humidity regulation, it is therefore necessary to measure the humidity of the air in the proximity where the breeding product is located. The humidity value will always depend on the ambient temperature. As the temperature increases, so does the humidity due to faster evaporation of the absorbed moisture from the substrate and material of the breeding station. The ventilation of the box also disturbs the humidity. Increasing the humidity in the box is possible in two different ways. The first, which is not covered in this chapter, is by increasing the temperature. In this case, higher humidity in the breeding environment is created by higher evaporation of moisture condensed in the substrate. The second, solved, method is the supply of water to the box, when the spontaneous evaporation of water occurs under the influence of a higher temperature. Water can be supplied to the system in two ways. The first is the connection to the water supply network, located in the building. This solution is suitable for large industrial use. In the event of a malfunction, at best only the given breeding box may be heated, at worst the entire breeding room may be heated. For this reason, water supply from a bottle container is chosen. This solution eliminates the destructive disorder. The disadvantage is the regular check for refilling this tank. A hose line will be led from this reservoir towards the peristaltic pump, which then leads towards the reservoir located in the box. The water will then spontaneously evaporate from this tank, which will increase the humidity. There is a built-in level sensor to check the overflow of the container. This will allow a certain amount of water to be supplied. Thanks to the use of a bidirectional peristaltic pump, it is possible to suck the water from this tank and return it to the external tank, which makes it possible to slow down and stop the increase in humidity.

Filing of drinkable water for animals is realized by supplying water to the breeding environment. In the box, there must be a porous stone under the outlet of the water supply, which absorbs water drops into its broken surface, thus eliminating the drowning of breeding individuals. The supply of loose feed is implemented using a feed dispenser designed again for terrarium needs and aquaristics. The feed is varied for different species and stages of development. To automate the entire environment, a container and a solid feed dispenser (granules, etc.) were designed. It is necessary to pay attention to the suitability and safety of the selected type of feed. Feeds that can degrade quickly and those that release excessive amounts of juice and liquid are not suitable. Such unsuitable feed can degrade over a long period of time and form molds that can destroy the entire farm. In smaller quantities, when consumption occurs faster than degradation, it is permissible.



Fig 2. Visualization of the experimental breeding station

The control unit has the task of managing the breeding, regulating the conditions of the components and passing on information to the authorized person. According to the specification, the ESP32 circuit from Espressif is placed and used for this purpose in the proposed device. It is a widely used microcontroller that includes many supporting systems such as Bluetooth, wifi, internal memory, timers and more. Everything is supported by sufficient documentation and instructions. The used parts will be described in the work below. Programming the ESP32 circuit can be done in many programming languages that support conversion for a given ESP32 architecture. MicroPython, C and the Arduino IDE can be used, for example. These languages support the use of libraries and functions where it is possible to write code without detailed knowledge of how the module works. The disadvantage is subsequently the comprehensiveness and size of the program. These parameters enable system robustness, versatility, reliability and application in a wide range of applications. The module contains supporting modules for remote wireless communication. For less demanding applications, it is possible to use Bluetooth communication. For other applications, communication via WIFI is used. The ESP32 module (Figure 2.1) is capable of several modes of working with this communication. The basic two modes are connection to an already existing network and creation of an access node (Access Point), when a custom network is created. This makes it useful in mobile applications and Internet of Things (IOT). By supporting many communications, the ESP32 module system is robust and supports many test and development connections.

The selected controller can be assigned to the SoC (system on chip) group, when all modules and connections are implemented on one circuit carrier. It was put on sale in 2016.

This unit records any changes in the environment and stores the data in its memory. This data can be downloaded from the control panel to an SD card or by connecting to a WIFI network. Entering values for regulation, which are read out further, is only possible via a wireless network. For this box, it is possible to enter the temperature and humidity of the breeding, when it is only possible to increase the temperature. Furthermore, the interval and dose of feed and fluids, ventilation and scanning of measured data for further processing.

Mains voltage 230 V is used as the input voltage, which is used to power the heating elements via a switching relay. This voltage is further transformed and stabilized to 12 V, which is again used as a

supply voltage for other components. Namely, two peristaltic pumps, a fan and a stepper motor for the feed dispenser. The control and sensor circuits are powered by a voltage of 3 V, so it is necessary to stabilize this voltage as well. The ESP32-WROOM32 module is used as the basic control unit. A display is connected to this control unit, which will enable the setting and reading of values. A rotary encoder is located here for moving around the display and other functions. There is an SD card reader for data recording. The circuit includes a connection with the RTC DS3231 as a source of the clock signal. Last but not least, the circuit will contain control elements for the feed dispenser, fan and for controlling the peristaltic pumps. Connectors will further emerge from the circuit, which will be connected to switching relays that will switch the heating elements.

III. TESTS

A. *Calibration of temperature reading for regulation*

Figure 3 shows a test measurement of temperature regulation and the effect on humidity in the breeding environment. For the test measurements, a reference sensor was placed in the breeding area where the tested insects will be found. The regulation was set to increase the temperature by 1 °C in the breeding area (green course). This temperature is only a reference for calibrating temperature sensors for feedback control. These sensors are located in the slats that hold the food container (dark blue). Due to the thermal capacity of the material used and the slow conduction of heat from the crate to the support bars, the measured temperature is lower and gradually copies the reference temperature. Based on this data, a conversion table was found for the correct temperature setting in the breeding area. Here, a large influence of the temperature in the room environment is promoted, when a slow increase is observed at night, on the contrary, a rapid increase in temperature is observed during midday. For this reason, it is recommended to place the experimental box in a room with a constant temperature and then repeat the temperature calibration.

It is also possible to observe the air temperature inside the box on the graph (light blue curve). This course copies the observed temperature in the supporting rails of the crate. Here there is no cooling from the bottom side of the experimental box.

The final orange curve demonstrates the effect of temperature on humidity in the experimental box. This is a measurement of relative humidity. In general, it can be said that as the temperature increases, the proportion of bound moisture in the air increases. When increasing the air temperature in the experimental box, it is noticeable that the air began to bind more water, thus increasing the relative humidity. When the temperature exceeds 33 °C, there is a noticeable decrease in the relative humidity in the monitored process. When converted to absolute humidity, the humidity is still increasing slightly.

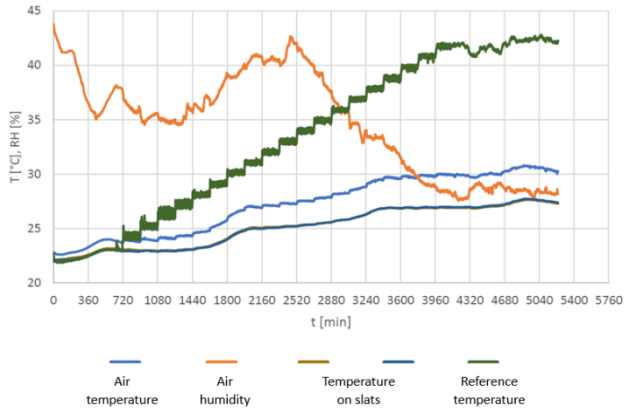


Fig 3. Test measurement of temperature control

B. *Experimental humidity control*

Measurement of humidity regulation and ventilation took place in three phases. These phases are described by the graph in Figure 4. Measurement of maximum humidity and the influence of ventilation with temperature. In the first phase, humidity was increased when 40 ml of water was added to the beaker inside the experimental box. The experiment showed only a small increase in relative humidity in the box. Humidity can be increased with the combination of adding water to the substrate to feed the insects. After 600 minutes forced ventilation was switched on with a fan at maximum speed. The relative humidity decreased slightly. After this phase, the water was again returned to the system to a level of 40 ml. After filling, the heating of the substrate was started. Here we observe the same phenomenon as in the previous test.

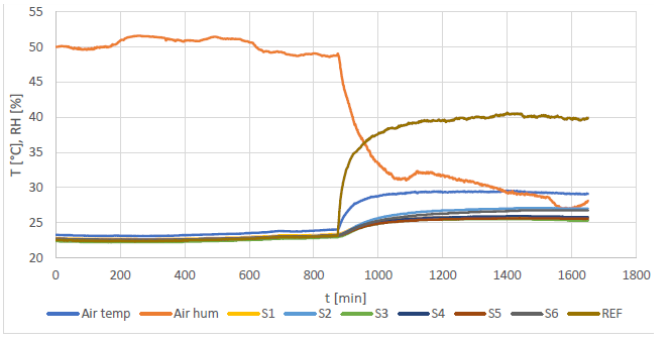


Fig 4. Humidity regulation test

C. *Experimental temperature control*

This measurement was made for a set temperature of 30 °C. The measurement took place with a reference sensor placed in the inner part of the breeding substrate. The set average temperature of the sensors intended for regulation was 25 °C, see figure 5. The resulting temperature fluctuated within ± 0.8 °C.

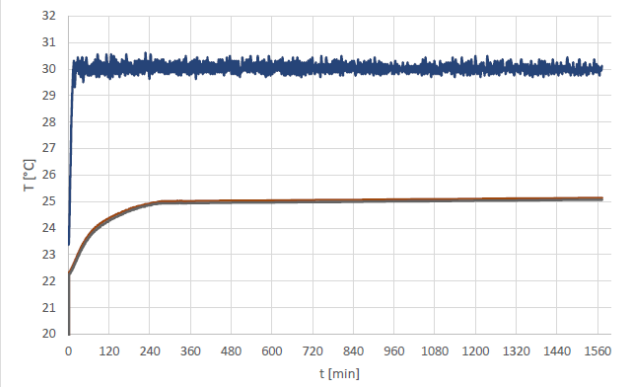


Fig 5. Temperature regulation test

IV. CONCLUSION

In 2021, the European Food Safety Authority (EFSA) authorized and approved a new food application where dried Mealworm larvae (lat. Tenebrio molitor) are allowed to be sold on the European market for human consumption. The Czech Republic issued a general breeding methodology in its decree. However, the specific optimal breeding steps are not publicly known and are part of the know-how of breeding companies.

The goal of the work was therefore the creation of an experimental breeding box for the breeding of edible insects, which would allow not only automatic regulation of temperature and humidity, but also

automatic dosing of feed and water, and which would allow good handling of the breeding.

The work summarizes basic information about the species of mealybug and its development. Breeding conditions are also generally summarized in this section. The practical part describes the design and implementation of a new type of experimental breeding box, in which it is possible to insert a food crate, in which the breeding of a certain type of insect will be placed. In this way, a rapid change of breeding can be achieved. The conditions for cleanliness and disinfection have also improved for this reason – only the transport food box needs to be cleaned.

The newly designed breeding box enables the measurement and regulation of basic breeding conditions, namely temperature and humidity. To measure the temperature of the substrate, there are two heating elements on the bottom side and a set of temperature sensors for regulation feedback. When measuring the temperature and humidity of the air, it is possible to regulate the level of humidity in the environment with the designed and implemented system. For this regulation, however, it is necessary to take into account the values of the surrounding environment, which significantly influence the conditions. Excessive heat and humidity in the breeding room cannot be reduced. An integral part of the design is also ventilation, which is a necessity for successful breeding.

The design is also equipped with a system for refilling water and feed. This is created using a hose line system and a peristaltic pump. A stepper motor-driven feed dispenser was designed and created for this box using 3D printing technology. This enables the dosing of a dry bulk mixture with a defined dose. However, larger feed particles can cause the paddles that move the feed from the hopper to the rearing area to jam.

A wireless LAN connection using WIFI communication was designed and used for communication. Parameter setting is also possible using the control panel on the top of the box. For this system, according to the catalogs and the recommended connection of elements, a circuit connection with selected components was designed and created.

The manufactured parts were tested with experimental measurements, the purpose of which was to verify the conditions and specify the settings of the created service program for controlling the regulation. The maximum temperature that was reached in the temperature measurement is 42.8 °C.

A control manual for the operator was created as part of the work, which aims to familiarize the user with the control and setting of individual parameters.

With the designed and implemented experimental breeding station, it is thus possible, for example, to verify the breeding parameters stated by companies, but also to obtain new information for increasing the yield of breeding not only in the developmental stage of the mealybug larva, but also for other stages and other insects.

The experimental box is made with handles located on the front and back for easier handling and transfer of the breeding box. The constructed breeding box is shown in Figure 6.



Fig 6. Finished product

ACKNOWLEDGEMENT

The arcticle was supported by project no. FEKT-S-20-6215, New approaches in the use of modern micro- and nanoelectronics.

REFERENCES

[1] DVORSKÝ, Adam. Inovace zařízení pro chov jedlého hmyzu. Brno, 2022. Dostupné také z: https://www.vutbr.cz/studenti/zav-prace/detail/139277. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav mikroelektroniky. Vedoucí práce Martin Adámek.

[2] Elhassan M., Wendin K., Olsson V., Langton M., Multidisciplinary Digital Publishing Institute, Quality aspects of insects as food-Nutritional, sensory, and related concepts. [online], 2019,[cit 12.10.2021], ISSN 23048158, dostupné z https://www.scopus.com/record/display.uri?eid=2-s2.0-85063266760&origin=reflist&sort=plf-f&src=s&sid=dae63d48c153e7781382c913d5920f&4b&s2stbTITLE+ech-KEYl-KEYl-CZE jedlý+hmyz%29

[3] Grau T., Vilcinskas A., Joop G., Walter de Gruyter GmbH, Sustainable farming of the mealworm Tenebrio molitor for the production of food and feed, [online], 2017, [cit 12.10.2021], ISSN 09395075, dostupné z: https://www.scopus.com/record/display.uri?eid=2-s2.0-85023778479&origin=reflist&sort=plf-f&src=s&sid=1930759a6371ddbc19e87ec31c5d8755&s%sd8ABBTITLEl+KEYl=6ebbTITLEl molitor+životnost%29

Valeo

VALEO.JOBS.CZ



A taky Vašek, Ondra, Lukáš, Eva, Olga... Ti všichni táhnou za jeden provaz, kterým přitahují novinky do automotive světa. Za každou naší inovací stojí konkrétní lidé, kterým Valeo vděčí za svůj úspěch. Díky nim mohou být auta autonomní, elektrická, komfortní i bezpečná. Postavte se také za svou myšlenku, kterou za pár let řidiči ocení v provozu.

Koukni na valeo.jobs.cz nebo naskenuj QR kód.

SMART TECHNOLOGY FOR SMARTER MOBILITY

Suppresion of FET gate parasitics using transient gate overvoltage

Pavel Tomíček
Department of Microelectronics
Brno University of Technology
Brno, Czech Republic
xtomic05@vutbr.cz

Jaroslav Boušek
Department of Microelectronics
Brno University of Technology
Brno, Czech Republic
bousek@vut.cz

Abstract - This **paper presents a method of driving field effect transistors that is capable of suppressing the effect of parasitic elements. These parasitics slow down the switching speed and provide a fundamental limit on the maximum switching speed of the FET transistor. This limit can be improved using overvoltage gate drivers, which provide a short pulse on the gate with amplitude higher than the maximum safe value specified by the manufacturer of the transistor.**

Keywords - FET, parasitic elements, overvoltage gate driver

I. Introduction

The switching speed of field effect transistors is important in almost all areas of electronics, but the most prominent fields that benefit greatly from improving performance are probably digital and power electronics. Here, the effects are increased processing speed and smaller and more efficient switching converters. Another area is instrumentation for experimental physics, where devices such as Pockels cell and particle accelerators require high voltage pulse power supplies with subnanosecond rise and fall times [1, 2].

II. Analysis

To turn the field effect transistor on, it is necessary to charge its gate capacitance. The rate at which it is charged is fundamentally limited by the parasitic gate resistance and inductance. Also, in a real circuit there are more components such as the driver and the PCB. These can further slow the switching speed. In practical design, it is therefore often necessary to optimise all three components to reach the desired switching speeds.

Optimalization of the transistor for maximum switching speed consists of reducing parasitic gate resistance and inductance and minimising gate capacitance. The first two elements are parasitic, and it is usually desirable to minimize them as much as possible; however, the gate capacitance is necessary for the function of the transistor itself. This is because the drain-source on-resistance $R_{DS(on)}$ is inversely proportional to it. Therefore, reducing the gate capacitance increases $R_{DS(on)}$. But $R_{DS(on)}$ is also inversely proportional to carrier mobility, therefore, by using modern wide-bandgap materials with higher carrier mobility and critical electric field, such as SiC or GaN, it is possible to keep the product of carrier mobility and gate capacitance constant. Another advantage is the shorter channel length due to a larger critical electric field, which also decreases gate capacitance. The result is that with a smaller gate capacitance, the $R_{DS(on)}$ did not change or even decreased. [3, 4]

Parasitic gate resistance and inductance are caused by the gate electrode itself and by the packaging of the device. These parasitics are often simplified in the parasitic model and are simply presented as a single resistor and inductor. Their reduction can be achieved by choosing modern transistors in packages optimised for low lead inductance and resistance, such as DFN or WLCSP. These packages have several downsides compared to older types, such as TO-220 or TO-263-7. Some of the downsides are reduced clearance, creepage, and lower power dissipation. These limit their usage to low voltage and low power applications. [3, 5, 6]

Conventional gate drivers use a push-pull output structure. They are usually made of two FETs or two bipolar transistors, where the FETs are connected as common source amplifiers and the bipolar transistors are connected as emitter followers. These configurations are in Fig. 1.

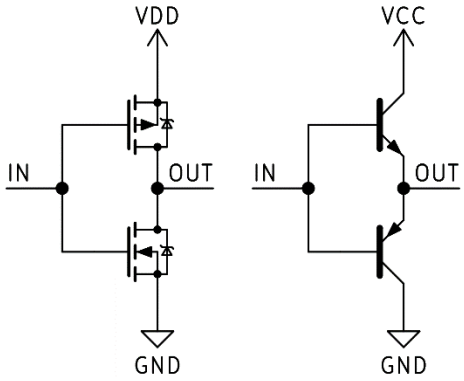


Fig. 1. FET and bipolar-type output stage of conventional gate drivers

These drivers are simple to use and widely available. However, the output voltage is limited by the power supply rails. This poses a restriction on the switching performance because even if the output voltage of the driver increases to the power supply voltage instantly, the rise time of the voltage on the gate of the transistor connected to the output is limited by the parasitic resistance and inductance. Therefore, it is not possible to further improve the switching speed of the output transistor using these conventional drivers.

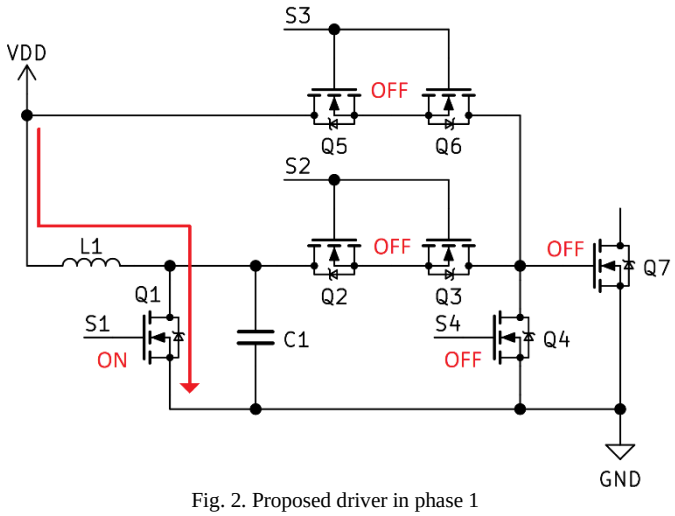


Fig. 2. Proposed driver in phase 1

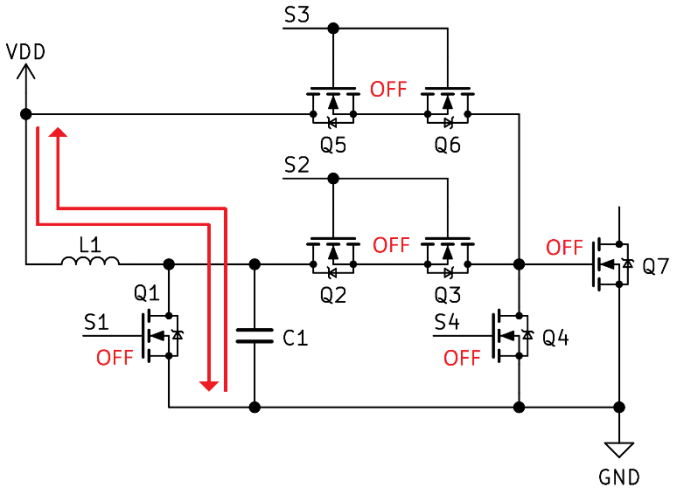


Fig. 3. Proposed driver in phase 2

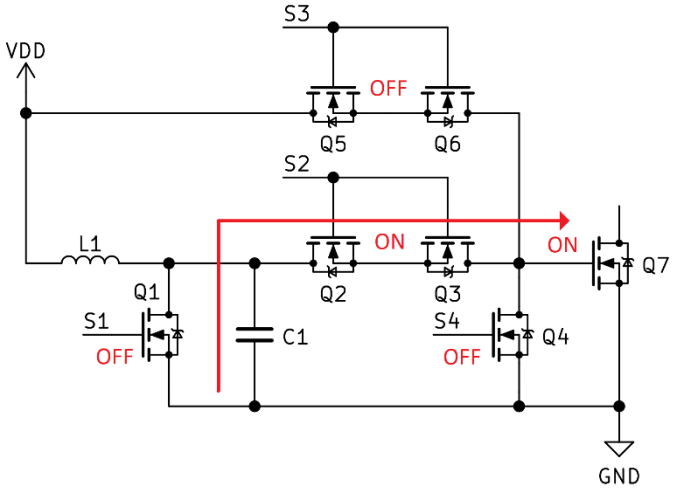


Fig. 4. Proposed driver in phase 3

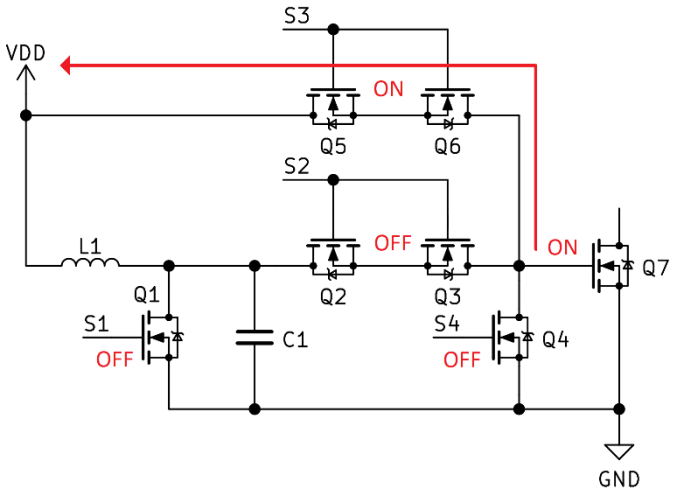


Fig. 5. Proposed driver in phase 4

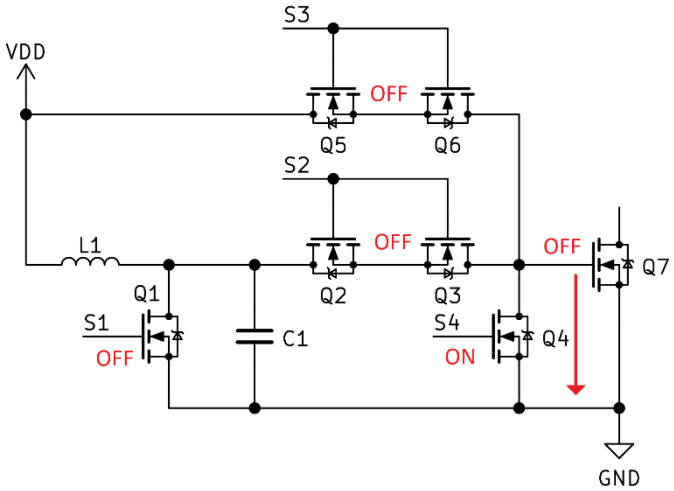


Fig. 6. Proposed driver in phase 5

III. Proposed Driver

The proposed driver uses a resonant LC tank to generate high voltage to charge the gate capacitance. The principal schematic of the driver is in Fig. 2 to Fig.6.

A. *Principle of operation*

The driver works in several phases. Their details are discussed as follows. All phases are represented in Fig. 2 to Fig.6.

Phase 1 [t_0-t_1]: Prior to this phase, only Q4 is turned on keeping the output transistor turned off. In this phase, Q1 is turned on at time t_0 and the current through the inductor L1 starts to rise linearly. Once the current reaches the desired value at time t_1, Q4 is turned off. This phase is one of the downsides of the proposed driver, as it introduces a relatively large propagation delay, in this case 100 ns.

Phase 2 [t_1-t_2]: The sudden turning off of Q1 induces resonance of the LC tank made of L1 and C1 with the frequency f_{RES} calculated by

$$f_{RES} = \frac{1}{2\pi\sqrt{L_1 C_1}} \quad (1)$$

For time t_1 *to* t_2, which is equal to a quarter of a period of f_{RES}, all transistors are turned off. During this time, resonance in the LC tank creates harmonic oscillations with amplitude greater than the supply voltage.

Phase 3 [t_2-t_3]: Once t_2 is reached the voltage on C1 reaches maximum value and Q2 and Q3 are turned on, which connects the LC tank to the gate of Q7, which starts charging from the generated high voltage.

Phase 4 [t_3-t_4]: After a few nanoseconds time t_4 is reached, Q2 and Q3 are turned off, while Q5 and Q6 are turned on and connect the gate of Q7 to a steady-state voltage. This causes the high voltage on the gate to rapidly decrease to a safe value and dampens any oscillations on the gate that may be caused by parasitic elements. Transistor Q7 is now fully switched on and can be turned on for an arbitrarily long time.

Phase 5 [t_4-t_5]: Once it is required to turn Q7 off, Q5 and Q6 are turned off and Q4 is turned on, shorting the gate of Q7 to its source, causing the gate voltage to decrease to zero. The driver is now prepared to turn on Q7 once required.

IV. Prototype

Transistors Q1- 6 are of type SSM6K361NU,LF by Toshiba and are driven two by ADuM6132 gate drivers by Analog Devices. The nominal value of the inductance of L1 is 100 nH and the capacitance of C1 is 10 nF. Due to the chosen parts, the resonant frequency f_{RES} is equal to 5 MHz. The picture of the whole driver is in Fig. 7. The driver is made of two boards. The first board implements the driver itself and its power supplies. The second board is an FPGA evaluation board that provides the timing signals to the first boards.

The driver can be triggered either by a button or from an external signal generator through an SMA connector.

A. *Timing generation*

Timing necessary for driving the four transistor switches Q1, Q2-3, Q4, and Q5-6 was generated using a four-channel digital delay generator with 5 ns timing precision implemented in a LCMXO3D-9400HC-5BG256C FPGA manufactured by Lattice Semiconductor. This allowed for precise adjustment of the driving signals with a resolution of 5 ns. In Table 1 the duration of the key phases 1 to 3 of the driver is presented. Other phases can be arbitrarily long and, therefore, their specific time is not important.

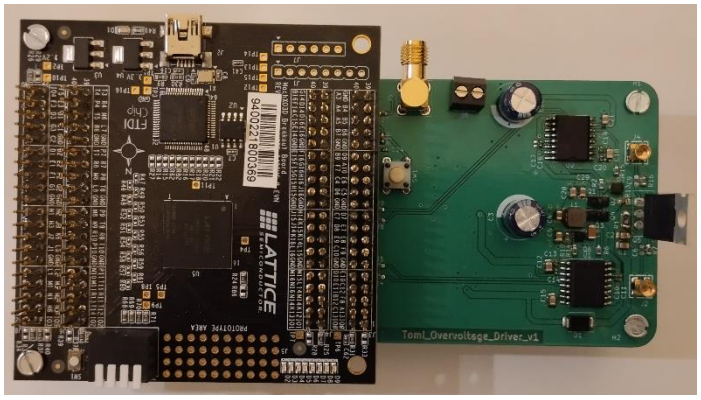


Fig. 7. Prototype of the proposed driver

Table 1 Duration of phases 1 to 3

Phase	Time [ns]
1	100
2	70
3	15

These durations were experimentally verified to provide optimal performance. For example, phase 2 was extended by 20 ns compared to a theoretical value of 50 ns. This is expected as component tolerances and parasitic capacitances caused by transistors Q1 and Q2 will affect the resonant frequency f_{RES}.

V. Test and Results

All measurements were performed using an oscilloscope Rigol DS1104Z with 100 MHz bandwidth and the gate driver was powered by a 15 V power supply. The first measurement was performed without the output transistor Q7 connected to test the theoretical maximum pulse amplitude delivered to the gate of Q7. The waveform captured by the oscilloscope is in Fig. 8.

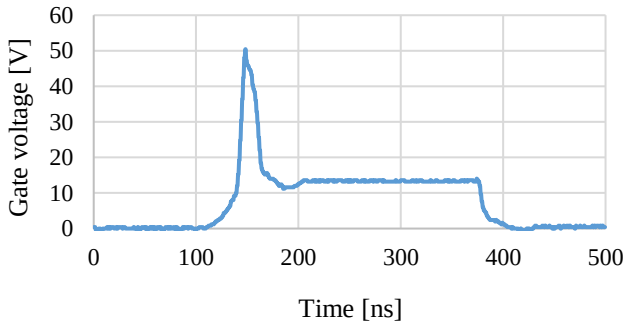


Fig. 8. Output voltage of the driver without connected output transistor Q7

From Fig. 8. it is clear that positive overvoltage with amplitude of 50 V is generated. It lasts for around 20 ns, before dropping to a steady-state value of 13,6 V.

Further measurements were made with transistor IRFB7437PBF by Infineon Technologies connected as Q7. The selected parameters of the transistors are in Table 2.

Table 2 Selected parameters of the output transistor Q7 IRFB7437PBF

Parameter	Value
Total gate charge	150 nC
Input capacitance	7330 pF
Drain breakdown voltage	40 V
Continuous drain current	195 A
Gate-to-Source Voltage	±20 V
Gate resistance	2,2 Ω

With this transistor connected, several tests were performed. These consisted of measuring the gate voltage and the drain voltage while driving using the gate overvoltage and without it. For the test without the gate overvoltage, phases 1 and 2 were skipped. This resulted in capacitor C1 being charged to steady state voltage when t_3 was reached. The theoretical value of the gate voltage V_{GS} once Q2 and Q3 are turned on can be determined by

$$V_{GS} = V_{C1t3}\frac{C_{iss}}{C_{iss} + C_1} \quad (2)$$

V_{C1t3} is the voltage to which capacitor C1 is charged when Q2 and Q3 are turned on. For overvoltage driving, that is 50 V as was measured in Fig. 7. For conventional driving, V_{C1t3} is equal to the steady-state value of 13,6 V. Capacitance C_{iss} is the input capacitance of the output transistor Q7. The theoretical values of V_{GS} once Q2 and Q3 turn on are 21,4 V for overvoltage driving and 5,8 V for conventional driving. The measured waveforms are shown in Fig. 9.

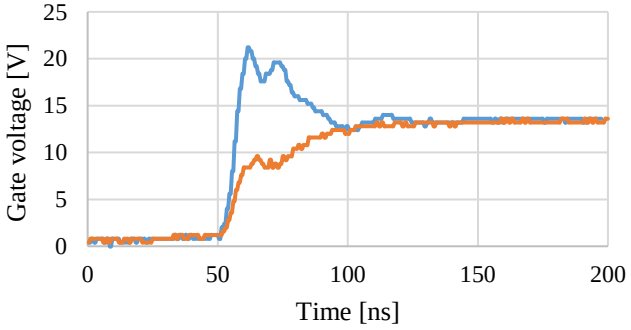


Fig. 9. Driver output voltage with the output transistor connected. With overvoltage driving (blue) and conventional driving (orange).

The peak voltage achieved with overvoltage driving is 21 V, which matches the value calculated using Equation 2. When overvoltage driving, the gate capacitance is charged faster, and therefore the transistor is turned on faster. This was observed directly by measuring the voltage of the drain of Q7. This measurement is shown in Fig. 10.

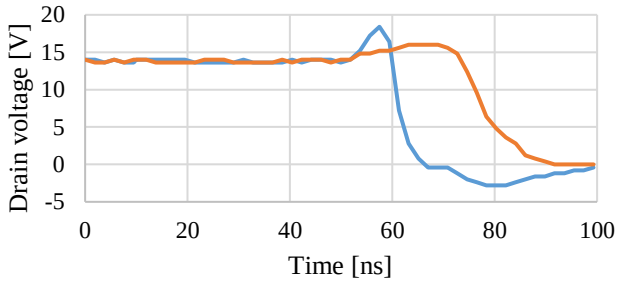


Fig. 10. Drain voltage of transistor Q7. With overvoltage driving (blue) and conventional driving (orange).

Using overvoltage driving, the delay from applying gate votlage to the transistor turning on is 16 ns shorter compared to conventional driving. The fall time of the output voltage is reduced from 12 ns to 4 ns.

VI. Conclusion

In this paper a method for faster driving of field effect transistors is presented. This method relies on the application of a high voltage pulse to the gate of the driven transistor. This suppresses parasitic inductance and resistance and charges the gate capacitance faster. The performance of the proposed driver was measured on a prototype PCB that is capable of overvoltage and conventional driving. Using the gate overvoltage, it was possible to reduce the delay between applying voltage to the gate and the transistor turning on by 16 ns compared to conventional driving. The drain voltage fall time was reduced from 12 ns to 4 ns using overvoltage driving.

Acknowledgment

The article was supported by project no. FEKT-S-23-8162, Modern micro- and nanoelectronics for future.

References

[1] M. LaPointe, L. Collier, T. Kajiwara, J. Dickens, J. Mankowski and A. Neuber, "Improving Fast SiC MOSFET Switching Using an Inductive Gate Drive Approach," 2019 *IEEE Pulsed Power & Plasma Sci. (PPPS)*, Orlando, FL, USA, 2019, pp. 1-4, doi: 10.1109/PPPS34859.2019.9009986.

[2] H. -S. Kim, C. -H. Yu, S. -R. Jang and G. -H. Kim, "Solid-State Pulsed Power Modulator With Fast Rising/Falling Time and High Repetition Rate for Pockels Cell Drivers," in *IEEE Trans. Ind. Electron.*, vol. 66, no. 6, pp. 4334-4343, June 2019, doi: 10.1109/TIE.2018.2866078.

[3] E. A. Jones, F. F. Wang and D. Costinett, "Review of Commercial GaN Power Devices and GaN-Based Converter Design Challenges," in *IEEE Trans. Emerg. Sel. Topics Power Electron.*, vol. 4, no. 3, pp. 707-719, Sept. 2016, doi: 10.1109/JESTPE.2016.2582685.

[4] A. Seidel and B. Wicht, "Integrated Gate Drivers Based on High-Voltage Energy Storing for GaN Transistors," in *IEEE J. of Solid-State Circuits*, vol. 53, pp. 3446-3454, Dec. 2018, doi: 10.1109/JSSC.2018.2866948.

[5] K. Shenai, "Future Prospects of Widebandgap (WBG) Semiconductor Power Switching Devices," in IEEE Trans. Electron. Devices, vol. 62, no. 2, pp. 248-257, Feb. 2015, doi: 10.1109/TED.2014.2360641.

[6] K. Umetani, K. Aikawa and E. Hiraki, "Straightforward Measurement Method of Common Source Inductance for Fast Switching Semiconductor Devices Mounted on Board," in *IEEE Trans. Ind. Electron.*, vol. 64, no. 10, pp. 8258-8267, Oct. 2017, doi: 10.1109/TIE.2017.2694

Design Adaptation of a 5-Phase Outer Rotor PMSM for Faulty Modes of Operation with Open Phases

1st Vitaliy Sizonenko
Department of Power Electrical and Electronic Engineering
Brno University of Technology
Brno, Czech Republic
vitaliy.sizonenko@vut.cz

2nd Petr Hutak
Department of Power Electrical and Electronic Engineering
Brno University of Technology
Brno, Czech Republic
hutak@vut.cz

***Abstract*—In this paper, a design adaptation of a 5-phase outer rotor permanent magnet synchronous machine for the operation with open phases is presented. The designed machine has limits on the outer diameter and permissible electromagnetic loadings. The paper will show that the fault-tolerant machine has higher permanent magnets, wider air-gap, lower number of turns in the stator winding and increased core length in comparison to the regular-design machine. Changes introduced to the regular-design machine result in the rise of its mass and decrease of the stator resistance and inductance, while keeping supply limits of the machine as well as its rated torque and speed constant.**

***Index Terms*—electrical machine, fault-tolerant, permanent magnet synchronous machine (PMSM), fractional slot concentrated nonoverlapping winding (FSCW), tooth-coil winding (TCW), outer rotor.**

I. Introduction

Common approaches during the design of a fault-tolerant (FT) permanent magnet synchronous machine (PMSM) include usage of 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 [1]. Furthermore, higher-phase solution is not only capable to bring FT properties into a machine, but also increase its power density, if this is relevant for the given application. For example, in [2] it was shown, that for the given outer rotor dimensions, only one optimal tooth-coil winding topology exists from the point of view power/mass density. Still, despite the fact that for the given requirements the 6-phase machine has shown higher power density among the considered designs, current paper continues development of the star-connected 5-phase 40-slots 36-poles PMSM, due to the local availability of the 5-phase supply converter. Selected machine has a single-layer fractional-slot concentrated nonoverlapping winding, its more detailed description is given in [1] and [3]. Such a winding has coils wound around each second tooth of the stator, and in the rest of this paper is called single-layer tooth-coil (SLTC) winding.

Nevertheless described above measures are very useful to be considered as general guidelines for the design of a FT PMSM, they alone might be not enough to achieve the required FT properties of the machine working at the particular application. Thereby, more specific approaches can be utilized during the design process of a FT PMSM, which can be found in the literature.

For example, the paper [5] is dedicated to the comparison of different shapes of segmented stator cores, utilizing tooth-coil windings. Such solutions allow to advance FT properties of an electrical machine even further, though slightly compromising with some of its characteristics. Another recent work [6] discusses utilization of PMSMs in the more-electric aircraft. Among other topics, it provides with the design adjustments to reduce the short-circuit currents in surface permanent magnet and interior permanent magnet machines.

Accordingly, current article will introduce specific changes, made to the regular-design (RD) PMSM, in order to make it FT during the operation with open phases. In particular, Sections II and III will share the initial design of the 5-phase RD PMSM with SLTC winding. Then, Sections IV and V will show the limitations of the initial design and introduce its adaptation for the achievement of the required FT properties. Finally, Section VI will compare the RD and the FT PMSMs.

II. Initial Design of the Machine

In comparison to [2], this paper is focused on the more precise design and adaptation of the selected machine from the mentioned article for the FT operation. In this regard, several changes are introduced into the new calculations, including the following:

- 0.2 mm silicon iron (FeSi) steel is used instead of 0.1 mm FeSi steel. The choice of the steel is based on its availability on the local market. The new steel has very similar magnetization curve in the knee area, which allows to keep dimensions of the machine before its temperature calculation.
- Copper losses are calculated for 150 °C.
- Iron losses include manufacturing factor $k_{Fe,MF} \approx 1.5$, according to the common practice during an electrical machine design.
- The rotor is 1 mm elongated in comparison to the stator core length $l_r = l_s + 1$ mm, in order to take into account fringes of magnetic flux at the ends of the rotor, and to ease its centering during assembly with the stator.
- Lamination space factor $k_{Fe} = 0.95$ is used, according to the information from the manufacturer of the machine.

All the regular requirements and additional limits on the designed PMSM are listed in Table I. Analytical design process of the machine follows [4]. After choice of the machine's topology and its calculation under the rated mode of operation, the next step is to check and adapt its design for the faulty mode of operation with two open phases, during which the machine must keep the torque constant. Additionally, the machine is required to withstand short-time overload, providing the torque of 212 Nm at 2653 rpm, resulting in the double output power of 59 kW in comparison to the rated 30 kW. Thereby, the most challenging for the machine overload mode with two open phases (Overload CE) and the rated operation mode with healthy phases (Rated HE) are calculated and compared. In order to automate calculation of the average torque and losses of the machine for different operation modes, calculation time of the finite element method (FEM) model is chosen to be depended on the supply frequency of the motor, in particular it equals 2 time periods of the supply frequency. Thus, e.g. for the rated operation mode with supply frequency $f = 600$ Hz the calculation time equals $T = 3.333 \cdot 10^{-3}$ s. The time step of the FEM model is chosen to be 1/100 of two time periods of the supply frequency, for the rated mode it is $t = 33.33 \cdot 10^{-6}$ s.

The limit for the rated line-to-line voltage 400 V is a new boundary in comparison to [2], to which, strictly speaking, existing machine doesn't fit. For a star-connected 5-phase winding the line-to-line voltage between not neighboring phases exceeds the line-to-line voltage between neighboring phases. However, since it is only necessary to slightly alter the number of turns in the existing machine, its initial design is left unchanged. Thus, the first harmonics of the rated peak stator phase voltage $\hat{U}_s$ and the rated peak stator phase current $\hat{I}_s$ for m-phase winding are:

$$\hat{U}_{s1} = \frac{\sqrt{2}U_n}{2\sin(\pi/m)} \approx \sqrt{2} \cdot 340 \approx 481 \text{ V}, \quad (1)$$

$$\hat{I}_{s1} = \frac{\sqrt{2}P}{(m\eta U_{s1}\cos\phi)} \approx \sqrt{2} \cdot 23.2 \approx 32.8 \text{ A}, \quad (2)$$

where U_n is the line-to-line voltage between neighboring phases, η is the machine's efficiency. Field distribution of the RD PMSM working at Rated HE mode of operation is provided in Fig. 1.

TABLE I
MOTOR TO DESIGN

Parameter	Value
Number of phases, m	5
Limit for the rated line-to-line voltage U_n, V	400
Rated power P_n, kW	30
Rated speed n, rpm	2000
Rated torque T, Nm	144
Efficiency η	0.95
$\cos\phi$	0.8
Limit for the external max. diameter $D_{re,max}$, mm	290
Limit for long-term peak line-to-line voltage $\hat{U}_{n,max}$, V	600
Limit for long-term peak phase current $\hat{I}_{s,max}$, A	400

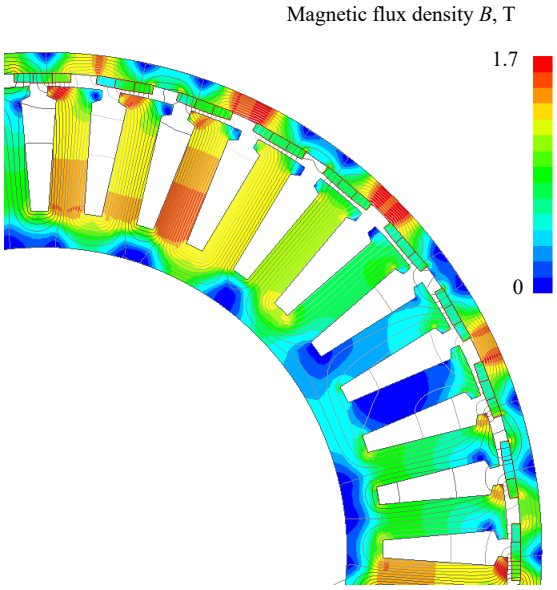


Fig. 1. Field distribution in the RD PMSM working at Rated HE mode.

Copper wire in the stator winding is made of several parallel conductors, its skin effect factor is estimated as $k_R = 1$. Segmented samarium-cobalt magnets are used due to their high Curie point. Coercivity of the real permanent magnet (PM) material for 20 °C varies within $840 - 890$ kA/m limits. For FEM calculations its approximate average value for 150 °C is calculated and used. It is assumed that the PMs are not demagnetized up to 200 °C if the magnetic flux density in the air-gap B_δ under the PMs is co-directed with the PMs' magnetization vector. Example of the more detailed demagnetization analysis, including determination of the most dangerous areas in PMs, is provided in [7]. Analogous work can be done at the next stage of the machine development, when the distribution of B_δ during the faults is already harmonized. Then, Table II summarizes parameters of the stator of the RD PMSM.

TABLE II
PARAMETERS OF THE STATOR OF THE RD PMSM

Parameter	Value
Air-gap length δ, mm	1
Stator core length l_s, mm	62.4
Type of the stator winding	SLTC
Number of phases m	5
Number of stator slots Q_s / number of poles $2p$	40 / 36
Number of slots per phase per pole $q = Q_s/2pm$	2/9
Number of coils per phase $Q_s/2m$	4
Number of parallel paths a	1
Number of turns in series per phase N_{ph}	144
Number of turns per coil $N_{t/c} = a \cdot N_{ph}/(Q_s/2m)$	36
Number of conductors in the stator slot $z_{Qs} = N_{t/c}$	36
Stator slot angle $\alpha_u = p \cdot 2\pi/Q_s$, $^\circ$ el.	162
Fundamental winding factor k_{w1}	0.9877
Stator copper conductor cross-sectional area S_{cs}, mm^2	4.05
Stator slot cross-sectional area S_{us}, mm^2	243
Stator phase resistance R_s, mΩ (at 150 °C)	200.670
Stator phase inductance L_s, μH	3593.134
Stator end winding leakage inductance L_{ew}, μH	228.754

The stator slot cross-sectional area in Table II gives sligtly different value from the one, calculated by the formula:

$$S_{us} = \frac{z_{Qs} S_{cs}}{k_{Cu}} = \frac{z_{Qs} I_s}{k_{Cu} a J_s}, \quad (3)$$

where S_{us} is the cross-sectional area of the stator slot, z_{Qs} is the number of conductors in the stator slot, S_{cs} is the cross-sectional area of a single copper conductor, k_{Cu} is the space factor for copper, I_s is the rated current, J_s is the rated current density. The reason for it is that the actual geometry of the computer-aided design (CAD) drawing has rounded this value.

The type of cooling of the machine is forced air ventilation, without ventilation ducts in the stator core and slots. Current density in the stator winding of the RD PMSM is $J_s = 5.7$ A/mm^2 for the rated operation mode. Then, during faulty modes of operation control algorithm for the designed machines is adjusted in a non-symmetrical way, in order the total current of the machine was minimal.

III. Calculation of R_s and L_s for the SLTC winding

According to Fig. 2, the average length of a single turn of the copper conductor in a single coil of SLTC winding is calculated as follows:

$$l_{Cu,turn} \approx 2l_s + 4\delta_{ew} + 2\pi r_{ew} = 2l_s + 4\delta_{ew} + \pi\tau_u, \quad (4)$$

where l_s is the stator length, δ_{ew} is technological gap, r_{ew} is the average radius of the end winding, τ_u is the slot pitch. Additionally, it is necessary to take into account the lengths of the connections between the coils within a single phase and its outputs. For 4 series-connected tooth-coils, evenly distributed along the air-gap circumference, there are 3 inter-coil connections and 2 output connections, each with the equal length, dependent on the diameter of the air-gap circumference D_δ:

$$l_{Cu,conn} \approx \frac{\pi D_\delta}{4}. \quad (5)$$

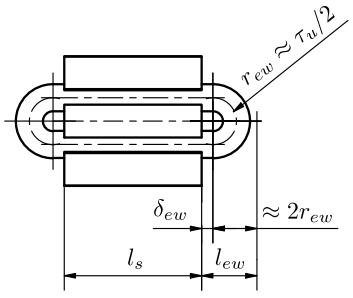


Fig. 2. Dimensions of a signle coil of the SLTC winding.

Then, knowing the lengths of the copper conductors, the stator phase resistance is calculated using common formulas.

Inductance of several coils connected in series within a phase can be found via geometry and properties of the materials utilized in the winding:

$$L_{ph} = N_{c/ph} \cdot \mu_0 \mu_r \frac{N_{t/c}^2 \cdot S_c}{h_c}, \quad (6)$$

where L_{ph} is the phase inductance, $N_{c/ph}$ is the number of coils per phase, μ_0 is the permeability of vacuum, μ_r is the relative permeability, $N_{t/c}$ is the number of turns per coil, S_c is the cross-sectional area of the coil, h_c is the height of the coil. From Eq. (6) the single-phase end winding leakage inductance L_{ew} for a SLTC winding, series connected coils, is calculated acc. to Fig. 2 as follows:

$$L_{ew} = \frac{1}{m}\frac{Q_s}{2} \cdot \mu_0 \mu_{r,ew} \frac{z_{Qs}^2 \cdot (\pi r_{ew}^2 + 2\delta_{ew} \cdot 2r_{ew})}{h_{ew}}, \quad (7)$$

where Q_s is the number of stator slots, $\mu_{r,ew}$ is the effective relative permeability of the end winding, h_{ew} is the height of the end winding. Finally, the total stator phase inductance is calculated as the sum of numerically evaluated inductance from a 2-D FEM software and analytically estimated L_{ew}.

IV. Demagnetization and Overvoltage Issues Occuring in the RD PMSM During its Operation with Open Phases

Fig. 3 demonstrates distribution of normal components of the magnetic flux density in the air-gap and in the middle of the stator teeth during Overload CE mode. Firstly, demagnetization issue is clearly seen in the figure, e.g. at the position around the pole 46.5°-53.5°, where positive half-wave of the air-gap magnetic flux density B_δ is almost completely eliminated and takes negative values instead. Due to the drop of B_δ, the stator current of the machine is increased, in order to keep the torque constant, which leads to oversaturation of the e-steel.

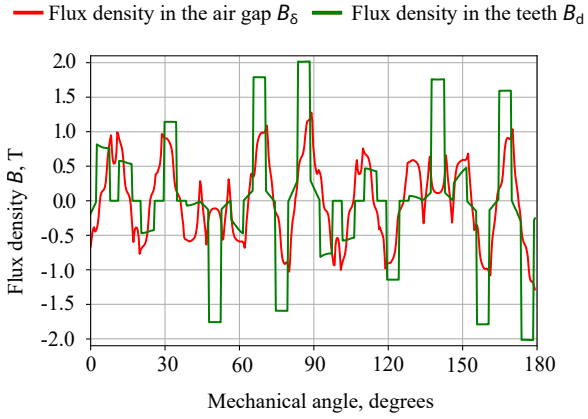


Fig. 3. Distribution of normal components of the magnetic flux density in the RD PMSM working at Overload CE mode.

TABLE III
ELECTROMAGNETIC LOADS OF THE RD PMSM WORKING AT RATED HE AND OVERLOAD CE MODES.

Mode of operation	$\hat{U}_{n,max}$, V	$\hat{I}_{s,max}$, A	$\Delta P_{Cu,}$ kW	$\Delta P_{Fe,}$ kW	$\Delta P_{PM,}$ kW
Rated HE	1061	32.8	0.540	0.597	0.234
Overload CE	14255	290	22.154	1.820	1.674

According to the catalogue data, B-H curve of the chosen FeSi steel ends at the value of 1.66 T for 200-1000 Hz frequency. Thus, for the level of magnetization 2 T the approximation of B-H curve is used, in order to demonstrate insufficience of the RD machine for the faulty mode of operation. In such a case the current $\hat{I}_s$ is increased in a nonlinear manner, taking into account saturation of the magnetic circuit. The real e-steel can require even higher stator current. Exact values of the coefficients for increase of current amplitudes and changed phase shifts for different faulty modes of operation are not discussed in the current paper.

Another issue occuring in the RD machine working under the faulty mode with open phases is overvoltage. Induced phase voltage $u_{2D,FEM}$ calculated in a 2-D FEM software includes voltage drop on all the inductances of the machine except the end winding leakage inductance L_{ew}, and doesn't take into account voltage drop on the stator resistance R_s. Thereby, the total phase voltage is defined in a 2-D FEM software manually as follows:

$$u_s = u_{2D,FEM} + i_s R_s + L_{ew}\frac{di_s}{dt}. \tag{8}$$

Then, Table III summarizes current, voltage and losses of the RD machine, working at Rated HE and Overload CE modes.

V. CHANGE OF THE MACHINE DESIGN TO MEET FT REQUIREMENTS

If only the air-gap δ of the machine is increased, its PM losses are decreased, however the total magnetic voltage remains constant:

$$\hat{U}_{m,total} = H_{cB}h_{PM} = \\ = \frac{\hat{B}_{Fe,n}}{\mu_0\mu_{r,Fe}}l_{Fe} + B_{PM}\frac{H_{cB}}{B_r}h_{PM} + \frac{\hat{B}_{\delta,n}}{\mu_0}\delta = \text{const.}, \tag{9}$$

where H_{cB} is coercivity of PMs, h_{PM} is the height of PMs, $\hat{B}_{Fe,n}$ is the peak value of the normal component of the magnetic flux density in the teeth, $\mu_{r,Fe}$ is the relative permeability of the e-steel, l_{Fe} is the average length of the flux line path in the e-steel, B_{PM} is the magnetic flux density of the PMs, B_r is the remanence of the PMs, $\hat{B}_{\delta,n}$ is the peak value of the normal component of the magnetic flux density in the air-gap. Therefore, the magnetic flux density along the magnetic circuit will start to decrease. This will lead to the change of $\cos\phi$, analogically to the classic synchronous machine when controlling excitation, and drop of the machine's torque, acc. to Maxwell's Stress Tensor:

$$\sigma_{F,tan} = \frac{\hat{B}_{\delta,n}}{\sqrt{2}}\frac{\hat{A}_a}{\sqrt{2}}\cos\phi, \tag{10}$$

where armature linear current density $\hat{A}_a$ can be calculated as a current sum flowing in a slot divided by the slot pitch [4]. For SLTC winding it is:

$$\hat{A}_a = \frac{\hat{I}_s z_{Qs}}{\tau_u} = \frac{\hat{I}_s N_{t/c}}{\tau_u}. \tag{11}$$

In such a case, the stator current of the machine must be increased in order to keep the required torque. However, these measures alone will not help with the demagnetization issue of PMs, since magnetic voltage created by the stator winding will also increase, keeping its peak magnetic field strength $\hat{H}_s$ high, for $\mu_{r,Fe} \approx \infty$:

$$\hat{H}_s = \frac{\hat{I}_s N_{t/c}}{\delta + h_{PM}}. \tag{12}$$

Thereby, it is advisable to increase the air-gap of the machine δ and height of the PMs h_{PM} simultaneously, increasing by this way magnetic voltage created by the PMs, and decreasing the peak magnetic field strength $\hat{H}_s$. The level of demagnetization in the air-gap, illustrated in Fig. 3, will start to decrease, and therefore, with increased resulted peak magnetic flux density in the air-gap $\hat{B}_\delta$, the stator current of the machine must be decreased, in order to keep the torque constant, according to Eq. (10) and Eq. (11).

Then, if there is a limit on the external diameter D_{re} of the machine, its air-gap diameter D_δ must be decreased in order to fit the machine into the required dimension. This will lead to decrease of the stator slot pitch τ_u, and consequently to increase of the linear current density $\hat{A}_a$ and the tangential stress $\sigma_{F,tan}$. However, since the volume around the air-gap V_δ will decrease, the resulted torque of the machine, defined as:

$$T = 2\sigma_{F,tan} \cdot V_\delta, \tag{13}$$

will not alter. If the relative width of PMs α_{PM} is desired to be kept constant, the width of PMs must be adopted accordingly.

Finally, in order to decrease overvoltage in the stator winding, its number of turns N_s must be decreased, and the stator current I_s increased proportionally, in order to keep the value of ampere-turns $I_s N_s$ constant. If the peak values of the stator voltage $\hat{U}_s$ and current $\hat{I}_s$ still don't fit into the required boundaries of Table I, the length and therefore the volume around the air-gap V_δ of the machine must be increased, allowing by this way lower electromagnetic loads due to the lower tangential stress $\sigma_{F,tan}$, with all the parameters kept constant, but less number of turns in comparison to the previous iteration.

VI. Comparison of Parameters of the RD and the FT PMSMs

Fig. 4 demonstrates distribution of normal components of the magnetic flux density in the air-gap and the teeth of the FT PMSM working at Overload CE mode. Drops of the air-gap magnetic flux density below 0 T are still visible, but now they take place between the PMs. These drops can be eliminated completely by further applying measures described in the article. Finally, electromagnetic loads of the FT PMSM are provided in Table IV, changed machine parameters are summarized in Table V.

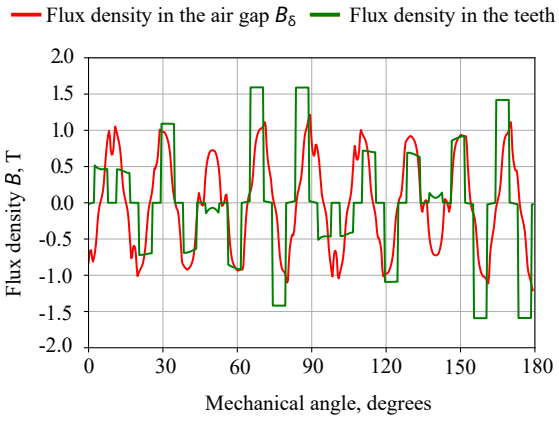


Fig. 4. Distribution of normal components of the magnetic flux density in the FT PMSM working at Overload CE mode.

TABLE IV
Electromagnetic Loads of the FT PMSM Working at Rated HE and Overload CE Modes.

Mode of operation	$\hat{U}_{n,max}$, V	$\hat{I}_{s,max}$, A	ΔP_{Cu}, kW	ΔP_{Fe}, kW	ΔP_{PM}, kW
Rated HE	325	105	0.324	1.054	0.246
Overload CE	530	393	2.390	1.237	0.595

TABLE V
Machine's Parameters Before and After Adoptation of the Design

Parameter	**RD PMSM**	**FT PMSM**
Rated phase voltage U_s, V	340	120
Rated phase current I_s, A	23.2	74.2
Air-gap length δ, mm	1	1.3
Height of PMs h_{PM}, mm	2.8	9
Stator core length l_s, mm	62.4	88
Number of turns in series per phase N_{ph}	144	28
Conductor cross section area S_{cs}, mm^2	4.05	18.20
Stator phase resistance R_s, mΩ (at 150 $^\circ C$)	200.670	11.753
Stator phase inductance L_s, μH	3593.689	149.454
Stator end winding leakage inductance L_{ew}, μH	228.754	8.033
Total mass of the machine, kg	19.3	28.3

VII. Conclusion

The article has showed, that the limits on the outer diameter of a PMSM and its permissible electromagnetic loads have an influence on its adaptation process for the operation with open phases. The changes in the machine design have been introduced under the condition of keeping its torque, speed and supply parameters within the rated values. Higher PMs, wider air-gap, lower number of turns in the stator winding together with the increased core length result in the rise of the machine mass and decrease of the stator resistance and inductance. Listed measures for the particular machine can be generalized and applied to a FT PMSM working under different modes of operation with open phases.

The following steps of electromagnetic design include comparison of different rotors with PMs magnetized in the radial and Halbach array directions, more detailed demagnetization analysis and optimization of the chosen machine.

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-23-8430). Mr. Vitaliy Sizonenko acknowledges Mr. Ondrej Vitek as his research supervisor for his valuable comments and help while writing this paper.

References

[1] 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.

[2] V. Sizonenko and O. Vitek, "Preliminary Design and Comparison of 5 Phase and 6 Phase Fault Tolerant Outer Rotor Permanent Magnet Synchronous Machines with Different Electrical Steel," *2022 IEEE 20th International Power Electronics and Motion Control Conference (PEMC)*, Brasov, Romania, 2022, pp. 390-395, doi: 10.1109/PEMC51159.2022.9962852.

[3] 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.*, San Antonio, TX, USA, 2005, pp. 902-909, doi: 10.1109/IEMDC.2005.195829.

[4] J. Pyrhönen, T. Jokinen, and V. Hrabovcova, *Design of Rotating Electrical Machines*, 2nd ed. New York, NY, USA: Wiley, 2014.

[5] 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)*, Vancouver, BC, Canada, 2021, pp. 4254-4259, doi: 10.1109/ECCE47101.2021.9596020.

[6] N. Bianchi et al., "Permanent Magnet Synchronous Motor Drives for More-Electric Aircraft," *2022 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM)*, Sorrento, Italy, 2022, pp. 871-876, doi: 10.1109/SPEEDAM53979.2022.9842284.

[7] M. Galea, L. Papini, H. Zhang, C. Gerada and T. Hamiti, "Demagnetization Analysis for Halbach Array Configurations in Electrical Machines," in *IEEE Transactions on Magnetics*, vol. 51, no. 9, pp. 1-9, Sept. 2015, Art no. 8107309, doi: 10.1109/TMAG.2015.2429645.

Topology Optimization of Line-Start Synchronous Reluctance Machine with Fully Filled Flux Barriers

Iveta Lolova
Department of Power Electrical and Electronic Engineering
Brno University of Technology, FEEC
Brno, Czech Republic
Iveta.Lolova@vut.cz

***Abstract*—This paper deals with topology optimization based on the normalized Gaussian network of line-start synchronous reluctance machines. The paper aims to discuss the results of performed topology optimization in detail. The individuals are compared from the point of quality (efficiency, power factor and torque ripple). Thus, several feasible geometries of the line-start synchronous reluctance machines with the ability to synchronize are offered and compared. Then the paper offers quantitative processing of geometries to define the probable distribution of iron and aluminum in the rotor to obtain a motor with synchronization capability. The main contribution of this paper is defining the possible geometries of line-start synchronous reluctance machine with fully filled barriers that can reach synchronism. These geometries are based on TO and this paper offers a detailed comparison of them.**

***Index Terms*—AC motors, evaluation process, line-start synchronous reluctance motor, normalized Gaussian function optimization, topology optimization.**

I. Introduction

The usage of electrical rotating machines is responsible for a significant amount of global energy consumption. There are increasing demands for more energy-efficient machines to address this matter, this has led to the IEC 60034-30 efficiency classification standard. Topology optimization (TO) is a promising approach for optimizing electrical rotating machines. The advance of TO is the possibility of obtaining geometry without any earlier expectation compared to parametric optimization (PO), where the geometry shape is strictly defined by the designer of the motor. The comparison of TO and PO is presented in [1].

Various approaches to TO are researched with application to electrical rotating machines. The detailed overview is provided by [2]. This paper deals with TO based on the normalized Gaussian network (NGnet). This method was already published in [3]–[7]. Topology optimization of synchronous reluctance motors (SynRM) and interior permanent magnet synchronous motors (IPMSM) is investigated in [6], [7]. Furthermore, localized material degradation caused by manufacturing is considered in [8].

This paper deals with TO based on the NGnet for line-start synchronous machines (LSSM). There are various types of LSSM with and without permanent magnets. LSSMs are becoming increasingly widespread as a replacement for induction machines due to their better performance, as discussed in [9]–[12]. Also, the development of single-phase line-start synchronous machines were presented in various papers [13]–[17].

The paper is organized into several sections. Section II shortly describes the methodology of used TO. The main Section III of this paper deals with discussing the obtained results of performed TO. This section is divided into two parts: the first part, Section III-A, describes the parameters and shapes of several chosen Geometries (1-8) in detail. While the second part, Section III-B, provides a description of the probabel distribution of iron and aluminum in the rotor to reach synchronism. Finally, Section IV concludes the paper and outlines the main contribution of this paper.

II. Topology optimization

TO based on NGnet offers the possibility of the creation of geometries without any previous expectation. TO divides the optimized space into small elements and assigns the material to them independently. In this paper, the LSSynRM is optimized. There is a challenge in synchronizing ability with this type of machine. Thus special methodology of application TO based on this machine was proposed. This paper develops further the methodology presented in previous authors' papers [3], [4], [18]. The main contribution of this paper is better processing of the data which can now be examined in detail as will be done in the following section. Moreover, there were some changes in the definition of an electromagnetic model in Ansys compared to previous work [4]. Especially, the better distribution of mesh is provided thanks to the unification of elements. So, TO based on NGnet was performed with objectives and constraints stated in Tab. I. The specification of the optimized motor is

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-23-8430). Author acknowledges Assoc. Prof. Jan Barta, Ph.D. as my research supervisor for his valuable comments and help while writing this paper.

This work has been supported by the COMET-K2 Center of the Linz Center of Mechatronics (LCM) funded by the Austrian federal government and the federal state of Upper Austria.

Iveta Lolová is Brno Ph.D. Talent Scholarship Holder – Funded by the Brno City Municipality.

TABLE I
OBJECTIVES AND CONSTRAINTS SETTING.

Stage	Name	Objective	Min Constraint	Max Constraint
Topology creation	Feasible geometry	YES	-	-
Initial analysis	Torque (Nm)	NONE	5.0	20.0
Initial analysis	Torque ripple (%)	NONE	0.0	55.0
Transient ana.	Efficiency (-)	MAX	0.0	1.0
Transient ana.	Power factor (-)	MAX	0.0	1.0
Transient ana.	Torque ripple (%)	MIN	0.0	100.0
Transient ana.	Speed (rpm)	NONE	1 400.0	1 600.0

TABLE II
SPECIFICATION OF THE OPTIMIZED MOTOR.

Rated power	kW	2.2
Number of phases	-	3
Number of pole-pairs	-	2
Rated speed	rpm	1 500
Rated line-to-line voltage	V	400
Terminal connection	-	Y

given in Tab. II. The obtained results of this performed TO are discussed in detail in the following section.

III. RESULTS OF TOPOLOGY OPTIMIZATION

The topology optimization process took about 20 days and was carried out on two Intel(R) Xeon(R) CPU ES-2660 v3 processors, which had 20 physical cores and 40 logical processors running at a maximal speed of 2.60 GHz. During the optimization process, the licence error for ANSYS occurred on the computation server, thus in the obtained result there is a nonnegligible number of errors. The obtained results are presented in Tab. III. Tab. III includes information about the numbers of individuals that were evaluated during TO as well as the stage of the evaluation algorithm of one individual in which the given individual was excluded. It can be seen that 12 969 individuals were evaluated by the evaluation algorithm of one individual fully described in [4]. One piece of iron sheet in geometry is the main condition for feasible geometry. Then almost 11 thousand individuals had feasible geometry. Only 16.15 % of individuals were excluded for unfeasible geometry. That proves that TO based on NGnet does not tend to create unfeasible geometries with many obstacles in geometries.

The next stage of evaluation of one individual is initial analysis, where average torque and torque ripple are considered with the current supply. In this stage there are almost 5 thousand errors, this error means that the geometry is feasible, but during FEM in Ansys Maxwell an error occurred and no results were given back to evaluate. Error in Ansys happened due to computation server problem as described above. It can be seen, that individuals, for which initial analysis was performed, are divided into two-thirds that do not fulfill the conditions given for average torque and torque ripple and the last third of these individuals meet the conditions. The individuals, that fulfill the conditions, approach the final transient analysis.

The transient analysis is the most time-consuming of all stages of evaluation of one individual analysis. During transient analysis, 87 individuals with errors appeared. Almost 2 thousand geometries were evaluated with FEM in Ansys

TABLE III
THE OVERVIEW OF CALCULATED INDIVIDUALS DURING TO.

Description	Note	Meeting the conditions	Number of individuals
All individuals	-	-	12 969
Feasible geometry	-	YES	10 875
Initial analysis	Performed + Error	-	10 875
Initial analysis	Error	-	4 620
Initial analysis	Performed	YES+NO	6 255
Initial analysis	Performed	NO	4 193
Initial analysis	Performed	YES	2 062
Transient analysis	Performed + Error	-	2 062
Transient analysis	Error	-	87
Transient analysis	Performed	YES+NO	1 975
Transient analysis	Performed	NO	579
Transient analysis	Performed	YES	1 396
Synchronized geometry	-	-	1 243

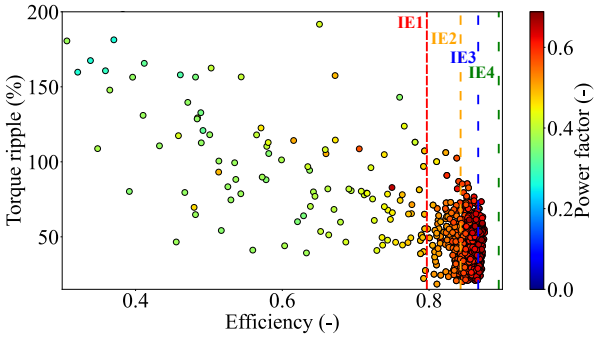


Fig. 1. All obtained designs of performed topology optimization. Red, yellow, blue and green lines represent the minimal efficiency for IE1, IE2, IE3 and IE4, respectively.

Maxwell. The transient analysis ensures that the final design has synchronization ability and can reach a synchronous state. Around 1400 individuals fulfilled conditions for efficiency, power factor, torque ripple and speed stated in Tab. I. The geometries that can not reach synchronous speed also fulfilled the conditions due to the wide range for speed from 1 400 rpm to 1 600 rpm. Such motors work in an asynchronous regime just under the synchronous speed. Thus, 1 243 geometries reached the synchronous state of all individuals. It means, that approximately 10 % of all 12 969 evaluated individuals can reach synchronous speed and fulfill all constraints and conditions.

Fig. 1 shows the results of performed topology optimization with all individuals that approached the transient analysis. Thus, there are also individuals with very high values of torque ripple, these individuals did not reach synchronous speed. Fig. 1 clearly demonstrates that the class efficiency standard IE3 is reached by some designs. No individual can reach the class efficiency standard IE4. In the following subsection, only the individual that can reach synchronous speed will be discussed.

A. Qualitative processing of geometries

This section is dedicated to showing a few chosen geometries that were obtained during performed TO and discussing

TABLE IV
THE OVERVIEW OF CHOSEN INDIVIDUALS.

	Description	IE	Efficiency (%)	Power Factor (-)	Torque Ripple (%)
1	MIN torque ripple	IE1	81.074	0.510	18.500
2	MIN torque ripple and reach IE2	IE2	84.567	0.578	20.151
3	MIN torque ripple and reach IE3	IE3	86.785	0.643	23.093
4	Chosen from Pareto front	IE3	87.023	0.656	36.871
5	Similar to Geometry (4)	IE2	86.560	0.630	43.343
6	MAX Efficiency	IE3	87.565	0.617	52.910
7	Similar to Geometry (8)	IE2	84.835	0.582	55.014
8	MAX Power Factor	IE2	84.858	0.688	58.441

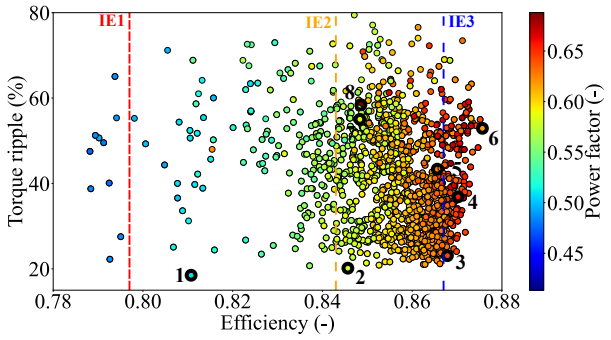


Fig. 2. Obtained designs of topology optimization with the ability to synchronize. Red, yellow and blue lines represent the minimal efficiency for IE1, IE2 and IE3, respectively. There are also highlighted individuals for further discussion with black circles and numbers.

their shapes and performances.

Fig. 2 shows the results of TO with all individuals that can reach synchronous speed. It is apparent from Fig. 2 that power factor and torque ripple rise with higher efficiency. Moreover, there are depicted the minimal values of efficiency for class efficiency standards IE1, IE2 and IE3 with red, yellow and blue lines, respectively. In this figure, the individuals for further discussion are highlighted with black circles and numbers. These individuals were chosen due to their performances. The numbers in Fig. 2 correspond to those in Tab. IV. From the figure, we can observe that values of efficiency can reach up to 87.565 %. The power factor values are from 0.4 to almost 0.7, but these low values are typical for line-start synchronous reluctance machines with fully filled flux barriers. The lowest torque ripple is under 20 %. After reaching IE3, the torque ripple increases dramatically for another small increase in efficiency. The data stated in Tab. IV inform about chosen individuals from Fig. 2.

On the one hand, the first Geometry (1) has the lowest torque ripple of all individuals with a value of 18.500 %. The geometry of Geometry (1) is shown in Fig. 3. It can be seen there is one main flux barrier and this barrier has an extended shape in the middle of the field towards the surface of the rotor.

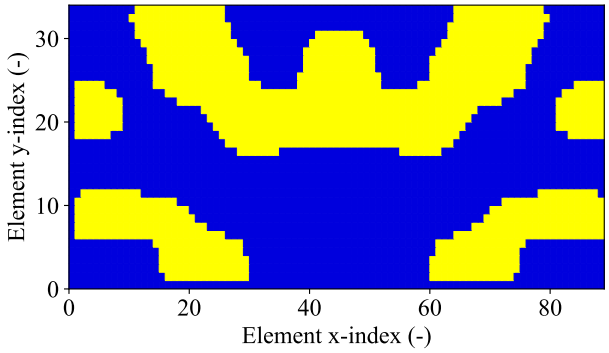


Fig. 3. Geometry (1) with minimal torque ripple. Yellow represents aluminum and blue represents iron sheets.
Efficiency = 81.074 %; Power Factor = 0.510; Torque Ripple = 18.500 %.

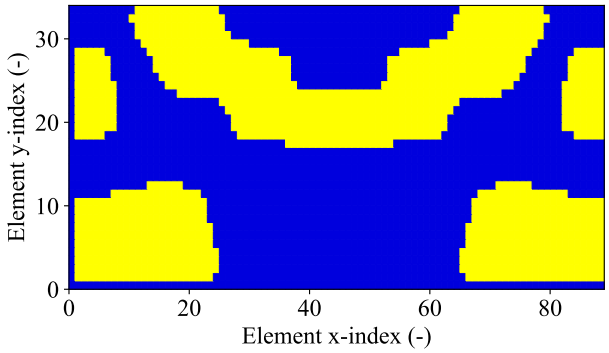


Fig. 4. Geometry (2) with minimal torque ripple and motor reach IE2. Yellow represents aluminum and blue represents iron sheets.
Efficiency = 84.567 %; Power Factor = 0.578; Torque Ripple = 20.151 %.

That creates higher slotting compared to other Geometries (2-8), thus the torque ripple is the lowest. On the other hand, also the efficiency and power factor is the lowest.

As the low torque ripple is desirable for electrical rotating machines two more geometries with the lowest torque ripple are chosen. Moreover, these two geometries: Geometry (2) and Geometry (3) (see Fig. 4 and Fig. 5) are I2 and IE3, respectively. Both Geometries (2,3) are quite similar to Geometry (1), but they lost the extended shape toward the rotor surface, which led to higher torque ripple. Geometry (3) has higher efficiency and power factor compared to Geometry (1,2). Moreover, Geometry (3) has less aluminum material in its geometry, which is also good from a manufacturing cost point of view.

Geometry (4) (see Fig. 6) is chosen from the Pareto front to show that also with such simple geometry, relatively high values of efficiency can be achieved. Geometry (4) has better efficiency, power factor and also torque ripple in comparison with Geometry (5) (see Fig. 7). That is even though Geometry (5) has a more complicated shape.

The next geometry to be shown is Geometry (6) (see Fig. 8), which has the highest efficiency of all individuals. Once again the shape is a quite simple shape with only three flux barriers as Geometry (4,8). Unfortunately, this shape leads to a significantly high torque ripple with a value of 52.910 %.

Finally, Geometry (8) (see Fig. 10) has the highest power

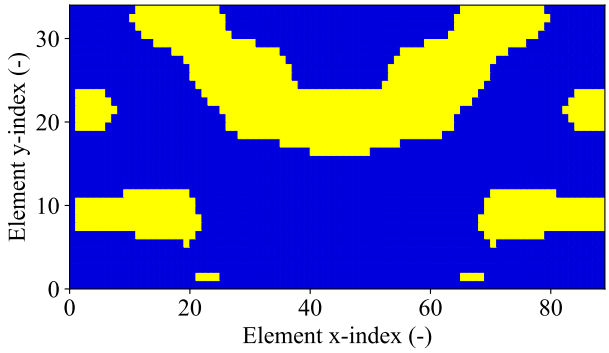


Fig. 5. Geometry (3) with minimal torque ripple and motor reach IE3. Yellow represents aluminum and blue represents iron sheets.
Efficiency = 86.785 %; Power Factor = 0.643; Torque Ripple = 23.093 %.

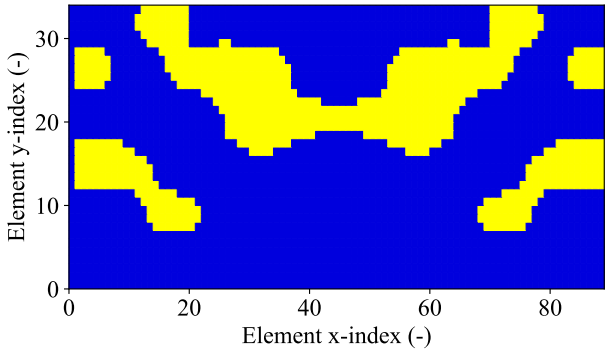


Fig. 7. Geometry (5) chosen with similar values to Geometry (4). Yellow represents aluminum and blue represents iron sheets.
Efficiency = 86.560 %; Power Factor = 0.630; Torque Ripple = 43.343 %.

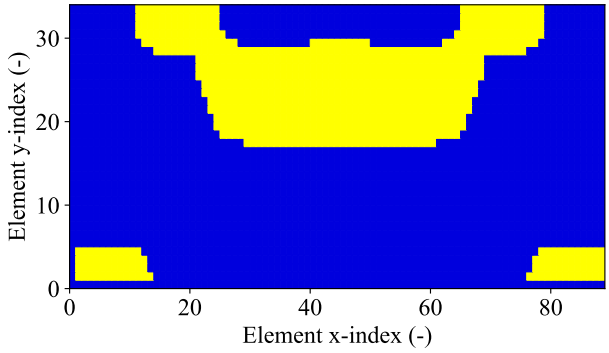


Fig. 6. Geometry (4) chosen on pareto front. Yellow represents aluminum and blue represents iron sheets.
Efficiency = 87.023 %; Power Factor = 0.656; Torque Ripple = 36.871 %.

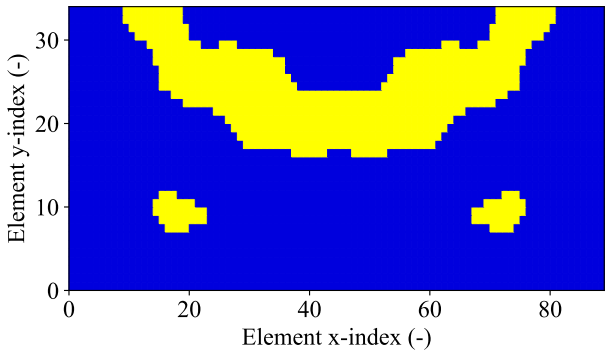


Fig. 8. Geometry (6) with maximal efficiency. Yellow represents aluminum and blue represents iron sheets.
Efficiency = 87.565 %; Power Factor = 0.617; Torque Ripple = 52.910 %.

factor with a value of 0.688 and also the highest torque ripple. Geometry (7) (see Fig. 9) has similar torque ripple and efficiency, but the power factor is significantly lower. Again, it can be seen that with simpler shape the power factor is better.

B. Quantitative processing of geometries

This section deals with the quantitative processing of geometries. Only the individuals with the ability to synchronize were included for processing. The aim of quantitative processing of geometries is to determine where aluminum or iron sheet is most likely to be found. This allows us to quickly compare all geometries to each other without examining all 1320 geometries in detail as was done in the previous section.

Fig. 11 shows the probability of occurrence of the element aluminum in geometry. Based on this figure it can be assumed where the elements aluminum are placed in the rotor to obtain the geometry of the rotor that can synchronize. Fig. 11 reveals, that the geometry has one main flux barrier filed with aluminum with the probability of almost 100 %. There is also a prolonged shape towards the rotor surface in the middle of the pole. It ensures a lower torque ripple referring to the previous chapter. Then there are two flux guides of iron. On both sides of the pole are also flux barriers filled with aluminum to ensure a higher power factor and efficiency. These barriers help to increase the saliency ratio of the rotor. Without them, the saliency ratio would be low and thus there is a lesser chance that the motor can reach synchronous speed.

IV. Conclusion

This paper deals with TO based on NGnet of LSSynRM. Firstly, the used methodology of TO is shortly mentioned. The objectives and constraints setting are described. The specification of the optimized motor is given. The main part of this paper deals with obtained results of performed TO as the aim of this paper is to provide a detailed discussion of these results. Thus, the results were first presented with a description of the process of optimization. From that, it can be seen that TO based on NGnet does not tend to create unfeasible geometries as it is problem with direct topology optimization (DTO). Moreover, around 10 % of all evaluated individuals have the ability to reach a synchronous steady state.

From the qualitative processing of geometries, several observations of geometries are done. Firstly, geometry tends to create one main flux barrier. For lower torque ripple the barriers have an extension towards the rotor surface. The power factor is higher with simple geometry. It can be seen from Tab. IV, the efficiencies are comparable to commonly used induction machines (IMs). Moreover, the power factor of LSSynRMs has a lower power factor than IMs. The low power factor is expected for LSSynRMs due to the high magnetizing current. The higher power factor can be achieved with a higher saliency ratio, but the used methodology tends to create geometries with a low saliency ratio. Thus, further improvements of the TO for LSSynRM should be carried out.

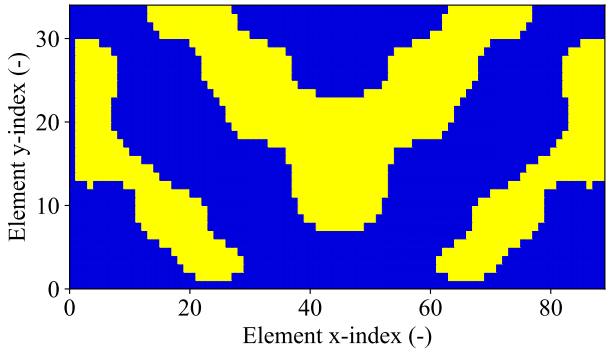


Fig. 9. Geometry (7) with similar values (Efficiency, Power Factor, Torque Ripple) to Geometry (8). Yellow represents aluminum and blue represents iron sheets.
Efficiency = 84.835 %; Power Factor = 0.582; Torque Ripple = 55.014 %.

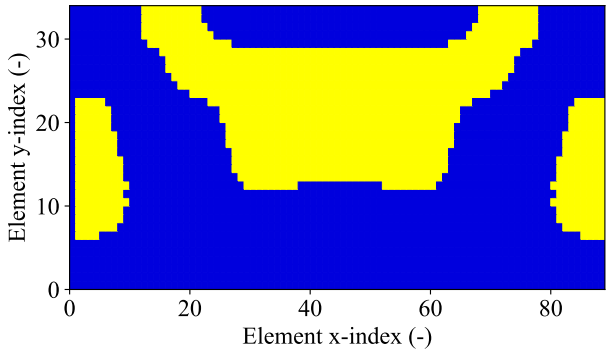


Fig. 10. Geometry (8) with maximal power factor. Yellow represents aluminum and blue represents iron sheets.
Efficiency = 84.858 %; Power Factor = 0.688; Torque Ripple = 58.440 %.

The quantitative processing of geometries reveals the probability distributions of iron and aluminum in the rotor. To create a motor with the ability to reach a synchronous state, there is almost 100 % probability to have a main flux barrier of aluminum and two flux guides of iron. Then there is a variation in the shape of the flux barriers. Moreover, there are flux barriers on the sides of the pole. The shape and presence of these barriers significantly affect efficiency, power factor and torque ripple.

The main contribution of this paper is defining the possible geometries of LSSynRM with fully filled barriers that can reach synchronism. These geometries are based on TO and this paper offers a detailed comparison of them.

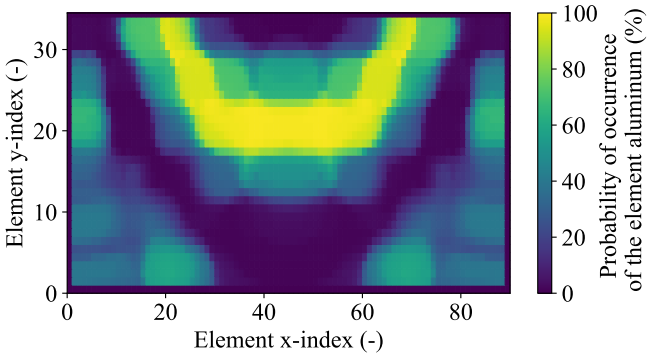


Fig. 11. Probability of occurrence of the element aluminum in geometry to achieve synchronous speed.

References

[1] A.-C.Zăvoianu, G. Bramerdorfer, E. Lughofer, and S. Saminger-Platz, "*Multi-objective topology optimization of electrical machine designs using evolutionary algorithms with discrete and real encodings*," in *Computer Aided Systems Theory – EUROCAST 2017*. Cham: Springer, 2018, pp. 331–338.

[2] F. Nishanth and B. Wang, "Topology optimization of electric machines: A review," in *2022 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2022, pp. 1–8.

[3] I. Lolova, "Topology optimization of the line-start synchronous machines," Brno, 2020. [Online]. Available: https://www.vutbr.cz/studenti/zav-prace/detail/125799

[4] I. Lolova, J. Barta, G. Bramerdorfer, and S. Silber, "Topology optimization of line-start synchronous reluctance machine," in *2020 19th International Conference on Mechatronics - Mechatronika (ME)*, 2020, pp. 1–7.

[5] J. Barta, L. Knebl, G. Bramerdorfer, I. Lolova, S. Silber, and O. Vitek, "Topology optimization of rotor bars geometry and arrangement for a line-start permanent magnet synchronous machine," *IEEE Access*, vol. 9, pp. 115 192–115 204, 2021.

[6] S. Sato, T. Sato, and H. Igarashi, "Topology optimization of synchronous reluctance motor using normalized gaussian network," in *IEEE Transactions on Magnetics*, vol. 51(3). IEEE, 2015.

[7] T. Sato, K. Watanabe, and H. Igarashi, "Multimaterial topology optimization of electric machines based on normalized gaussian network," in *IEEE Transactions on Magnetics*, vol. 51(3). IEEE, 2015.

[8] Y. Hidaka and H. Igarashi, "*Topology Optimization of Synchronous Reluctance Motors Considering Localized Magnetic Degradation Caused by Punching*," in *IEEE Transactions on Magnetics*, vol. 51, no. (6). IEEE, 2017, pp. 1–4.

[9] M. Gamba, G. Pellegrino, A. Vagati, and F. Villata, "Design of a line-start synchronous reluctance motor," in *2013 International Electric Machines & Drives Conference*. IEEE, 2013, pp. 648–6557.

[10] V. Abramenko, I. Petrov, J. Pyrhonen, and L. Aarniovuori, "Design aspects of direct-on-line synchronous reluctance motors," in *2018 XIII International Conference on Electrical Machines (ICEM)*. ICEM, 2018, pp. 138–145.

[11] M. Gamba, E. Armando, G. Pellegrino, A. Vagati, B. Janjic, and J. Schaab, "Line-start synchronous reluctance motors: Design guidelines and testing via active inertia emulation," in *2015 IEEE Energy Conversion Congress and Exposition (ECCE)*. IEEE, 2015, p. 4820–4827.

[12] Y. Hu, B. Chen, Y. Xiao, J. Shi, L. Li, and X. Li, "Rotor design and optimization of the three-phase line-start synchronous reluctance motor," in *2019 22nd International Conference on Electrical Machines and Systems (ICEMS)*, 2019, pp. 1–6.

[13] I. Lolova, J. Barta, G. Bramerdorfer, V. Bilek, and O. Vitek, "The optimization of single-phase line-start permanent magnet synchronous motor for household applications," in *2022 International Conference on Electrical Machines (ICEM)*, 2022, pp. 774–779.

[14] K. Kurihara, H. Fukuda, and T. Kubota, "Trade-off with starting capability to maximize efficiency of a single-phase line-start permanent magnet motor," in *2017 20th International Conference on Electrical Machines and Systems (ICEMS)*, 2017, pp. 1–4.

[15] W.-S. Lee, M.-C. Kang, G.-T. Kim, H.-K. Shin, and C.-J. Kim, "Optimal design for characteristics improvement of single phase line-start permanent magnet synchronous motor," in *2016 19th International Conference on Electrical Machines and Systems (ICEMS)*, 2016, pp. 1–5.

[16] S.-w. Baek and B.-i. Kwon, "Optimum design of a single-phase line-start pm motor considering efficiency, maximum torque, and starting torque," *IEEE Transactions on Magnetics*, vol. 48, no. 12, pp. 4850–4859, 2012.

[17] J. Shang, W. Tang, and C. Liu, "Optimization of electromagnetic performance of single phase line-start permanent magnet synchronous motor based on the finite element method," in *2019 22nd International Conference on Electrical Machines and Systems (ICEMS)*, 2019, pp. 1–6.

[18] I. Lolová, "Electromagnetic analysis of line-start synchronous reluctance machines optimized by means of topology optimization," in *Proceedings II of the 27th Conference STUDENT EEICT 2021*. BUT FEEC, 2021, pp. 245–249.

Performance Parameter Identification and Stability Analysis of Homopolar Electrodynamic Magnetic Bearings for Suspension System

David Rura
Department of Power Electrical and Electronic engineering
Brno University of Technology
Brno, Czech Republic
david.rura@vut.cz

***Abstract*—This research paper focuses on the performance parameter identification and further analysis of homopolar electrodynamic magnetic bearings, which utilize levitation to achieve suspension. This levitation is made possible through the interaction between the magnetic field generated by a set of permanent magnets and the one excited by induced eddy currents in the rotor caused by its relative motion. As the electrodynamic bearings problem is not fully understood, and mathematical descriptions are lacking, models are often utilized to describe the performance of the magnetic bearings. To determine the model parameters, finite element analysis is adopted. Additionally, the stability of the magnetically levitated system is described by the Jeffcott rotor. The objective of this paper is to present the operational range within which the electrodynamic bearings can perform stably.**

***Index Terms*—electrodynamic magnetic bearings, radial magnetization, levitation, suspension, magnetic field, permanent magnet, induced eddy currents, rotor, modeling, stability analysis, finite element analysis, Jeffcott rotor, performance, stability**

INTRODUCTION

Magnetic bearings have gained interest recently due to their advantages over traditional mechanical bearings, including reduced friction, oil-free operation, lower maintenance costs, and more. The key to levitation is a magnetic field. Based on the principle of operation they can be divided into two main types: reluctance and Lorentz force. This paper focuses on the bearings based on Lorentz force generated by induced currents, known as electrodynamical magnetic bearings (EDB). These bearings operate based on the interaction between magnetic fields and electric currents, creating forces that support the rotating shaft. However, the dynamic behavior of the shaft in an electrodynamic bearing system is a complex task due to the limited operational performance of electrodynamic bearings. The complexity comes from the fact that usually, these bearings have insufficient stiffness, damping, or both as will be shown in this paper as well.

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-23-8430). Author acknowledges Assoc. Prof. Jan Barta, Ph.D. as my research supervisor for his valuable comments and help while writing this paper.

Therefore, the stability analysis of electrodynamic bearings is crucial for ensuring their reliable and safe operation. The stability analysis involves examining the behavior of the shaft and the electromagnetic forces acting on it under different operating conditions to determine whether the system operates stably. The analysis can be done through various analytical, numerical, and experimental approaches.

In addition to the analysis techniques, several factors can affect the stability of electrodynamic bearings, including the magnetic field strength, the geometric design of the bearing, and the type and magnitude of external forces acting on the system. The presence of nonlinearities, such as saturation, eddy currents, or imperfection of magnetic materials, can also influence the system's stability and dynamic behavior.

To perform a detailed study of the stability of electrodynamic magnetic bearings, an analytical model has been developed. While a few models for these bearings already exist, most of them substitute the rotor with a circuit containing RL elements, representing resistance for the current path and inductance due to the natural circular path of induced currents. By performing finite element analysis, the RL parameters can be identified, and the stability of motion can be examined. However, these models are built under certain assumptions, limiting their use.

This paper starts with a literature review on the modeling of EDBs in which are discussed the available geometries and the research progress toward reliable EDBs. The second chapter introduces the description of the electromechanical model and further model parameter identification. The partial stability analysis is presented in the last section and the entire work is concluded at the very end.

I. LITERATURE REVIEW

The concept of electrodynamical bearings is not new, but researchers have long faced the challenge of high eddy-current losses. Eddy currents are induced by changes in the magnetic field, and initial attempts to incorporate them into the system resulted in high losses despite sufficient stiffness and damping. As a result, EDBs were not efficient for industrial use [1] and

others that also dealt with high eddy-current losses caused by the use of alternating fields in the airgap.

Recent research has demonstrated the potential of using radial magnetization to address the challenges of eddy-current losses in electrodynamical bearings [1]. However, due to the high cost of custom-made and low-batch radial magnets, an alternative approach has been proposed using a pair of axially magnetized magnets placed in opposition to create quasi-radial magnetic flux. The main challenge with this approach is designing an effective flux pathway, as the radial flux is present on both the inner and outer sides of the stator, while the rotor is typically located inside. The research proposes several methods to address this issue. Examples of these null-flux electrodynamic bearings can be found in [1], [2] and [3].

Further progress in radial field excitation has been made by introducing radially magnetized magnets [4]. These magnets can reinforce the overall damping and rotating stiffness of EDBs and can be arranged in the commonly known Halbach configuration, which concentrates the magnetic field above one of the surfaces. This leads to a high-magnetic field with low leakage in the non-used direction, as demonstrated in [1].

Analytical models derived from Maxwell's equations, as presented in previous literature such as [1] or [3]. It is noted that the derivation process of these models is subjected to certain limitations and assumptions. The derivation of these models are highly restrictive in a sense of hypotheses that are taken. These models do not offer a holistic comprehension of all the physical parameters involved, nor do they conduct an in-depth analysis of their effects on the system under investigation.

A more convenient approach in the field of magnetic bearings involves using an analogy between the rotor's RL (resistance-inductance) nature. This model is further improved by incorporating a spring-damper model. The electrical resistance is analogous to a mechanical spring, while the inductance is akin to a damper, and vice versa. The cited literature, [5] and [6], employs the Jeffcott rotor model to study dynamic stability through root loci analysis. The research papers demonstrate that the EDB (electromagnetic bearing) performs unstably when used alone. Thus, external non-rotating damping must be introduced. One possible way of introducing damping into the system is outlined in [7], followed by [7] and [8].

In this study, the EDB is modeled using the RL (resistance-inductance) model presented in [9] and [10]. Previous research by the same authors, namely [11] and [9], addressed the identification of parameters for the electromechanical model. Further details on the electromechanical model will be presented later in this paper. The availability of an existing model enables the examination of the stability of the entire system. Unfortunately, due to the lack of experimental data in the field of electrodynamic bearings, it is not an easy task to find models that can be adapted to experiments and matched with their results. Most of the data obtained are from finite element (FE) analyses. One example of an extensive experiment on a homopolar EDB is provided in [1], and a few additional studies have been based on this data, with enhanced models applied. Conducting FE analyses of the EDB is also challenging, especially due to meshing. False eddy-currents may develop in the rotor if the mesh is too coarse or if the mesh type is improperly selected, as reported for example in [12].

The upcoming section will address the construction of a homopolar EDB based on an electromechanical model, as well as a description of the process for obtaining identified parameters for the model.

II. Principle of Electrodynamic Suspension

The operation principle of EDBs is based on the induction of eddy currents in the moving body, which is the rotor. Various approaches exist for describing the principle of force development in EDBs. From the author's perspective, the most straightforward approach is through the Lorentz law, which is adapted from the force F acting on charged particles to the force acting on a wire with length $d\mathbf{l}$, flowing current i, and located in a magnetic field B.

$$\mathbf{F} = i(d\mathbf{l} \times \mathbf{B}) \tag{1}$$

With the help of the Lenz law, which describes the direction of the induced current that opposes the change in magnetic flux through the area enclosed by the loop, there are two opposing magnetic fields, resulting in the interaction described by Eq. (1). Due to the inductive nature of the current loop, there is a phase shift between the induced magnetomotive force and the induced current. As a result, the direction of the force will also be phase-shifted from rotor deviation. The force vector can be decomposed into two components: the lift force F_{lift}, which effectively repels the rotor towards the stable center, and the drag force F_{drag}, which is perpendicular to the lift force and has a destabilizing effect.

III. Electromechanical model

The previous section provided an introduction to the principle of force development in the EDB. In this section, the electromechanical model and the mechanical model using the Jeffcott rotor model will be discussed.

A. Electromechanical model of electrodynamic bearings

The rotor can be modeled as a single loop coil with both inductance and resistance. However, this model faces a major problem regarding the direction of the developed force in the EDB. As the rotational speed increases, the reactance of the coil also increases, while the resistance remains relatively stable. The skin effect causes the current to be squeezed towards the surface, resulting in a thinner effective conductive path. This phenomenon is responsible for the discrepancy between simulated and experimental results. To address this issue, we introduced a speed-resistant dependency into the model by adopting the decomposition of tanh functions in [10]. Although we tested various methods, this method yielded the lowest error.

The equations describing the forces developed by the EDB using the RL model of the motor can be expressed as follows

$$\bar{F} = -\frac{c_m^2 \omega}{\left(R_r^2 + (\omega L_r)^2\right)} \left(\omega L_r - j R_r\right) \bar{z} \tag{2}$$

This complex equation can be split into real and imaginary parts describing the lift force acting in the direction towards the magnetic center and the drag force perpendicular to the lift force. The equations for the lift and drag forces developed by the EDB can be expressed as follows:

$$F_{lift} = \frac{\omega^2 \left(L_r/c_m^2\right)}{\left(R_r/c_m^2\right)^2 + \left(\omega \left(L_r/c_m^2\right)\right)^2} |\bar{z}| \tag{3}$$

$$F_{drag} = \frac{\omega \left(R_r/c_m^2\right)}{\left(R_r/c_m^2\right)^2 + \left(\omega \left(L_r/c_m^2\right)\right)^2} |\bar{z}|. \tag{4}$$

The equations shown above represent the resistance and inductance of the conducting rotor with R_r and L_r, respectively. We rewrite these equations to obtain the parameters R_r/c_m^2 and L_r/c_m^2. Here, c_m represents a scaling factor that can be set to 1 to fix one degree of freedom. This is possible because the global impedance depends on the chosen point of view for the flux variation exciting the equivalent circuit [10]. However, it is later fine-tuned for specific dimensions and operating points.

The Jetcoff rotor's EDB behavior is modeled using a published model from [9]. As the counteracting force generating rotation is rotated by δ, the non-rotating damping and stiffness are mapped in the inertial frame, while the rotating damping in the rotating frame is rotated from the inertial frame by $\omega t + \delta$. By framing the parameters into the appropriate frame, the equation of motion can be obtained.

The motion is described in the reference frame named $\hat{\mathbf{I}}$, where the position is represented by a complex vector $\bar{z}_I = x_I + j y_I$. The general equation of motion in the frame $\hat{\mathbf{I}}$ can be written as:

$$m\ddot{\bar{z}}_I + c\dot{\bar{z}}_I + k\bar{z}_I = 0. \tag{5}$$

Here, m is the mass, c is the damping, and k is the stiffness. If there is no external force acting on the rotor, the equation equals 0.

To account for the angular shift between the direction of the levitation force and the actual offset vector, the equation 5 must be rewritten by adding a phase angle δ to the rotating damping $\bar{c}_r = c_r \cos\theta + j c_r \sin\theta$. The final equation of motion for the Jeffcott rotor supported by EDB can be expressed as:

$$\begin{aligned} m\ddot{\bar{z}}I + (c\mathrm{nr} + c_r \cos\theta + j c_r \sin\theta)\, \dot{\bar{z}}_I \\ + (k + c_r \omega \sin\theta - j c_r \omega \cos\theta)\, \bar{z}_I = 0 \end{aligned} \tag{6}$$

This equation is formed under certain conditions that are discussed in [10]. Here, x_I and y_I are the Cartesian coordinates in the inertial frame, m stands for the mass of the rotor, c_n and c_{nr} are the rotating and non-rotating dampings, k is the stiffness, and δ is the angle between deviation vector and the force vector.

The modeling of the Jeffcott rotor is based on the work published in [13]. Due to the phase shift between the direction of the offset and force developed in the bearing, the modified model must be considered as presented in Eq. 6 from [9] and further discussed in [10]. Both papers also describe the identification of parameters for the EDB electromechanical model.

B. Electromechanical model

The electromechanical model combines the electrical and motion models of the EDB to describe the overall motion of the system. A stability analysis can be performed to evaluate the EDB's performance across a wide speed range. The derivation of the linked model is based on an analysis of the energy balance in the system. Mechanical vibrations are typically damped by energy injected into the system. The advantage of the EDB type considered in this work is that if the rotor is spinning around the magnetic center, the bearings do not inject any power into the system. This also implies that there is no problem with stability due to insufficient stiffness and damping very close to the center position.

Equations that join both models are derived from the power balance in the system:

$$c_r \left(\omega - \lambda_{\mathrm{Im}}\right)^2 |\bar{z}I|^2 \cos\theta = \frac{c_m^2 \, |\bar{z}I|^2 \left(\omega - \lambda \mathrm{Im}\right)^2}{\sqrt{R_r^2 + \left(\left(\omega - \lambda \mathrm{Im}\right) L_r\right)^2}} \cos\varphi \tag{7}$$

$$c_r \left(\omega - \lambda_{\mathrm{Im}}\right)^2 |\bar{z}I|^2 \sin\theta = \frac{c_m^2 \, |\bar{z}I|^2 \left(\omega - \lambda \mathrm{Im}\right)^2}{\sqrt{R_r^2 + \left(\left(\omega - \lambda \mathrm{Im}\right) L_r\right)^2}} \sin\varphi \tag{8}$$

From these linking equations, the following conclusions can be drawn:

$$\theta = \varphi = \arctan\left(\frac{\left(\omega - \lambda_{\mathrm{Im}}\right) L_r}{R_r}\right) \tag{9}$$

$$c_r = \frac{c_m^2}{\sqrt{R_r^2 + \left(\left(\omega - \lambda_{\mathrm{Im}}\right) L_r\right)^2}} \tag{10}$$

As shown, the equations return the mechanical parameter c_r (rotating damping) from identified electrical model parameters. In the next section, a method for identifying all the parameters needed for the electromechanical model will be discussed.

C. Parameter Identification

To identify the parameters required for the equation of motion for the EDB, the stiffness, rotating and non-rotating damping, and global resistance and inductance need to be calculated. All of these parameters can be determined from FEM analysis of the EDB.

1) Stiffenss: The stiffness k can be identified by placing the rotor off-center and detecting the lift force. The stiffness can be calculated using the equation:

$$k = \frac{F_{lift}}{\Delta x} \tag{11}$$

For purely electrodynamic suspension, the stiffness is zero, as there is no physical contact between the rotor and stator.

2) Non-Rotating Damping c_{nr}: There are two approaches to identify non-rotating damping in the EDB. The first approach introduces the whirl constant velocity to the off-centered rotor spinning at the same frequency. The values of the forces are noted for different velocities, and the following equation is obtained:

$$\bar{F} = -jc_{nr}\lambda_I \bar{z} \tag{12}$$

The second approach involves fixing the rotor at an offset position and spinning the stator instead. The reference frames are changed, but the equation remains the same. The first method is easier for experimentation, and the second is more convenient for FE simulation. Non-rotating stiffness is also zero for purely electrodynamic suspension.

3) Rotating Damping c_r: Rotating damping can be obtained by spinning the rotor at a fixed offset position. In this case, $\dot{\bar{z}} = 0$ and $\lambda_I = 0$. The equation of motion for the EDB is modified to the following form:

$$\bar{F} = -(c_r\omega \sin\delta - jc_r\omega\cos\delta)\bar{z} \tag{13}$$

The values of c_r and δ are obtained by applying the least-mean-square criterion to Eq. (13). It should be noted that both parameters are usually applicable to a wide range, but after a threshold, both become frequency-dependent, as shown in [11].

IV. Case Study

In this study, we further investigated the model developed in [4] by utilizing the Halbach magnet arrangement, which improved overall performance parameters. The Halbach magnet arrangement rectified the magnetic field by concentrating it on one side of the magnet arrangement structure, obviating the problems of the Lembke design [1] that uses two opposing axially magnetized magnets to generate quasi-radial magnetization, and pole shoes to rectify the magnetic flux and reduce leakage flux and remanence of the magnetic circuit, achieving higher radial flux.

In this work, calculated stiffness and damping using the model presented in [1] and applied an enhanced electromechanical model to study the performance of the EDB, representing a step forward towards stability analysis. The heteropolar EDB is known for its low damping, and systems employing this type of EDB usually require an external damper to ensure sufficient damping under load. An interesting study in [14] investigated the use of AMB with feedback control to introduce damping into the system. Another method of introducing external damping is to place a damper between the stator of the EDB and the frame, which is sometimes used in air bearings to enhance damping.

TABLE I: Identified model parameters.

R_r (mΩ)	65.5
L_r (mH)	317.1
k (N/m)	0
c_{nr} (Ns/m)	0
c_r (Ns/m)	4.21
δ (°)	68
m (kg)	0.5

V. Model Parameters

A. RL model parameters

Identification of RL parameters comes from the examination of results obtained from FE analysis. With the help of Eq. (3) and Eq. (4), a minimization algorithm is applied to find the global error minimum. In this case, a least-mean-square method has been applied. The search for RL parameters must be done for both forces at the same time due to the coupling of both equations.

Let us consider the data from previous work [4]. The data are shown in Fig. 1.

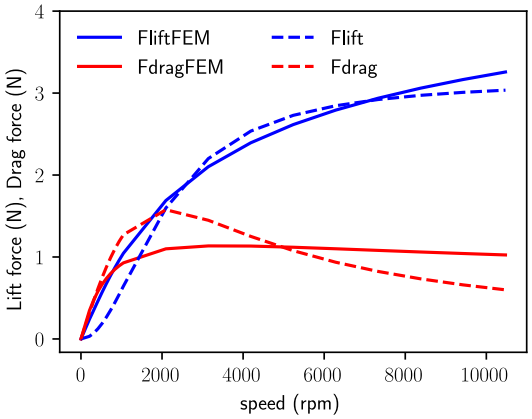


Fig. 1: Lift and drag forces from FE analysis and using Eq. (3) and Eq. (4) with obtained RL parameters.

RL parameters obtained from the force analysis give the following values listed in Table I.

It can be noted that there is a mismatch between the forces obtained from FE analysis, which is also in agreement with the work [11]. Further steps for tuning the model parameters have been presented in [10], but they were not applied in this work.

B. Stability

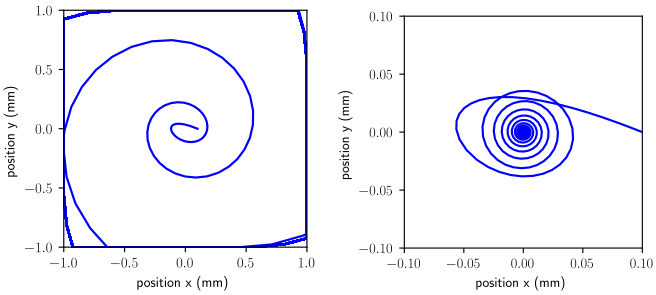
The stability analysis is conducted by considering the Jeffcott rotor and the parameters listed in the Table I. These parameters are obtained from FE analysis but can also be obtained experimentally.

In general, there are stable and unstable rotations. The critical state between these two states is when the rotor is

rotating in the orbit. In the stable state, the rotor returns to its center when an external force or perturbation is fully damped. The unstable operation describes the operation in which the system doesn't provide enough stiffness or damping.

Fig. 2a shows the rotor orbit while operating at 10000 rpm. The initial offset starts at $\Delta x = 0.1mm$, which is 10% of the airgap. The orbit gradually rises, and in a real-world scenario, the rotor would crash with the stator. Therefore, drop-down or safety bearings are usually adopted in the system. The bearing itself is unstable in the studied range (1000-100000rpm). Higher speeds are limited by material strength. The higher the speed, the lower the radius allowed. To provide stability to the system, external damping and stiffness must be introduced, which was reported in previous research.

Fig. 2b shows the studied EDB with an introduced external stiffness of $k = 55$ N/m and non-rotating damping of $c_{nr} = 55$ Ns/m. Under these conditions, the system can provide stable levitation at 10000 rpm.



(a) Unstable (b) Stable

Fig. 2: Rotor orbit during unstable and stable operation with external $k = 55$ N/m and $c_{nr} = 55$ Ns/m at 10000 rpm.

To examine the stability of the system, a stability analysis needs to be performed. Due to the low insternal damping of the EDB, an extreme amount of external damping has to be introduced to ensure stability. The additional non-rotating damping with respect to speed is shown in Fig. 3.

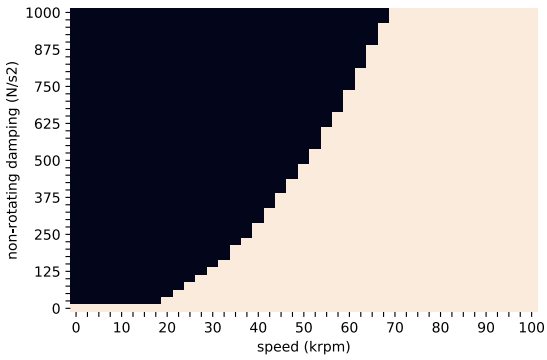


Fig. 3: Additional damping introduced to the system (dark are represents unstable operation) while stiffness is zero

CONCLUSION

Significant strides in understanding the performance of the EDB system have been provided by this study. The model parameters have been identified through FE analysis and the overall stability under obtained parameters has been studied by utilizing the presented electromechanical model. It has been shown that for a purely electrodynamical suspension, the stiffness and non-rotating damping are zero, and thus external stiffness and non-rotating damping must be introduced to provide stable levitation. The values for both parameters under which the system is stable have also been proposed by our research. Stable operation of the suspension system can be delivered, for instance, by using $k = 55$ N/m and $c_{nr} = 55$ Ns/m at 10000 rpm, as has been shown.

REFERENCES

[1] T. A. Lembke, *Design and analysis of a novel low loss homopolar electrodynamic bearing*. PhD thesis, KTH, Stockholm, 2005.

[2] K. Davey, A. Filatov, and R. Thompson, "Design and analysis of passive homopolar null flux bearings," *IEEE Transactions on Magnetics*, vol. 41, pp. 1169–1175, 3 2005.

[3] A. V. Filatov, *"Null-E" magnetic bearings*. PhD thesis, University of Virginia, Virginia, 2002.

[4] D. Rura, J. Barta, and P. Klima, "Design and Analysis of Homopolar Electrodynamic Bearing with Radial Magnets," in *2022 International Conference on Electrical Machines (ICEM)*, pp. 321–327, IEEE, 9 2022.

[5] N. Amati, X. De Lépine, and A. Tonoli, "Modeling of Electrodynamic Bearings," *Journal of Vibration and Acoustics*, vol. 130, 12 2008.

[6] J. G. Detoni, F. Impinna, A. Tonoli, and N. Amati, "Unified modelling of passive homopolar and heteropolar electrodynamic bearings," *Journal of Sound and Vibration*, vol. 331, no. 19, pp. 4219–4232, 2012.

[7] A. Tonoli, N. Amati, G. Genta, F. Impinna, and J. Girardello Detoni, "A new stabilization technique for electrodynamic bearings," pp. 551–560, 3 2010.

[8] Q. Cui, "Stabilization of electrodynamic bearings with active magnetic dampers," 2016.

[9] V. Kluyskens, B. Dehez, and H. B. Ahmed, *Dynamical electromechanical model for magnetic bearings*. PhD thesis, 2007.

[10] V. Kluyskens and B. Dehez, "Dynamical Electromechanical Model for Magnetic Bearings Subject to Eddy Currents," *IEEE Transactions on Magnetics*, vol. 49, pp. 1444–1452, 4 2013.

[11] V. Kluyskens and B. Dehez, "Analysis of the stability and the consumption of an electrodynamic bearing for different operation conditions," *13th International Symposium on Magnetic Bearings 2012 (ISMB13)*, 8 2012.

[12] A. Kameari, "Magnetic Field Analysis Using Hybrid Mesh of Linear Hexahedral and Tetrahedral Edge Elements," *IEEE Transactions on Magnetics*, vol. 44, pp. 1182–1185, 6 2008.

[13] V. Fedak, P. Zaskalicky, and Z. Gelvanič, "Analysis of Balancing of Unbalanced Rotors and Long Shafts using GUI MATLAB," in *MATLAB Applications for the Practical Engineer*, InTech, 9 2014.

[14] F. Impinna, J. G. Detoni, A. Tonoli, and N. Amati, "Test and Theory of Electrodynamic Bearings Coupled to Active Magnetic Dampers," *14th International Symposium on Magnetic Bearings*, pp. 263–268, 8 2014.

BIOGRAPHIES

David Rura was born in Presov, Slovak Republic in 1995. He received his M.Sc. in power electrical electronic engineering from Brno University of Technology, Brno, the Czech Republic in 2019.

He is currently a postgraduate student and technical staff in research and development at the Brno University of Technology, Brno, Czech Republic. His research interest includes electromagnetic design of various types of magnetic bearings including mainly active and electrodynamical magnetic bearings.

Analysis of Permanent Magnet Segmentation Influence on Eddy-Current Losses in High-Speed PM Synchronous Machine

Daniel Pribulla
Department of Power Electrical and Electronic Engineering
Brno University of Technology
Brno, Czech Republic
Email: xpribu02@vutbr.cz

Petr Klíma
Department of Power Electrical and Electronic Engineering
Brno University of Technology
Brno, Czech Republic
Email: petr.klima2@vut.cz

***Abstract*—An accurate prediction of eddy-current losses in electrical machines remains challenging. This paper studies the influence of segmentation on these losses in case of a high-speed surface mounted permanent magnet synchronous machines. Main sources of air-gap harmonics, causing the eddy currents were discussed as well as typical analytical methods used to model them. A 2-D finite element model of the studied machine with various number of segments was created to study the influence of these harmonics. It was found that pulse width modulation harmonics become dominant source of eddy-current losses with low switching frequencies. Certain combinations of switching frequencies and number of segments resulted in higher losses, than permanent magnet layer without the segmentation.**

***Index Terms*—high-speed, synchronous, machine, motor, PMSM, eddy currents, segmentation, loss, PWM, harmonics, 2-D FFT**

I. Introduction

The surface-mounted permanent magnet synchronous machine (PMSM) has become one of the most popular topologies when it comes to design of a high-speed machine. It is mainly because its superior efficiency, high achievable speeds and relatively simple construction [1].

Although this type of high-speed machine is now considered mature and well studied, there are still a few culprits that represent a challenge during the design of a PMSM. One of them is an accurate estimation of eddy-current losses in the volume of the permanent magnets [2]. While their impact on overall efficiency of the machine is generally minor, heating up of the permanent magnet (PM) material affects strength of adhesives, correct interference fit of the retention sleeves and in extreme cases may lead to demagnetization or release of the magnet and catastrophic failure of the machine. Thus correct estimation of eddy-current loss in PM is very important during the machine design.

II. Induction of Eddy Currents

Ideally, the air gap magnetic field distribution should appear stationary in the rotor reference frame, but in reality various effects introduce spatial harmonics rotating along the air-gap with different speeds. Resulting fluctuation of magnetic field on the surface of the PM causes eddy currents to form in the volume of the PM. The conductivity of the NdFeB magnet material is comparable with the conductivity of stator sheets, but laminated magnets are still not easily accessible, thus even small amplitudes of the harmonics can be an issue. There are three major sources of harmonics present in the air-gap of the machine - permeance, saturation and winding harmonics [3].

A. Permeance Harmonics

Permeance Harmonics are caused by a decrease of magnetic permeability in the stator slots. They create a series of harmonic multiples of number of stator slot in reference to the mechanical frequency of the rotor [4]. The high-speed machines usually have wide air-gap in the relation to the stator slot opening width, so their effect is low.

B. Saturation Harmonics

Saturation harmonics are caused by the non-linearity of the stator and rotor core material. Only odd harmonics are created [3]. A magnetic circuit of high-speed machines is usually designed to keep the magnetic field way below the saturation in order to suppress the core losses [5], therefore suppressing this harmonics source.

C. Winding harmonics

The ideal sine distribution of magnetic field in the air gap is modulated by distribution of conductors into a finite amount of stator slots. Only certain harmonics can be present in the distribution, based on the number of phases, slots and winding pitch [6]. For a three phase, integral slot winding -5th, 7th - 11th etc. harmonic is typical. This effect is eminent in case of high speed-machines with a small stator bore leading to minimal amount of slots.

Another very important source of winding harmonics is a pulse width modulation (PWM) of the voltage, supplied from an inverter. The switching frequency is usually kept above 20 kHz to avoid an audible noise. Unlike in case of low speed machines, these modulation harmonics are in the same frequency range as the other types mentioned above.

III. Analysis of Eddy Currents

Various analytical models have been created to estimate the eddy-current losses. Since the induction of eddy currents is inherently a 3-D problem, some simplifications have to be made to be able to find a solution based on Maxwell's equation. The simplest solution, that assumes homogeneous layer of PM material is e.g. in [7]. This model is accurate, but neglects the effect of the segmentation. Models in [8] assume homogeneous magnetic field above the magnet and calculates the resistance of the eddy-current path. The improved model in [4] accounts for a penetration depth in relation to the depth of a magnet. In order to sum up the influence of all above mentioned harmonics sources, the harmonic superposition principle is used [9]. This means that amplitudes and frequencies of harmonics from various sources are found, then the eddy loss is calculated and their power is summed up.

These models suggest that eddy-current losses are inversely proportional to the length and width of segments of the magnet, and for large number of segments approaching the formula used for core sheet iron. A 3-D finite element analysis of eddy-current losses from PWM modulation harmonics done in [10] and [11] suggest, that for very high frequencies, i.e. when depth of penetration is lower than thickness of a magnet, a small number of segments may result in higher losses than when not using the segmentation at all. Moreover, a study in [12] indicates, that the harmonics superposition principle doesn't hold for segmented structures well.

IV. Analyzed Machine

The analyzed machine is a high-speed surface-mounted permanent magnet synchronous motor. It's main parameters are shown in Table I. The PMs used in the machine are NdFeB and their temperature in the simulations is assumed to be 120°C, which was calculated by the early temperature model for 1 magnet segment with sine current supply. The temperature of the magnets was kept constant for all the calculations to include only the influence of the supply current in the results. The winding is single layer and is made of the Litz wire to suppress the eddy-current losses in it. The stator core material is NO-20, which is suitable for the high fundamental frequency of the stator flux. As the PMs are on the surface of the rotor, they are held by carbon fiber sleeve to ensure the total force affecting them is compressive and not tensile.

To calculate the eddy-current losses, a 2-D transient analysis using finite element method was used. The time step size of the simulation was 1/1000 of the period, which is smaller than recommended 1/400 of the period by [6]. The smaller step size was chosen to properly calculate the losses caused by high order harmonics. To accelerate the simulation time, only a quarter of the machine was modeled by a usage of symmetries. The model used for simulations is shown in Fig. 1.

The mesh in the model was created to be very dense because smaller depth of penetration of the high order harmonics. The detail of the mesh in the PM area is shown in Fig. 2.

TABLE I
Main parameters of analyzed machine

Parameter	Unit	Value
Rated Power	W	9 000
Rated Speed	rpm	60 000
Number of Poles	-	4
Stator Slots	-	24
Stator Outer Diameter	mm	120
Stator Axial Length	mm	50
Rotor Outer Diameter	mm	60
Radial Air-Gap Length	mm	2
Permanent Magnet Height	mm	2
Permanent Magnet Material	-	N40UH

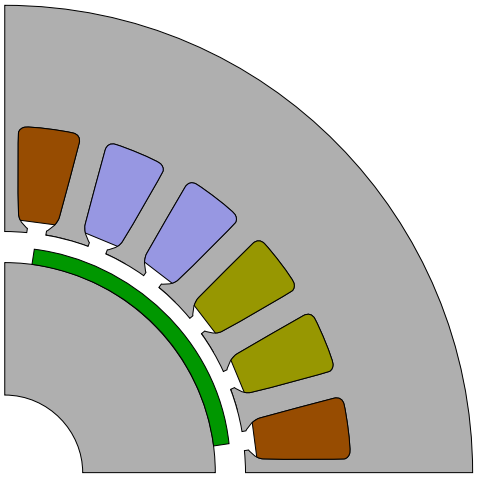


Fig. 1. Model of the analyzed machine used in FEM calculations.

The 2-D model used in the simulation assumes, that all the faces are electrically interconnected in infinity. This means that the end resistance of the PMs was neglected. Because of this, it is necessary to set condition of

$$\int^{S} \mathbf{J}\, dS = 0$$

to each face of the PM segment in order to account for the segmentation of PMs in the eddy-current calculation.

V. Segmentation Analysis

The influence of PM segments was analyzed in three steps. First, the total losses in PMs for different number of segments and different input currents were studied. Then, the 2-D Fast Fourier Transform (FFT) was performed to uncover the cause of the losses in the first step. At the last step, the reduction of losses in PMs was taken into the context of total losses in the machine.

A. Influence of PM Segments on PM Losses

The calculations were made for different PWM frequencies of the converter. This mainly influences the order of higher harmonics and their amplitudes of the machine input current. The assumed q-axis input currents for different PWM frequencies are shown in Fig. 3.

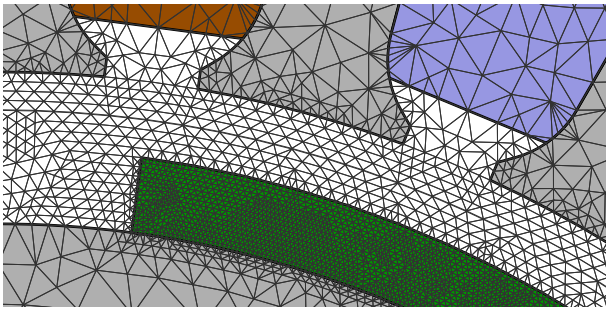


Fig. 2. Detail of the mesh in the PM area.

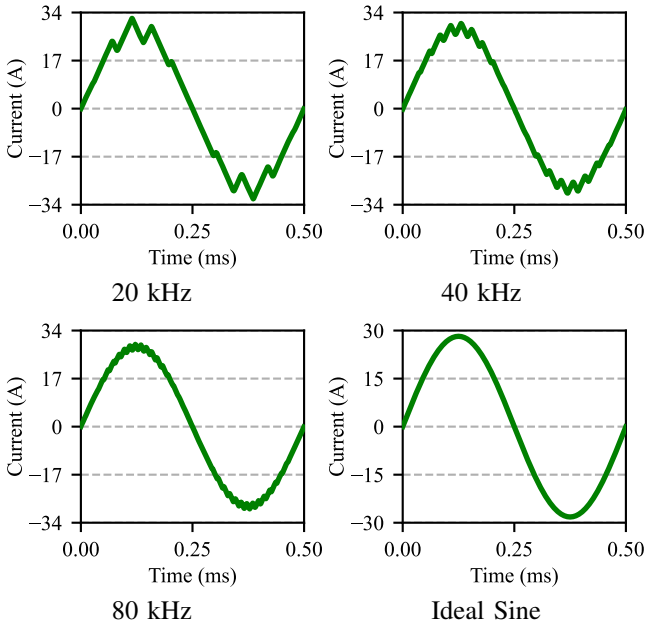


Fig. 3. Input currents for the specified frequencies of PWM and the ideal input current.

Fig. 4 shows the harmonic content of the input currents. It is clear, that for 2 kHz frequency of the carrier, 20 kHz switching frequency is very low and the current ripple is very high. All the modulation harmonics are very close to the fundamental frequency of the machine.

The calculated total eddy-current losses in PMs per number of segments for different input current PWM frequencies are shown in Fig. 5. The highest losses occurred when the motor was supplied by the 20 kHz PWM due to high amount of modulation harmonics. The absolute peak was achieved with two segments and 20 kHz PWM where the eddy loss was 1.8 times higher than the pure sine supply. In this case the eddy-current loss caused by PWM was almost the same as the sum of all the other eddy current sources described in section II.

For all supply currents but the ideal sine, the losses are not the highest when using 1 segment of PM. The highest losses with 20 kHz PWM is at 2 segments, for 40 kHz at 3 segments and for 80 kHz at 5 segments. Only after these "critical" number of segments, the losses start to decrease.

For the ideal sine current, the losses are the highest for

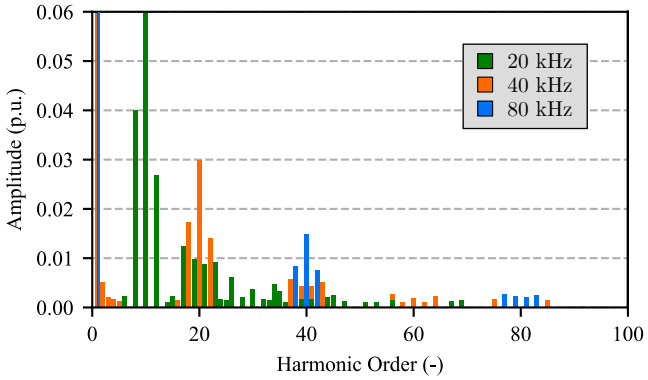


Fig. 4. Harmonic content of the input currents. The first harmonic for all the frequencies is 1.0.

1 segment. The losses are constantly decreasing with the increase in number of segments, but the decrease in losses is more steep after the number of segments is higher than 6. Losses at no-load operation are nearly the same for 1 to 5 segments, while the highest losses are at 1 and 5 segments. They are rapidly decreasing for more than 5 segments.

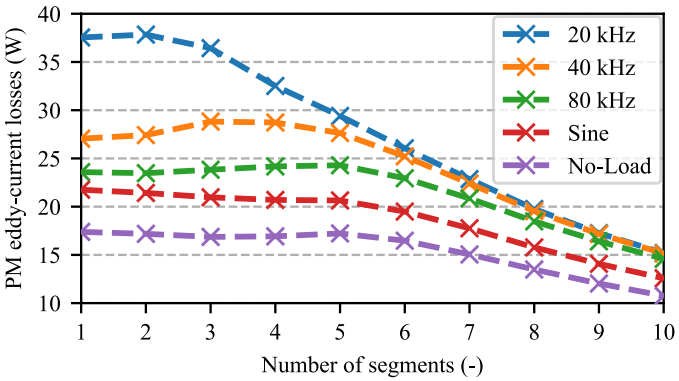


Fig. 5. Losses per number of PM segments.

According to these results, 8 segments and 40kHz switching frequency were selected into the final design, as they provide a notable suppression of the PM eddy-current losses for all the tested frequencies.

Fig. 6 shows current density for 1, 4, 7 and 10 segments. The time of the field save (the end of second period) is the same for all the subfigures. The maximal values do not rapidly change, but the area with high values decreases at 7 and 10 segments, which results in lower losses.

B. Harmonic Analysis

To properly analyze the causes for the loss dependence in Fig. 5, a 2-D FFT in analysis of flux density in the air-gap is required. Two dimensions of the analysis are needed, because the flux density is time and space dependent. If only classic 1-D FFT was used, the results would be dependent on the time of the flux density data save.

Fig. 7 shows a composition of harmonics without segmentation at nominal load with the ideal sine voltage supply. The

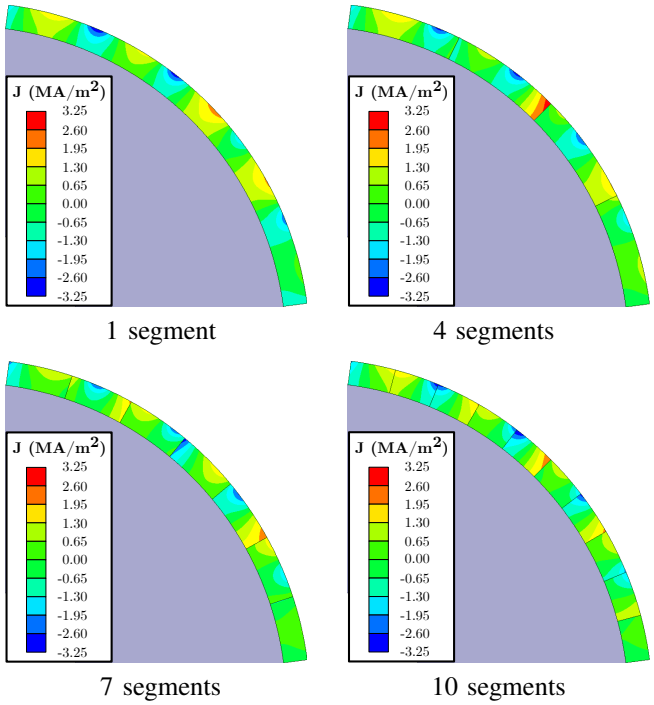


Fig. 6. Current density in PM for different number of segments with input current from 40 kHz PWM at the end of the 2nd period of the simulation.

blue harmonics result from trapezoidal distribution of PM magnetic field, but since this is bound to the PM rotation they have zero frequency and cannot generate any eddy-current losses. The orange peaks come from the slot permeance and winding harmonics.

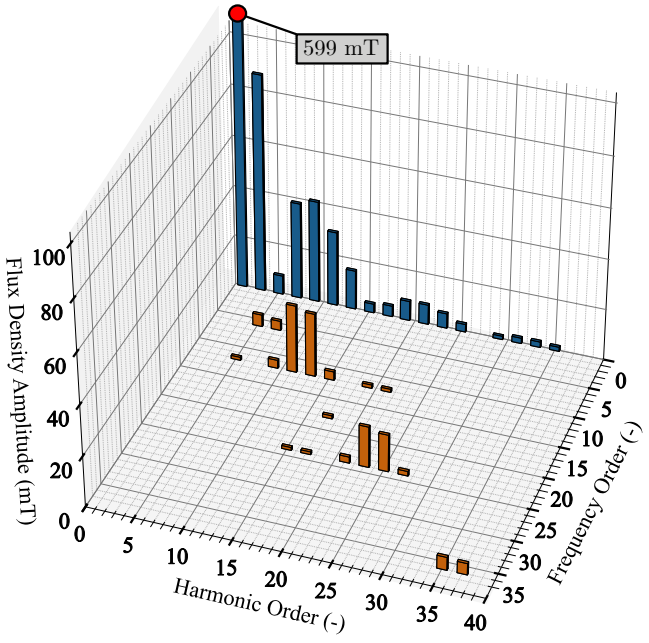


Fig. 7. Results of 2-D FFT analysis of the flux density in the middle of the air-gap for ideal sine input current and 1 segment.

In order to analyze and explain the results in figure 6, a series of 2-D FFT plots was created for various PWM switching frequencies. The result with pure sine supply in Fig. 7 was subtracted to show the differences only (Fig. 8). As expected, the main difference was found at the frequency order, which is a fraction of the PWM switching and fundamental frequency of the machine. They have relatively small amplitudes, but since their harmonic order is low, they have to rotate at a high angular frequency along the air-gap, compared to winding and slot harmonics.

C. Total Machine Losses

The losses produced in the machine are listed in Tab. II. The main source of the losses are the iron losses. This is mainly because of high fundamental frequency of the machine, which causes high losses even at low values of flux density in the stator.

TABLE II
AN OVERVIEW OF LOSSES IN THE ANALYZED MOTOR

Losses Type	Unit	Value
Copper Losses	W	124.7
Iron Losses	W	243.3
Mechanical Losses	W	143.0
PM Eddy Losses	W	18.6
Total Losses	W	529.6

The mechanical loss, caused by the high-speed rotor rotation, is the second biggest source of the losses. They were calculated using the analytical equations in [6]. Both the active rotor part and rotor ends mechanical losses were accounted for. Based on measurements in [13] the value of the roughness coefficient used in the calculation was estimated to 1.2, as the space between the PM poles is supposed to be solid, filled with non-magnetic material and the carbon sleeve should be smooth.

The copper losses are nearing the mechanical losses. The length of the machine is restricted by the critical speed of the machine, which causes a relatively high ratio of end winding to active length and increases the stator winding resistance. A low copper fill factor of the Litz wire also increases the resistance.

As expected, the PM eddy-current losses are low compared to the other sources (3.5% of total losses), but given the small volume of PM material, 18.6 W of heat will cause a significant increase in the PM temperature.

VI. CONCLUSION

An eddy-current losses in surface-mounted high-speed PMSM were studied using 2-D, transient FEM model in order to choose number of PM segments. It was found, that when machine is supplied with sine voltage, there is no benefit in dividing the PM layer into less than 6-8 segments. This can be explained by short pole-pitch of all significant harmonics present in the air-gap magnetic field distribution. Relatively wide PMs create no barrier for these harmonics. Thus the segmentation design is mainly a technological task, to simplify

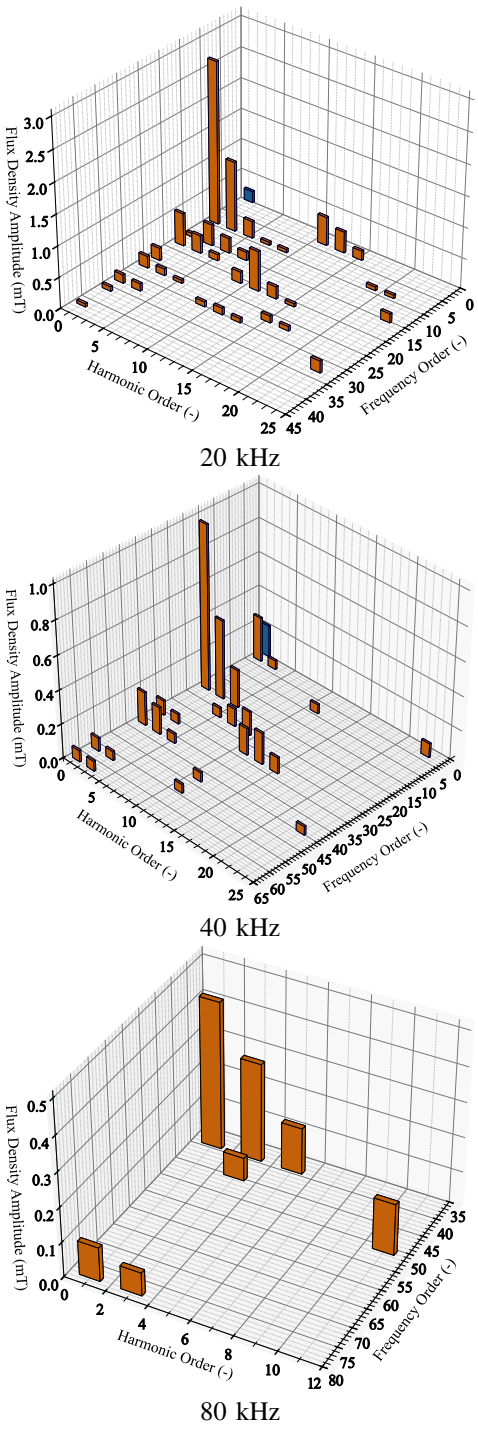


Fig. 8. Result of 2-D FFT of flux density in the middle of the air-gap for different PWM frequencies after subtraction by the sine current flux density.

the manufacturing of the rotor. The PWM inverter supplied voltage can almost double the losses, when too low PWM switching frequency is used. The eddy-current loss is not decreasing monotonously, but has a peak value for certain PWM switching frequencies. In [10], it is suggested, that the highest eddy-current losses are achieved, when the depth of penetration of the harmonics is equal to the half of the PM width. A future analysis of current density distribution in PM volume is needed. It is clear, that certain combinations of PWM frequencies and numbers of PM segments should be avoided. Because of this, 8 segments and 40kHz switching frequency was selected as the final design. The resulting 18 W of a dissipated heat seams acceptable, but a further thermal analysis is necessary. If technologically possible, a PM without segmentation may also be a good choice in case of the high-speed PMSM.

Acknowledgments

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-23-8430).

Authors acknowledge Mr. doc. Ing. Ondřej Vítek Ph.D. as their research supervisor for his valuable comments and help while writing this paper.

References

[1] GERADA, David, Abdeslam MEBARKI, Neil L. BROWN, Chris GERADA, Andrea CAVAGNINO and Aldo BOGLIETTI. 2014. High-Speed Electrical Machines: Technologies, Trends, and Developments. In: *IEEE Transactions on Industrial Electronics*. **61**(6), p. 2946-2959.

[2] LI, Silong, Yingjie LI, Wooyoung CHOI and Bulent SARLIOGLU. 2016. High-Speed Electric Machines: Challenges and Design Considerations. *IEEE Transactions on Transportation Electrification*. **2**(1), 2-13.

[3] DI, Chong, Ilya PETROV and Juha J. PYRHONEN. 2019. 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*. **7**, 27746-27755.

[4] PYRHONEN, Juha, Hanne JUSSILA, Yuliya ALEXANDROVA, Pavol RAFAJDUS and Janne NERG. 2012. Harmonic Loss Calculation in Rotor Surface Permanent Magnets-New Analytic Approach. In: *IEEE Transactions on Magnetics*. **48**(8), p. 2358-2366.

[5] PRIBULLA, Daniel. 2021. Design Of A High-Speed Generator For A Helium Expansion Turbine. In: *Proceedings I of the 27st Conference STUDENT EEICT 2021: General papers*. Brno.

[6] PYRHÖNEN, Juha, Tapani JOKINEN and Valéria HRABOVCOVÁ. 2014. *Design of rotating electrical machines*. 2nd edition. Chichester: Wiley.

[7] HELLER, Bedřich and Václav HAMATA. 1961. *Přídavna pole: síly a ztráty v asynchronním stoji* [Additional fields: forces and losses in asynchronous machine]. Prague: Czechoslovak academy of science.

[8] RUOHO, S., T. SANTA-NOKKI, J. KOLEHMAINEN and A. ARKKIO. 2009. Modeling Magnet Length In 2-D Finite-Element Analysis of Electric Machines. *IEEE Transactions on Magnetics*. **45**(8), 3114-3120.

[9] TESSAROLO, A., M. BORTOLOZZI and M. MEZZAROBBA. 2012. On the validity of the harmonic superposition principle for computing rotor eddy current losses in permanent magnet machines. In: *2012 XXth International Conference on Electrical Machines*. IEEE, p. 1369-1373.

[10] WAN-YING HUANG, Adel BETTAYEB, Robert KACZMAREK and Jean-Claude VANNIER. 2010. Optimization of Magnet Segmentation for Reduction of Eddy-Current Losses in Permanent Magnet Synchronous Machine. *IEEE Transactions on Energy Conversion*. **25**(2), 381-387.

[11] YAMAZAKI, Katsumi and Atsushi ABE. 2009. Loss Investigation of Interior Permanent-Magnet Motors Considering Carrier Harmonics and Magnet Eddy Currents. *IEEE Transactions on Industry Applications*. **45**(2), 659-665.

[12] CUSTERS, C. H. H. M., T. T. OVERBOOM, J. W. JANSEN and E. A. LOMONOVA. 2015. 2-D Semianalytical Modeling of Eddy Currents in Segmented Structures. *IEEE Transactions on Magnetics*. **51**(11), 1-4.

[13] SAARI, Juha. 1998. *Thermal Analysis of High-speed Induction Machines*. Helsinki. Dissertation thesis. Helsinki University of Technology.

Comparing of Evolutionary and Non-parametric Algorithms and Data Preparing

Martin Světlík
Dep. of Power Electrical
and Electronic Engineering
Brno University of Technology, FEEC
Brno, Czech Republic
xsvetl06@vutbr.cz

***Abstract*—This paper deals with modern approaches for electrical machine designing and data preparation for these methods. The discussed methods are the genetic algorithm, an algorithm from a group of evolutionary algorithms, and a decision tree, which belongs to non-parametric algorithms. This paper describes the genetic algorithm's main benefits, robustness, and high precision. Moreover, decision tree benefits like low computing time or low requirements for computing power are also discussed. Mathematical filters used for data preparation are necessary for reaching high accuracy of calculations. Bad chosen or used filters can lead to higher computing power requirements or increased computing time.**

***Index Terms*—Decision tree, Genetic algorithm, mathematical filters**

I. Introduction

Nowadays, the popularity of many approaches suitable for analyzing thermal and electrical parts of electrical machines designing increased. Many different approaches and countless methods have theoretically potential to become new revolutionary methods used for machine analysis and design. The main reason for using different approaches is the increasing complexity of the theoretical model used. That phenomenon is the consequence of many factors, e.g. higher requirements for machines efficiency or decreasing the number of manufactured prototypes. Increasing of real conditions and manufacturing imperfections approximation leads to saving more money during the designing process [1].

The high complexity of used models (lumped parameter thermal networks, electromagnetic descriptions of skin-effects etc.) requires high mathematical and physical essence knowledge. The methods based on mathematical and physical background are usually really robust and precise. Nevertheless, their calculations need high computing power and time which is the toll for being precise and robust. The evolutionary algorithms, especially genetic algorithms, are typically representative of this approach [2].

Non-parametric methods are based on machine learning and use so-called predictions to identify machines' performances. As mentioned, these methods are really useful for optimizing of electrical machines models, especially thermal, because they are very good in predictions in similar areas for which they have been trained. That gives these methods ability of fast optimizing, especially in cases where a lot of data have been calculated or measured. A typical representative of non-parametric methods is decision tree. Nevertheless, quality prepared data are still necessary [3], [4].

II. Evolutionary algorithms

Generally, the evolutionary algorithms are based on Darwinian evolution. These algorithms are usually very robust and powerful tools suitable for solving complex global problems, such as optimizations. The robustness is the main advantage in finding the local minimums where other methods may fail. On the other hand, these algorithms need more computing time, and the process of finding the global minimums requires high computing power. Moreover, knowing the mathematical and physical description of the solving problem is necessary. On the other hand, the merit is growing linearly with the increasing problem size [5].

The evolutionary group of algorithms includes, for example, the Genetic algorithm. The genetic algorithm is able to solve constrained and unconstrained problems. It is necessary to know the mathematical and physical background, so the electrical machine models must be created to allow this approach to design. The block scheme of the Genetic algorithm inspired by [7] is shown in Fig. 1. The principle of the function shown in the blocking scheme is as follows: First, the individuals from the first calculation are used to create the initial population, so-called parents. Then, these parents make the next generations which are scored, and the

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).

Furthermore, this work has been supported by the COMET-K2 "Center for Symbiotic Mechatronics" of the Linz Center of Mechatronics (LCM) funded by the Austrian federal government and the federal state of Upper Austria

Author acknowledges Assoc. Prof. Jan Barta, Ph.D. as my research supervisor for his valuable comments and help while writing this paper, and Assistant Prof. Marek Toman, Ph.D. as my Master thesis superivisor who gives me valuable advices and helps me reach knowledges neccesary for writing this paper.

children that reach better results than their parents are used for the creation new generation of parents. Third, creating new generations is repeated until the required results are not reached.

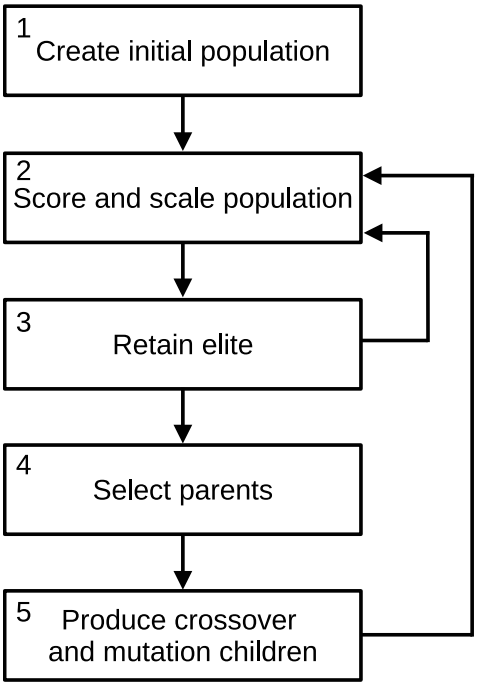


Fig. 1. Block scheme of Genetic algorithm [6].

There are three rules applied for creating a new generation. These rules are:

- Selection rules
- Crossover rules
- Mutation rules

Block diagram of selection rules is shown in Fig. 2. The principle of function is based on the selection of parents to become children directly. The parents are chosen due to individual scores or even randomly.

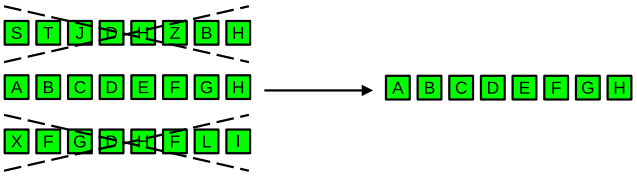


Fig. 2. Selection rule block scheme inspired by [6], and [7].

Crossover rule is based on two parents combining. The parents used for the combining rule are chosen randomly, and a completely new individual is created after that combination. A block diagram is shown in Fig. 3. [6] and [7].

Last used rule is called the mutation rule. That rule is based on the application the random changes to the parents' genome. The changed parents become the new children. This rule is graphically shown in Fig. 4.

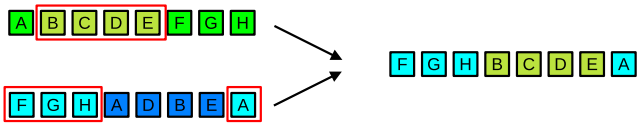


Fig. 3. Crossover rule block scheme.

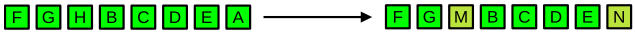


Fig. 4. Mutation rule block scheme [6], [7].

The genetic algorithms are not perfect. For example, some of the disadvantages of the genetic algorithm are high requirements for computing power and computing time. These disadvantages can be reduced by chosen data or by precise initial setup. However, mentioned disadvantages are not present in the next described approach, which is an approach that uses non-parametric methods.

Data, used as input for the genetic algorithm, where the interference is present can negatively influence the computing time of calculation and also require computing power. For that reason, it is convenient to use at least some type of mathematical filter that improves the input data and makes the curve more smooth. Using a relatively simple mathematical filter can lead to much faster calculations.

III. Non-parametric methods

Non-parametric methods are not based on mathematical models, which means it is not necessary to know the mathematical and physical background as well as in the genetic algorithm case. These methods allow well-known input data with continuous values and not so complex models to be modeled. On the other hand, the methods are really fast and for the well known areas are able to reach a high accuracy of predictions.

Decision tree is a method used for modeling complex patterns in data. The goal of decision trees is to create models that are based on learning simple decision rules. Learned rules, so-called features, are created due to input data. Block diagram of a simple decision tree is shown in Fig. 5. The presented decision tree is created for a single binary target Y and two variables X_1 and X_2, which are continuous. Outputs R are defined by the combination of input and variables. Decision tree presented in Fig. 5 consists of nodes and branches [8].

The data used for decision tree learning can be based on calculation considering the physical nature of the machine or on the measurement where the physical nature does not have to be fully understood. Nevertheless, the decision tree essence does not use a physical nature of things. Features for the decision tree can be chosen from almost every machine's input or output. However, it makes sense to choose inputs that have a significant impact on the machine. Therefore,

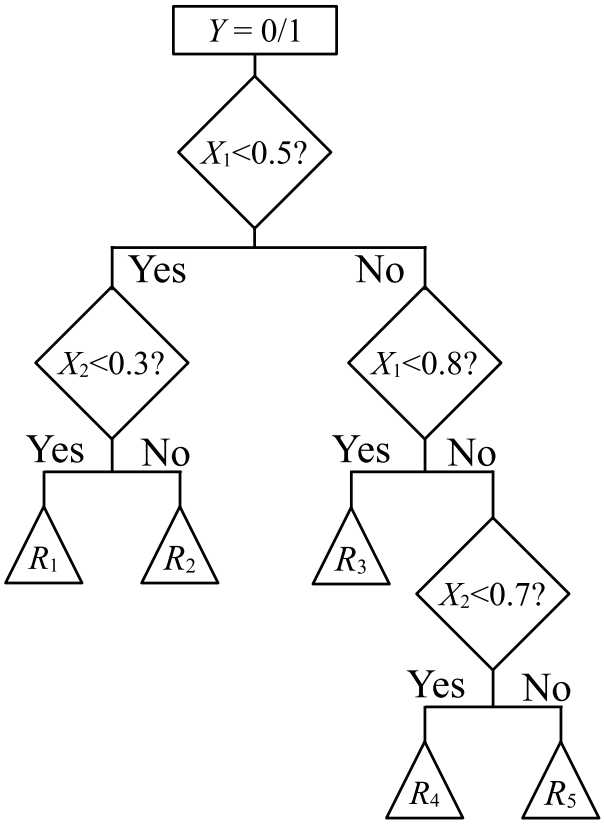


Fig. 5. Block scheme of Decision tree [8]

it makes sense to focus on variables like speed or current. Examples of usable features are shown in Tab. I.

TABLE I
TABLE OF FEATURES USABLE FOR THE DECISION TREE CREATING

Feautre	Used data
f_1	I
f_2	I^2
f_3	In
f_4	n

It is necessary to keep the Decision tree simple enough. So much complex Decision trees may not generalize the data well. The phenomenon when the decision trees are so much complex is called overfitting. There are several existing ways of protecting trees before overfitting. First, setting the minimum number of samples. second, setting the maximum depth of the tree. The reason for overfitting is dangerous is that the predictions are only piecewise constant approximations.

A. Nodes

Nodes are one of the two basic building blocks of Decision trees. There exist three types of nodes in the construction process of decision trees. First node type is so-called decision node or root note. Root nodes represent a choice into two or more mutually exclusive subsets. Second type of node represents one of the possible choices at the point in the structure. Top of the node is connected to the parent node, and the bottom is connected to the child or leaf node. That node is called an internal node or o chance node. Third type of node is called the leaf node and represents the final result of the Decision tree.

B. Branches

The second basic building block of the Decision tree structure is the branches. The branches' hierarchy is used for forming the Decision tree. The branches represent random results from internal or root nodes. The decision rule is represented by the connection way from the internal to the leaf node. These paths consist of several branches that can be imagined as the "if" rule.

Three steps must be used to keep the Decision tree detailed enough but out of the overfitting point. These steps are:

- Splitting
- Stopping
- Pruning

The splitting is applied to parent nodes. The nodes modified by splitting are divided into purer child nodes. The splitting is used repeatedly until are not reached only homogenous nodes. The stopping steps are necessary to use in the Decision tree creating process. The reason for its necessity is to protect the three from becoming too complex and overfitting. Too complex trees lose their ability of accurate predictions. The last criterion of Decision tree designing is called Pruning. The Pruning criterion is an alternative to the Stopping criterion. It is because The stopping criterion does not work very well in some situations. The alternative builds the large tree first and then optimizes it to the ideal size. The process of optimizing consists of removing nodes that provide less additional information [9], [10] .

The Decision trees are much more sensitive to input data than the Genetic algorithms. The main reason is that this method is more simple and works only on simple "if" rules. The interference of input data can lead to a significant decrease in the accuracy of their predictions. Moreover, using mathematical filters has almost zero impact on the speed of the analysis proved by the Decision tree. So same as for the Genetic algorithm case, it is important to prepare input data and remove the interference.

IV. DATA PREPARATION

Preparation of input data used for the analysis based on described methods is one of the most essential parts of any calculation type. This paper deals with two different calculation methods, and several data preparation methods will be discussed in the next chapter. Generally, it depends on the type of used data. The data mined by simulations and then used for other calculations compared to data mined by measurement usually does not need many preparations.

Therefore, a curve of capacitor charging voltage with non-filtered interference has been created for this paper.

For the purpose of the paper, as mentioned, the data simulating the capacitor charging and discharging with a high level of interference have been generated. The effect and accuracy of chosen mathematical filters are demonstrated on these curves of voltage. Moreover, the accuracy of the signal reconstruction is presented in Tab. II.

This paper is focused on two types of mathematical filters. These types are:

- Moving average (MA)
- Exponential moving average (EMA)

The moving average is one of the simplest ways of data filtrating. It is very easy to use and understand and is also one of the most used filters for the digital signal processors (DSP). This filter is good for reducing the measured inferences, so despite its simplicity, it is very efficient for data preparation. The formula describing the moving average is according to [11] defined as

$$x_{\mathrm{n}} = \frac{y_{\mathrm{n}} + y_{\mathrm{n+1}} + y_{\mathrm{n+2}} + \ldots + y_{\mathrm{n+w}}}{w}, \tag{1}$$

where the x_{n} is the averaged value of n-th sample, the y_{n} is the value of n-th data sample, and w is the number of values used to average so-called the window. The results of moving average application are shown in Fig. 6.

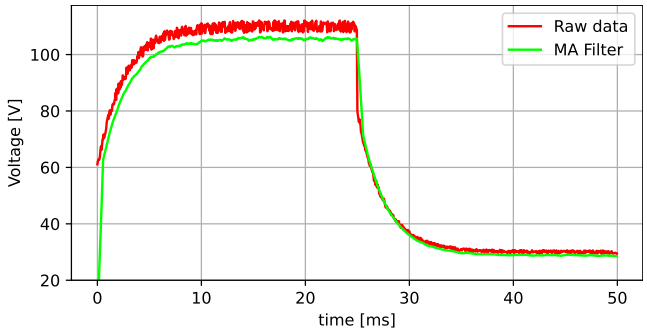


Fig. 6. Curves of measured data and applied classical approach of moving average.

The accuracy of the classical moving average approach calculated by 4 is equal to 3.349 V. The main reason for the low signal reconstruction accuracy is due to starting the reconstruction at the zeros value.

Therefore, the better way of using the moving average is ”move” the window into the past. This approach to moving the average is more practical because the filter does not start at zero but starts near the real measured values. Moreover, the Decision tree does not consider the zeros in the first steps, so reorientation of the window into the past makes the signal reconstruction more accurate. This phenomenon is defined as

$$x_{\mathrm{n}} = \frac{y_{\mathrm{n\text{-}w}} + y_{\mathrm{n\text{-}w+1}} + y_{\mathrm{n\text{-}w+2}} + \ldots + y_{\mathrm{n}}}{w} \tag{2}$$

and is shown in Fig. 7.

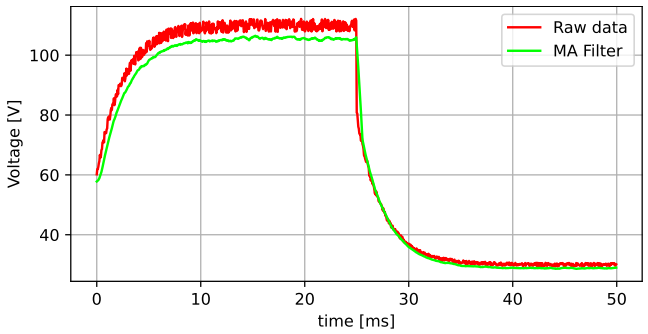


Fig. 7. Curves of measured data and applied moving average with window reoriented to past.

The accuracy increased from 3.349 V to 3.305 V. That means increasing reconstruction accuracy by 0.5 A was enough to change the window’s position.

The exponential moving average is another type of mathematical filter. Thanks to the different signs of the curve’s points have this filter potential to reach higher accuracy of results. The higher accuracy is reached thanks to the curve smoothing and, as a consequence removing the interference. [12], [13]. The exponential moving average is according to [14] defined as

$$x_{\mathrm{n}} = y_{\mathrm{n}} w + (1 - w) x_{\mathrm{n\text{-}1}}, \tag{3}$$

where the $x_{\mathrm{n}-1}$ is previous calculated value.

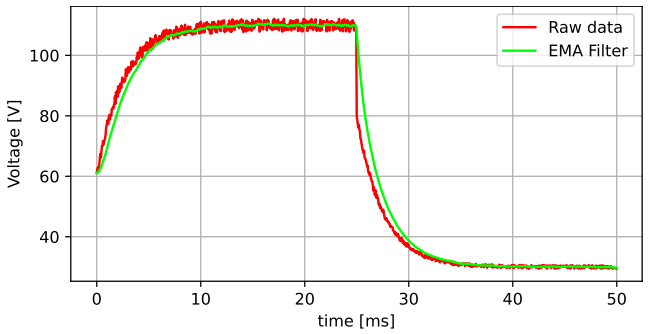


Fig. 8. Curves of measured data and applied exponential moving average

At first glance, it is clear that the accuracy has increased. The final value of signal reconstruction based on the exponential moving average is equal to 1.885 V.

The accuracy of signal reconstruction has been calculated

like mean squared error, which is according to [15] defined as

$$\varepsilon = \frac{1}{n}\sum_{i=1}^{n}\left(U - U_{\mathrm{r}}\right)^2, \tag{4}$$

where n represents the number of measured points, U represents the values of interfered voltage, and the U_{r} represents the reference Voltage.

The mean squared errors calculated for all three cases are shown in Tab II. There is proven that the worst prediction accuracy has been reached by using the classical moving average filter. On the other hand, the best accuracy has been reached by using the exponential moving average.

TABLE II
TABLE OF SIGNAL RECONSTRUCTION ACCURACY

Error	Moving average (first case)	Moving average (second case)	Exp. moving average	Unit
ε	3.349	3.050	1.885	V

V. CONCLUSION

First, two different methods used for electric machine analysis have been discussed. One of them is the Genetic algorithm which is part of the evolutionary algorithms. This method is really robust and precise. On the other hand, needs high computing power and computing time. The following method, the Decision tree, has been discussed. This method is part of the group called non-parametric methods. This method is much faster and does not need so much computing power. On the other hand, it is relatively sensitive to input data. In light of this information, the best way of designing electrical machines is combining those two methods to use both advantages. This will be the subject of future research.

Second, the mathematical filters suitable for preparing input data have been described. These filters are convenient to eliminate the disadvantages of discussed methods. The two approaches of moving average and exponential average have been used for the data preparations. The accuracy of these filters has been tested by the capacitor charging and discharging waveform. This waveform has been calculated and considered interference also. The best accuracy has been reached by exponential moving average. Graphically representations of these filters are shown in the Fig. 6, Fig. 7 and Fig. 8. The accuracy is summarized in Tab. II.

REFERENCES

[1] S. Paul and J. Chang, "Fast model-based design of high performance permanent magnet machine for next generation electric propulsion for urban aerial vehicle application," in CES Transactions on Electrical Machines and Systems, vol. 5, no. 2, pp. 143-151, June 2021, doi: 10.30941/CESTEMS.2021.00018.

[2] E. Schmidt, M. Sušć and A. Eilenberger, "Finite element analysis of a permanent magnet synchronous machine with an external rotor for a position sensorless control," 2009 Australasian Universities Power Engineering Conference, Adelaide, SA, Australia, 2009, pp. 1-6.

[3] BROWNLEE, Jason. Parametric and Nonparametric Machine Learning Algorithms: Nonparametric Machine Learning Algorithms. *Welcome to Machine Learning Mastery* [online]. 15.8.2020 [cit. 2023-03-09]. Available from: https://machinelearningmastery.com/parametric-and-nonparametric-machine-learning-algorithms/

[4] Decision Trees. *Scikit-learn: Machine Learning in Python* [online]. 2007 [cit. 2022-05-02]. Available from: https://scikit-learn.org/stable/modules/tree.html

[5] MATSUND, K., A. ECER, J. PERAIENUX, N. SATOFUKA a P. FOX, 2002. *Parallel Computational Fluid Dynamics 2002*. 1. Japonsko. ISBN 978-0-444-50680-1.

[6] Genetic Algorithms. *GeeksforGeeks* [online]. Noida, 22.2.2022 [cit. 2022-03-26]. Available from: https://www.geeksforgeeks.org/genetic-algorithms/

[7] *MathWorks* [online]. The MathWorks, 1994 [cit. 2022-03-26]. Available from https://www.mathworks.com/?s_tid=gn_logo

[8] Decision trees in more detail. *Clustering and Classification methods for Biologists* [online]. Menchester Metropolitan University: Menchester Metropolitan University [cit. 2022-05-02]. Available from: http://www.alanfielding.co.uk/multivar/crt/dt_example_04.htm

[9] *Towards Data Science* [online]. Towards Data Science, 2018 [cit. 2022-05-02]. Available from: https://towardsdatascience.com/decision-tree-in-machine-learning-e380942a4c96

[10] TAHIR, Nooritawati Md, Aini HUSSAIN, Salina Abdul SAMAD, Khairul Anuar ISHAK a Rosmawati Abdul HALIM. Feature Selection for Classification Using Decision Tree. *2006 4th Student Conference on Research and Development* [online]. IEEE, 2006, 2006, 99-102 [cit. 2022-05-02]. ISBN 978-1-4244-0526-8. Available from: doi:10.1109/SCORED.2006.4339317

[11] SMITH, Steven W. Moving Average Filters. *Digital Signal Processing* [online]. Elsevier, 2003, 2003, 277-284 [cit. 2022-04-27]. ISBN 9780750674447. Available from: doi:10.1016/B978-0-7506-7444-7/50052-2

[12] Understanding Moving Averages. *CME Group* [online]. 2022 [cit. 2022-04-27]. Available from: https://www.cmegroup.com/education/courses/technical-analysis/understanding-moving-averages.html

[13] What is a Moving Average?. *CFI Education Inc.* [online]. 2015 [cit. 2022-04-27]. Available from: https://corporatefinanceinstitute.com/resources/knowledge/other/moving-average/

[14] BURGSTAHLER, Lars a Martin NEUBAUER. *New Modifications of the Exponential Moving Average Algorithm for Bandwidth Estimation* [online]. Institute of Communication Networks and Computer Engineering: University of Stuttgart [cit. 2022-04-27]. Available from: http://content.ikr.uni-stuttgart.de/Content/Publications/Archive/Bu_ITCSS15-49_34635.pdf

[15] Machine learning: an introduction to mean squared error and regression lines. *FreeCodeCamp.org* [online]. USA, 2018 [cit. 2022-05-11]. Available from: https://www.freecodecamp.org/news/machine-learning-mean-squared-error-regression-line-c7dde9a26b93/

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.

e·on

Nastartuj svou kariéru

#spolujsmeeon

Studentské programy
- Letní energetická akademie – poslední týden v červenci
- Letní technické brigády – od června do září
- Stipendijní program
- Diplomové/bakalářské práce
- Praxe a brigády

Absolventské programy
(plný úvazek)
- Trainee program – od září
- Junior technik – od září

eon.cz/kariera

E.ON Kariéra E.ON Kariéra



Eddy Currents Additional Losses Comparison in Oil Distribution Transformer Windings

1st Miroslav Mrajca
Department of Power Electrical and Electronic Engineering
Faculty of Electrical Engineering and Communication, Brno University of Technology
Brno, Czech Republic
Miroslav.Mrajca@vut.cz

***Abstract*—This paper examines the additional losses caused by eddy currents in various types of windings typically used for oil distribution transformers. Two models were selected for 2D simulation using the finite element method, one with only a core leg and the other with the core leg and yoke. The parameters of the models are based on real manufactured transformers of rated power 250, 630 and 1000 kVA. The final results are discussed and compared with measurements. The highest percentage of additional losses is found in the foil winding, followed by round, rectangular wire, and finally the lowest percentage is found in the flattened round wire. The presence of the yoke near the winding end significantly reduces these losses in foil and, conversely, increases them slightly in wire winding.**

***Index Terms*—eddy current, additional losses, foil, round wire, flattened wire, rectangular wire, winding, transformer**

I. Introduction

In 2021, the European Union implemented new limits on the load of power transformers and no-load losses in order to meet ecodesign requirements. An increase in material usage is necessary to ensure compliance with these limits. Therefore, it is essential to estimate losses accurately in order to reduce the costs associated with transformer oversizing, as there is no tolerance for tested loss values.

The load loss of a transformer consists mainly of a loss of ohmic resistance of the winding, which can be measured when powered by direct current (DC). However, losses created by the flow of alternating current (AC) are higher, and the difference is called additional losses (AL). These additional losses can be explained as stray losses attributed to the magnetic leakage field of the windings, the field of high current carrying leads and busbars. These fields create eddy currents in winding conductors, structural steel parts, e.g. clamps or oil tank, and also own leads and busbars. It is assumed that the most significant portion of these AL arises directly in the winding conductors. This paper analyzes them, as there is no comparison between AL in the different winding types used in oil distribution transformers, as their exact estimation is a very difficult task.

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-23-8430).

For a better comparison between different transformer designs, additional losses can be expressed as a percentage value $P_{AL\%}$ using the following equation:

$$P_{AL\%} = \frac{P_{AC} - P_{DC}}{P_{DC}} \cdot 100\% \tag{1}$$

P_{DC} and P_{AC} represent the losses generated by the DC and AC current when transmitting the same power, respectively.

Distribution transformers are typically used to reduce medium voltage (MV) to low voltage (LV). The ratio of turns is approximately equivalent to the voltage ratio and the current ratio is inversely proportional to this ratio. To accommodate these parameters, different technologies are used for both windings. The inner winding is commonly used for LV and is wound from the foil due to its high cross-sectional area, which is necessary for high currents. The foil width defines the height of the winding, which is usually higher than the depth of penetration, and also means the current density distribution is highly ununiform across the winding's cross-section. As only a few turns are required, the winding production is simple and fast. For this type, the approximative analytical [1] and the finite element method (FEM) of AL [2] can be found.

The outer winding is typically for MV and is wound as a multi-layer winding. Each layer consists of tens or hundreds of enamelled wire turns, separated by insulation paper. The wire can have a round, flattened round, or rectangular shape, but its dimensions are small, usually smaller than the depth of penetration but sufficient to conduct the current. From the view of the entire winding, the current density can be considered uniform. There are many research studies available on this type of winding, both analytical methods [3] and FEM [4], as this type is used more often in other transformer applications.

II. Parameters of Analyzed Transformers

For the case study, three three-phase hermetically sealed oil distribution transformers of rated power 250, 630 and 1000 kVA were chosen. These are very commonly used across Europe, and the voltage ratio chosen was 22/0.4 kV, which is the most common in the Czech Republic and their need can be counted in hundreds of pieces in this country. Examining different rated powers will show how the size of conductors influences AL. All parameters used for analysis come from real manufactured designs of transformers to better

TABLE I
PARAMETERS OF ANALYZED TRANSFORMERS

Rated power	**250 kVA**	**630 kVA**	**1000 kVA**
Rated voltage ratio		22/0.4 kV	
Short-circuit impedance	4%	6%	6%
Frequency		50 Hz	
Vector group connection		Dyn1	
Core diameter	140 mm	160 mm	190 mm
Straight core part length	153.18 mm	133.86 mm	138.00 mm
Window height	555 mm	745 mm	785 mm
Distance core - winding		2.5 mm	
LV - Foil thickness	0.8 mm	1.8 mm	2.2 mm
LV - Foil width	520 mm	710 mm	750 mm
LV - Foil insulation		0.15 mm	
LV - Number of turns	32	24	19
Distance between windings		9 mm	
MV - Wire diameter	2.5 mm	4.3 mm	5.0 mm
MV - Wire flattening	36%	32%	36.2%
MV - Wire elongation	5.318%	7.303%	9.812%
MV - Winding height	505 mm	695 mm	735 mm
MV - Winding width	42.39 mm	63.71 mm	61.44 mm
MV - Layer insulation	0.21 mm	0.28 mm	0.21 mm

verify the used method for simulation and also for possible redesigning in further production and reducing costs or improving transformer performance. The material of conductors in both windings is aluminium and the core consists of high permeability grain-oriented silicon steel H075-23L. The MV windings are shorter than the LV windings and are mutually axially centred, meaning that the shortenings on both outer winding ends are equal. All needed electrical and mechanical parameters for reproducing one phase of the active part in the vertical section of the analyzed transformer are provided in Tab. I.

III. SIMULATED MODELS AND ITS SIMPLIFICATIONS

A. Simplifications for Maximal Simulation Acceleration with Preserved Accuracy

For the simulation, Ansys Electronics software using FEM was used. Due to the large size of the real transformer and the small size of the analyzed parts, the resulting mesh is very fine. To obtain results in a reasonable time frame, the simulation has to be simplified as much as possible. The only applicable method is to use a 2D analysis similar to [5]. However, the windings do not have a purely cylindrical shape for the chosen rated powers of the transformers, as they are oval to reduce the pitch of the core legs and the no-load loss and the transformer price. This meant that the vertical cross-section is not the same around the core leg axis, resulting in the simulated model not being able to be simplified into only cylindrical symmetry around the axis Z.

Based on the literature referenced, the eddy currents in the windings are caused by a leakage magnetic field, which is created by the load current [6]. This field is greatly affected by the winding's dimensions and current magnitude, as well as the magnetically conductive (steel) parts such as core laminations, tightening construction, or tank walls. Other parts, such as oil or insulation paper, are represented by vacuum properties in

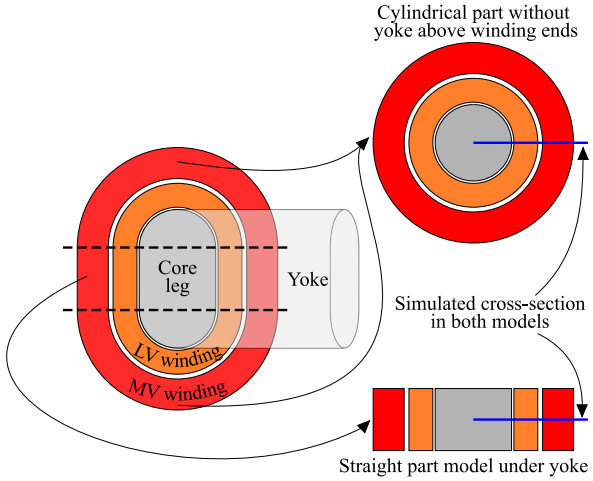


Fig. 1. Dividing oval winding on the side core leg into two models.

the model for the investigated problem. The oval shape of the windings can be divided into three parts, two of them are semicylinders, and the remaining part consists of straight block parts between them, as shown in Fig. 1. Most of the volume of block parts is located under the core yoke (middle phase) and tightening construction (side phase). We can assume that the leakage magnetic field and thus AL in the straight parts will be different from that in the cylindrical parts, which have no magnetically conductive body near their ends. The leakage field is also a little affected by the tank wall. However, its radial distance from the winding around its circumference is not the same, so it is impossible to include this factor in a 2D simulation.

B. Model Design

Based on these considerations, it appears that the most effective way to create an equivalent model to accurately calculate AL in winding conductors is to split the simulation into two models, visualized in Fig. 1, and then to sum both resulting losses for comparison with the real transformer. The first model utilizes a cylindrical shape for circular symmetry around the Z-axis, and the solution setup is based on an RZ cylindrical coordinate system. This model does not take into account the yoke above the windings' ends. It consists of cross-sections of the core leg, LV and MV windings, as shown in Fig. 2. The second model depicts blocks of core and winding in a Cartesian XY coordinate system, and the model is extended by the core yoke, as seen in Fig. 2, due to the majority of windings' ends here being located near the core yoke or tightening construction. It is important to note that using only two dominant geometries for creating an equivalent model does not fully capture the three-dimensional geometry of the transformer. This is because the leakage magnetic field is also affected in the third direction, which is not included in the 2D simulation. Furthermore, the simplification does not take into account the tiny angle of the half-cylindrical winding part that is under the core yoke, and the distance between them differs due to the stepped cross-section of the yoke. Therefore,

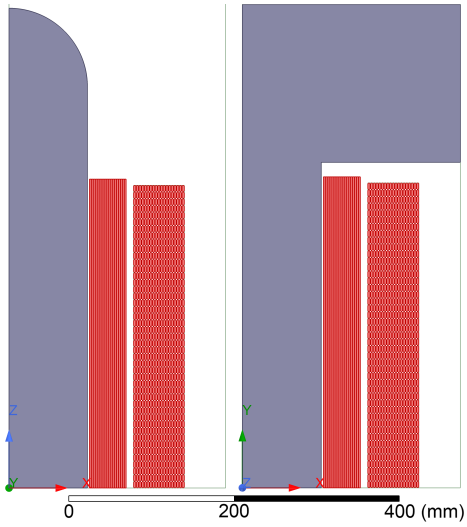


Fig. 2. Model of the cylindrical part without core yoke on the left and model of the straight part with the yoke on the right.

to obtain more precise results without using a 3D simulation, the simulation can be cut into more 2D models.

C. General Simulation Conditions

The model simulates the state during a short-circuit test, so the magnetizing current is negligible due to the much lower flux density than the transformer is built for. The BH curve of the core laminations can be linearized with a very high permeability of around 30,000, which should accelerate simulation. A sinusoidal voltage source powers the transformer and, since the outputs are shorted, the only load is the transformer's own impedance, which can be considered linear. Therefore, the load current should also be almost perfectly harmonic. For this reason, a transient simulation is not necessary and an eddy current analysis is sufficient. Further setup settings use adaptive mesh with maximal 10 passes or error 0.001% when the mesh reaches 4-5 million elements and the initial mesh inside the foil is set with the maximum element length half of its thickness and 0.5 mm inside the wire.

Another benefit of a short-circuit test is that the transformer is typically at ambient temperature. Since the test is very brief, and oil transformers have a large time constant, all elements can be assumed to be at the same temperature. For oil transformers, the load loss is then recalculated to the reference temperature of 75°C [7], and conductivity at this temperature is set for simulation. This setting does not accurately simulate a real short-circuit test, as the temperature during the test is usually around 20°C, and the measured losses are adjusted by the temperature coefficient of resistance. However, this coefficient is only correct for the ohmic loss; Eddy losses change disproportionately to the temperature rise. This fact has to be disregarded because the exact test temperature is not determined, so the reference temperature is used.

As the material and temperature properties are equal at the top and bottom of the windings, we can simplify the calculations by taking advantage of the symmetry of one phase, so the effect of other phases, including core legs, windings, and also their leakage field is omitted, which is allowed after conducting a trial simulation of the whole core window in 2D. Therefore, we only need to simulate half the height of the transformer's phase, with Neumann boundary conditions applied to the edge representing the middle of the winding's height. This is because the vector of magnetic flux should have only an axial component and be perpendicular to the symmetry line. Both models employ vertical symmetry of the core leg and windings, so for simulation purposes, only the radius of the core leg is used with a Dirichlet boundary condition on the bisector in the middle of the leg.

Each turn of foil in the LV winding cross-section is represented by a solid rectangle defined by the real width and thickness of the foil. The number of turns is determined by the measured designs. The turns of wire in the MV winding are simulated with three variations described in the following subsection III-D. However, the number of turns may vary depending on the wire shape to fill all layers of the winding, as an incomplete layer can change the magnetic leakage field and may not be the same on both ends of the winding for using symmetry. Additionally, in real winding on the MV side, the active number of turns can be adjusted by a tap changer. As a compromise, the number of turns in the model is determined by wire dimensions to fill the whole winding cross-section. However, the final ratio of LV and MV turns should be very close to the voltage ratio value. For simulation efficiency, all LV turns are excited by the value of the rated current in amplitude. On the MV side, this value is recalculated by the turns ratio to ensure equal ampere-turns.

D. Wire Shape Variations in MV Winding

For comparison, three different types of wire commonly used in MV windings of distribution transformers have been selected, illustrated in Fig. 3. The first type is a simple round wire, which is typically available up to a diameter of 5.3 mm for aluminium and 5 mm for copper. The second type is a flattened round wire, which is often used in the production of transformers instead of a full round wire. This process has the advantage of providing a higher filling factor and, as this paper shows, significantly reducing AL. The advantage of round wire

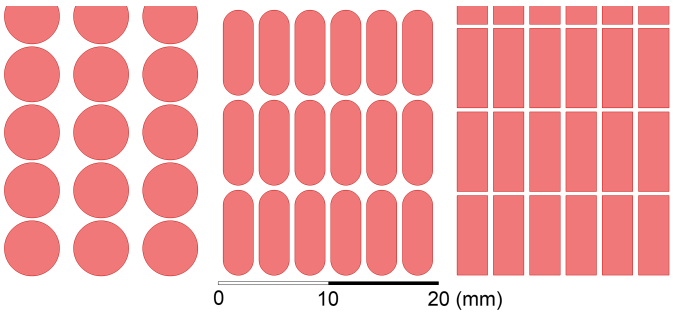


Fig. 3. Selected wire shapes for simulation, from the left: round wire, flattened round wire, and rectangular wire.

is that, with adjustable flattening, it is possible to cover a wide range of transformer designs up to a rated power of 2 MVA (depending on voltage ratio) with a small number of diameter series and can be easily accessed from stock.

It is challenging to mathematically describe the shape of a flattened wire, as the shape changes with tension and winding speed. However, a proven method is to use the same shape as the oval core introduced previously, which can be modelled as two solid semicircles and a rectangle between them. The difference between the new thickness and the original round diameter is defined by the percentage flattening factor. Additionally, the area of the cross-section is smaller as the wire is elongated during flattening, which affects the resulting width of the wire. For the comparison of conductor types, a flattening factor of 45% was chosen, as this is the maximum for the aluminium wire to show the maximal effect on AL. The elongation factor is included for comparison with measurement but is excluded for the winding types comparison to preserve the same cross-section area of the conductor.

The third type of wire selected for the comparison is rectangular, which is typically used when two parallel round wires cannot achieve the required load loss for transformers with usually a rated power of more than 2 MVA. Its disadvantage is that it is not suitable for more transformer designs simultaneously, and manipulation and size of its reel are reduced, so it often needs to be welded for connection during winding. Rectangular wire has defined rounded edges, but for this paper, it is not involved, as dimensions to meet the same cross-section area are not even usually available. The chosen thickness for comparison is the same as for flattened wire, and the width is calculated from the cross-section area. For all types of wire, it is essential to maintain a distance between turns of at least 3% of the wire width during the winding process.

IV. Simulation Results

A. AL Comparison Depending on the Type of Winding

The results of AL from the cylindrical model without the core yoke and the maximum flattening factor 45% for flattened wire and rectangular wire types are shown in Tab. II. Where, for example, the first row is calculated by (1) from the DC loss of 193.0 W and the AC loss of 203.3 W for the foil winding and similarly for the round wire of 292.3 W and 302.3 W. It is evident that the highest relative value of AL is in LV foil winding for all rated powers. Additionally, the type of MV winding does not appear to have any effect on AL in LV winding. Furthermore, it is noteworthy that AL in foil winding are not proportional to the foil dimensions in its cross-section of one turn, as can be seen when comparing MV winding types. For example, the relative AL are higher for the 630 kVA transformer than for the 1000 kVA transformer, despite the fact that the thickness and width of the foil are larger for the 1000 kVA transformer. A possible explanation for this could be that the resulting value of AL in foil winding does not depend on the conductor dimension but on the dimensions of the entire winding, as the radial dimension of the winding is larger for the 630 kVA than for the 1000 kVA transformer.

TABLE II
Results for the Model without Core Yoeke with Different Types of Wire on MV Side

Rated power	Type of MV winding	LV winding	MV winding
250 kVA	Round wire	5.40%	3.43%
	Flattened round wire	5.40%	0.80%
	Rectangular wire	5.40%	0.59%
630 kVA	Round wire	14.97%	7.65%
	Flattened round wire	14.97%	3.71%
	Rectangular wire	14.97%	4.23%
1000 kVA	Round wire	14.43%	9.16%
	Flattened round wire	14.44%	4.52%
	Rectangular wire	14.44%	5.33%

When comparing only MV wire winding types, for all selected rated powers, the flattened round wire has the lowest percentage of AL. This demonstrates why flattening is so advantageous. Using a maximum of 45% flattening of the wire can reduce AL by half or even a quarter for 250 kVA, compared to the round wire. Additionally, the flattened round wire has slightly lower AL compared to rectangular wire, as the value depends mainly on the thickness of the conductor due to the dominant flux density component in the axial direction [3] and rectangular wire has more volume exposed to this component. However, since rectangular wire has a higher filling factor, a more detailed study of this problem would be needed to determine whether it is more beneficial to use flattened or rectangular wire for material savings.

B. AL Comparison Depending on the Chosen Model

Comparing the two models presented in Section III-B reveals that the presence of a core yoke above the winding end significantly reduces AL in the foil winding to one-third of the value without the yoke, as can be observed in Tab. III, which shows the comparison for designs with foil and maximally flattened wire, but similar differences could be provided for other conductor types. On the other hand, AL in the multi-layered MV winding have slightly increased. This is because the AL in foil mainly depend on the radial component of the leakage flux [1], while those in the MV winding are predominantly dependent on the axial component of the leakage flux. The presence of the core yoke near the windings ends straightens the leakage field in the axial direction, resulting in the corresponding AL values.

TABLE III
Results for Both Models with and without Core Yoke for Comparison on Flattened Wire Design

Rated power	Type of model	LV winding	MV winding
250 kVA	Without yoke	5.40%	0.80%
	Yoke inclusive	1.85%	0.86%
630 kVA	Without yoke	14.97%	3.71%
	Yoke inclusive	4.88%	4.23%
1000 kVA	Without yoke	14.44%	4.52%
	Yoke inclusive	5.41%	5.11%

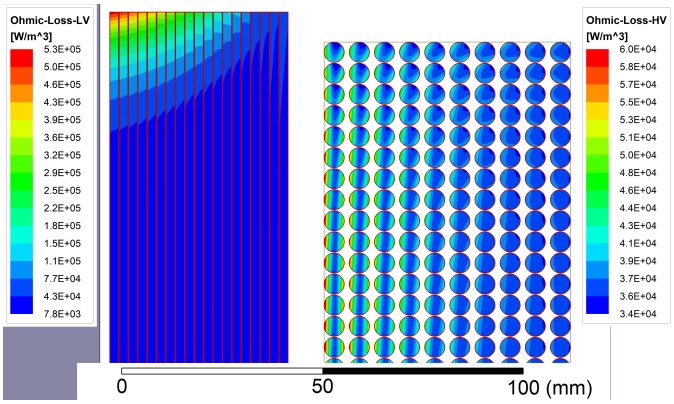


Fig. 4. Loss density distribution in detail of the 1000kVA transformer without yoke (different scale for both windings! - LV on the left and MV on the right).

C. Distribution of Loss Density

A significant difference is evident in the distribution of the loss density between the foil LV and wire MV windings, as shown in Fig. 4. The increased loss density in the HV winding is spread evenly across the entire winding, and decreases radially from the inner to the outer diameter of the winding, due to the trapezoidal course of the leakage flux density in the middle of the winding [3]. Additionally, the flux density's axial component decreases toward the winding's bottom and top ends, meaning that the point with the highest density of AL in the MV winding is in the middle of the winding height, next to its inner diameter. In contrast, almost all AL in LV winding are located at its ends' edges next to the core, and the density is at least 68 times higher than in most areas with the lowest density, so a dangerous hot spot can be created.

D. Verification of Results by Measurements

Verification of the chosen simulation method with measurements is essential. To ensure accuracy, the real flattening factor and wire elongation were used to maintain the same wire cross-section and dimensions as the measured transformer. Unfortunately, there is no available method to divide the load loss between the primary and secondary windings from the measurement, so only the total value of AL can be compared. As indicated in Tab. IV, the simulated relative AL are visibly smaller than the measured values, but increase similarly with the rated power. The difference can be attributed to the remaining additional losses created in the structural steel parts and bus bars described in Section I. Furthermore, the measured load loss is augmented by the temperature evaluation discussed in Section III-C.

There is also a non-negligible difference between the absolute values of measured DC losses and simulated DC losses. This is due to the fact that the simulation does not take into account the resistance of outlets from the LV winding, as well as dimensional or conductivity tolerances, and contact resistance of all connections such as between foil, busbars, copper ropes, bushings or wires, tap changer, etc. However, an important value for comparison is the AL ratio.

TABLE IV
COMPARISON BETWEEN MEASURED AND SIMULATED ABSOLUTE DC LOSSES AND PERCENTAGE OF ADDITIONAL LOSSES

Rated power	Measured DC losses	Simulated DC losses	Measured add. losses	Simulated add. losses
250 kVA	2189.2 W	1956.6 W	4.42%	2.41%
630 kVA	3917.2 W	3673.1 W	12.9%	8.23%
1000 kVA	6422.0 W	6116.1 W	14.6%	8.39%

V. CONCLUSION

This paper examines the AL caused by eddy currents in various types of windings typically used in oil distribution transformers and outlines a chosen model for FEM simulation. To compare the winding types and verify the method, three commonly produced transformers are selected. Simulations are essential for distinguishing AL between windings and other losses, such as stray losses, as there is no feasible way to separate them during measurement or postprocessing.

The results indicate that the highest percentage of AL is created in the LV foil winding, which is multiple times higher than in the MV winding. However, in general, adding a core yoke above the winding end can reduce these AL in foil by three times. Of the different types of conductors used in MV winding, the lowest AL were obtained with flattened round wire. Flattening the wire is important, as it can reduce these losses by more than half compared to the round shape. The rectangular wire has slightly higher AL compared to flattened wire. However, adding a core above the MV winding will slightly increase AL. Ultimately, it has been proven that the chosen method is applicable to the investigated problem.

ACKNOWLEDGMENT

The author acknowledges Ing. Radoslav Cipín, Ph.D. as my research supervisor for his valuable comments and help while writing this paper.

REFERENCES

[1] J. Turowski, "Additional losses in foil-wound and bar-wound transformers," *IEEE Transactions on Power Apparatus and Systems*, vol. 95, no. 3, p. 762, 1976.

[2] B. Ram, "Loss and current distribution in foil windings of transformers," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 145, no. 6, p. 709–716, 1998. [Online]. Available: https://digital-library.theiet.org/content/journals/10.1049/ip-gtd_19982364

[3] S. Kulkarni and S. Khaparde, *Transformer Engineering: Design, Technology, and Diagnostics*, 2nd ed. Boca Raton, Florida, USA: CRC Press, 2013.

[4] M. Krasl, R. Vlk, and J. Grosiar, "Eddy current losses of winding of transformer," in *EUROCON 2005 - The International Conference on "Computer as a Tool"*. Belgrade, Serbia: IEEE, 2005, pp. 1434–1437. [Online]. Available: http://ieeexplore.ieee.org/document/1630232/

[5] L. Jiang, H. Yan, J. Meng, Z. Yin, W. Wei, and Y. Wang, "Analysis of eddy current effect and loss calculation of transformer winding based on finite element algorithm," in *2017 International Conference on Computer Systems, Electronics and Control (ICCSEC)*. Dalian, China: IEEE, 2017, pp. 405–410.

[6] L. Rabins, "Transformer reactance calculations with digital computers," *Transactions of the American Institute of Electrical Engineers, Part I: Communication and Electronics*, vol. 75, no. 3, pp. 261–267, 1956.

[7] *IEC 60076-1: Power transformers - Part 1: General*, 3rd ed., Geneva, Switzerland, 2011.

Advanced Turowski's Method for Eddy Loss Calculation in Transformer Foil Winding

1st Miroslav Mrajca
Department of Power Electrical and Electronic Engineering
Faculty of Electrical Engineering and Communication, Brno University of Technology
Brno, Czech Republic
Miroslav.Mrajca@vut.cz

***Abstract*—This paper presents an improved Turowski's method for calculating additional losses caused by eddy currents in foil winding, which is widely used in oil and dry distribution transformers. These losses are typically significant, and their estimation is essential for transformer output parameters and heating during the design process. The original method is the only one known to provide an analytical solution, which is necessary for design optimization. However, as this paper demonstrates, it is quite imprecise. Therefore, the final equation is derived without simplifications again. Then the new utilization combines Turowski's equations with the finite element method to describe the current density in each winding turn from the magnetic field strength obtained at the winding end, making the simulation faster and still accurate in comparison with the original method. The results are compared with the reference values and discussed.**

***Index Terms*—foil winding, eddy current, additional losses, transformer, current distribution, proximity effect**

I. Introduction

Recently, most distribution transformers, both oil and dry, with high current and low voltage (LV) on at least one side, have been wound with foil or a thin sheet as a conductor. This type of winding is popular due to its simplicity and speed of the winding process, which can be highly automated. At the beginning and end of the foil, the thick bus bars are cold welded to create outlets. The foil, which can be made of copper or aluminium, is wound together with insulating paper impregnated in resin to create inter-turn insulation. After being wound, the shape is fixed by baking, which hardens the resin in the paper, allowing the winding to resist short-circuiting forces even when the shape is oval or rectangular.

The width of the foil winding is typically the same as the entire height of the transformer winding. As the radial flux cuts through the foil sheets, the current is pushed out to the edges, resulting in a significantly nonuniform current density [1]. Unfortunately, this leads to additional losses occurring even at network frequency and creating hot spots at the ends of the winding. This is the biggest disadvantage of foil winding, and as confirms [2] this problem is still very up to date.

LV winding is usually the inner winding and is allied to the outer medium voltage (MV) winding, as presented in Fig. 1(b).

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-23-8430).

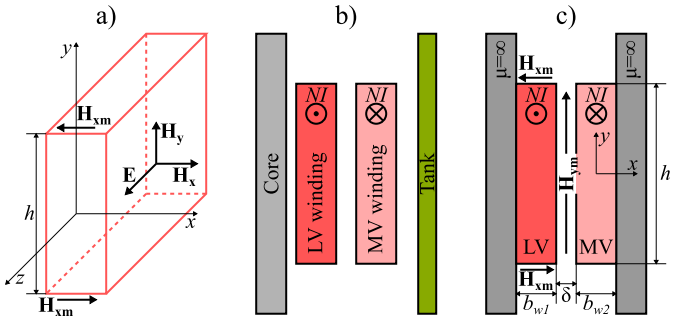


Fig. 1. (a) Illustration of the components of the leakage magnetic field strength and the electric field strength inside the foil sheet. (b) Cross-section of real transformer windings. (c) Model for analytical calculation.

MV winding is typically made up of insulated wire or several sections wound from foil, but its width is a smaller fraction of the total winding height, meaning that the magnetomotive force distribution can be considered substantially uniform. This presents a different kind of problem from the calculation of additional losses that occur in the foil winding.

For the optimization of the transformer design, a fast calculation of the current distribution in the foil winding and thus the resultant additional losses is essential. This can be provided so far only by the analytical method because the finite element method (FEM) is too time-consuming for applicable optimization in production. Until now, there has been only one analytical method, which is described in the following section.

II. Turowski's Original Method

A. Current Density and Power Losses in the Foil Winding

Turowski's original method from [3] solves the current distribution in the foil with infinite length only in the axial direction $J = f(y)$ inside the winding cross-section while disregarding the changes in the current density in the radial direction, and considers it constant as the thickness of the foil is limited to zero. To solve Maxwell's equations for the sinusoidal process inside the conducting plate, it applies:

$$\nabla \times \mathbf{H} = \sigma \mathbf{E} = \mathbf{J}, \tag{1}$$

$$\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H}. \tag{2}$$

Given that **H**, **E** and **J** are vectors of the magnetic field strength of the leakage field, electric field strength, and current density

at their amplitudes, respectively. Furthermore, σ and μ are the conductivity and permeability of the material, respectively, and ω is the angular frequency.

Because the problem of fields is solved only in the XY axes and current flows only in the Z axis, as illustrated in Fig. 1(a), then (1) can be simplified to the form:

$$J = \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}. \tag{3}$$

According to Turowski, the final current density J is composed of two components: J_0, which is uniform and is created solely by the axial magnetic field strength component H_y, and varies only in the radial direction x; and J_1, which is nonuniform and is generated by the radial magnetic field strength component H_x, and varies in the axial direction y, as expressed below:

$$J = J_0 + J_1, \tag{4}$$

$$J_0 = \frac{\partial H_y}{\partial x}, \tag{5}$$

$$J_1 = -\frac{\partial H_x}{\partial y}. \tag{6}$$

When analyzing only H_x then (1) and (2) can be combined and adjusted in the form of a differential equation:

$$\frac{\partial^2 H_x}{\partial y^2} = \alpha^2 H_x, \tag{7}$$

where α can be formulated as:

$$\alpha = (1+j)k, \tag{8}$$

$$k = \sqrt{\frac{\omega\mu\gamma}{2}}. \tag{9}$$

Taking into account that the vertical coordination axis is placed in the middle of the winding height $y = 0$, then taking into account the boundary conditions $H_x(-\frac{h}{2}) = -H_x(\frac{h}{2})$ and $H_x(\frac{h}{2}) = -H_{xm}$, the solution of (7) is given by:

$$H_x(y) = -H_{xm}\frac{\sinh \alpha y}{\sinh \alpha \frac{h}{2}}, \tag{10}$$

where h means the height of the winding and H_{xm} is the radial component of the magnetic field strength at the edge of the foil sheet. Consequently, using (6), the induced component of the current density J_1 can be determined:

$$J_1(y) = H_{xm}\alpha\frac{\cosh \alpha y}{\sinh \alpha \frac{h}{2}}. \tag{11}$$

At the same time, Ampere's law requires that the for uniform current density component J_0 has to apply:

$$2b\int_0^{\frac{h}{2}} J_0 + J_1(y)dy = \sqrt{2}I, \tag{12}$$

with I meaning rated current and b the foil thickness.

In [3] numerous simplifications were considered during further derivation. The main simplification involves the assumption that the uniform component of the current density J_0 in (12) is approximately equal to the DC current density $J_0 \approx \frac{I}{hb}$. Presented power losses are as follows:

$$P_{AC} = \frac{I^2}{hb}\frac{l}{\gamma}\left(1 + \frac{4bH_{xm}}{\sqrt{2}I}\left(1 + \frac{kbhH_{xm}}{2\sqrt{2}I}\right)\right). \tag{13}$$

B. Magnetic Field Strength on the Winding Ends

To ensure the validity of the consideration of dividing current density into two components, uniform J_0 and nonuniform J_1, each dependent only on one direction of magnetic field strength, axial H_y and radial H_x in (4), (5) and (6), H_x must be constant across the entire width of the winding. This means that the radial magnetic field strength component at the winding ends H_{xm} must remain constant along the entire winding end edge, and the current density distribution, as well as additional losses, are equal at every turn.

To determine H_{xm} at the winding edges, Turowski applies Amper's law:

$$\sqrt{2}NI = H_{ym}h + 2H_{xm}b_{w1}, \tag{14}$$

where N is the number of turns, H_{ym} the axial magnetic field strength component on the outer diameter of the LV winding and b_{w1} is the width of the LV winding, as presented in Fig. 1(c). The average value of magnetic field strength in the gap H_{ym} can be calculated in a simplified way as:

$$H_{ym} = \frac{\sqrt{2}NI}{h}k_R, \tag{15}$$

where k_R stands for Rogowski's coefficient defined as:

$$k_R = 1 - \frac{b_{w1} + \delta + b_{w2}}{\pi h}, \tag{16}$$

where δ is the size of the gap between the windings and b_{w2} is the width of the MV winding. The expression of the radial magnetic field strength component at the winding ends H_{xm} from (14), (15) and (16) is then:

$$H_{xm} = \frac{NI}{\sqrt{2}b_{w1}}(1 - k_R) = \frac{NI}{\sqrt{2}b_{w1}}\frac{b_{w1} + \delta + b_{w2}}{\pi h}. \tag{17}$$

III. New Derivation of Turowski Equations

This section introduces the new approach to derive the final equations without multiple simplifications, as included in [3]. The final equations are compared on examples of calculations further in the paper. Adjusting (12) the uniform current density J_0 can be obtained:

$$J_0 = \frac{\sqrt{2}I}{hb} - \frac{2}{h}\int_0^{\frac{h}{2}} J_1 dy = \frac{\sqrt{2}I}{hb} - \frac{2}{h}\int_0^{\frac{h}{2}} H_{xm}\alpha\frac{\cosh \alpha y}{\sinh \alpha \frac{h}{2}}dy, \tag{18}$$

and after integration is the final form:

$$J_0 = \frac{\sqrt{2}I}{hb} - \frac{2H_{xm}}{h}. \tag{19}$$

Setting (19) in (4), the final resulting current density J is:

$$J(y) = \frac{\sqrt{2}I}{hb} - \frac{2H_{xm}}{h} + H_{xm}\alpha\frac{\cosh \alpha y}{\sinh \alpha \frac{h}{2}}. \tag{20}$$

When the current density in the foil winding is described, then the power losses P_{AC} can be expressed as:

$$P_{AC} = \frac{2bl}{\gamma} \int_0^{\frac{h}{2}} J(y) \cdot \overline{J(y)} dy \tag{21}$$

where l means the conductor length and $\overline{J(y)}$ is the conjugate complex number of J. The complete result of (21) after substitution and simplification is as follows:

$$P_{AC} = \frac{bhl}{\gamma} \left(\frac{\left(4H_{xm}^2 - \frac{2I^2}{b^2}\right)(\cos hk - \cosh hk)}{4h^2 \sin\left(hk\left(\frac{1}{2} - \frac{1}{2}j\right)\right) \sinh\left(hk\left(\frac{1}{2} + \frac{1}{2}j\right)\right)} + \frac{2H_{xm}^2 hk(\sin hk + \sinh hk)}{4h^2 \sin\left(hk\left(\frac{1}{2} - \frac{1}{2}j\right)\right) \sinh\left(hk\left(\frac{1}{2} + \frac{1}{2}j\right)\right)} \right). \tag{22}$$

To quantify the degree of nonuniformity of current density in the foil and the eddy losses P_{AC} created thereby, a percentage value of additional losses $P_{AL\%}$ is expressed with respect to ohmic losses P_{DC} as:

$$P_{AL\%} = \frac{P_{AC} - P_{DC}}{P_{DC}} \cdot 100\%. \tag{23}$$

IV. Case Study

A. Selected Transformers for the Comparison

For the case study, three real different three-phase hermetically sealed oil distribution transformers and their electric parameters and mechanical dimensions, listed in Tab. I. are used for the comparison between the additional losses obtained from (13), (22), FEM and the new method using their combination. The material of the conductors is the same in both windings for each design, and aluminium is selected for a wider analysis because of its recent great popularity. The core is built of high-permeability grain-oriented silicon steel H075-23L. The MV windings are shorter compared to the LV windings, and they are mutually axially centred.

TABLE I
Parameters of Analyzed Transformers

Rated power	**100 kVA**	**400 kVA**	**1000 kVA**
Rated voltage ratio		22/0.4 kV	
Short-circuit impedance	4%	4%	6%
Frequency		50 Hz	
Vector group connection		Dyn1	
Core diameter	120 mm	150 mm	190 mm
Straight core part length	143.52 mm	131.1 mm	138.00 mm
Window height	445 mm	685 mm	785 mm
Distance core - winding		2.5 mm	
LV - Foil thickness	0.4 mm	1 mm	2.2 mm
LV - Foil width	410 mm	650 mm	750 mm
LV - Foil insulation		0.15 mm	
LV - Number of turns	49	28	19
Distance between windings		9 mm	
MV - Winding height	395 mm	635 mm	735 mm
MV - Winding width	40.07 mm	50.18 mm	61.44 mm

B. FEM Model

Ansys Electronics has been chosen as the FEM software for simulations of the current density distribution in foil winding. As the most reasonable approach, the current density problem is analyzed in 2D, also used in [4] or [5], and responds with a comparison of the analytical method. For comparison, three models are simulated. In all models, the effect of other steel parts, such as the tightening construction and tank walls, is omitted.

The first model consists only of windings cross-sections and core legs surrounding both windings, as well as in an analytical model as shown in Fig. 1(c), but in the real transformer, the core near the MV winding is usually so far as to be neglected. For this case, both windings have the same height to eliminate enlargement of the radial leakage field component due to the height difference [6] and the Cartesian coordinate system is chosen to simulate the model as closely as possible to the analytic model. The second model is more similar to a three-phase transformer, but only in a part where the windings stick out of the core window, which is the majority case. The model consists of only the windings cross-sections and one coaxial core leg. As this part is usually cylindrical, the coordinate system is also cylindrical around the Z axis. The last model represents the winding part inside the core window of the three-phase transformer, so the model is similar to the second model but with the addition of a core yoke near the winding ends. Therefore, a Cartesian coordinate system is chosen for this part, which is usually straight. The last two models should demonstrate the difference of the core yoke near windings ends on the current density, and thus additional losses in the foil. All the described models can be seen in Fig. 2.

Models use as much symmetry as possible, so only one-half of the height of one phase is simulated with the Neumann boundary condition and one-half of the width with the Dirichlet boundary condition on the core leg bisector. The remaining boundaries are set as a balloon, and the rest of the volume is assigned to the vacuum properties. The simulation setup is set to solve eddy currents, since only sinusoidal current flow is analyzed. Because the simulation should represent a short-circuit test, the core lamination properties are assigned with a linearized BH curve from 0 to 1 T and the conductivity of the conductors is at the reference temperature for oil transformers,

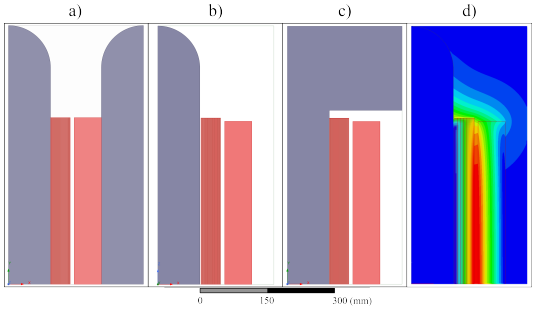


Fig. 2. (a) Analytical model. (b) Winding part sticking out of the core window. (c) Winding part inside the core window. (d) Resultant magnetic field

TABLE II
COMPARISON OF ADDITIONAL LOSSES IN FOIL WINDING OBTAINED BY ANALYTICAL AND FEM CALCULATIONS

Rated power	DC losses	Turowski simplified	Turowski unsimplified	FEM Turowski	FEM only core	FEM with yoke
100 kVA	103.9 W	11.47%	2.67%	2.77%	3.88%	1.39%
400 kVA	295.2 W	11.53%	3.57%	4.02%	5.68%	1.95%
1000 kVA	631.5 W	15.08%	5.70%	10.12 %	13.37%	5.47%

which is 75 °C in the windings. Each turn of foil is represented by a solid rectangle defined by the real width and thickness of the foil. The number of turns is determined by the selected designs, and each turn is excited with the amplitude of the rated current on the LV side. The initial mesh within the foil is set with an element length of no more than half of its thickness. In contrast, MV winding is formed by a single stranded rectangle and is excited with a current to produce the same magnetomotive force as LV winding. Furthermore, the configuration settings utilize an adaptive mesh with a maximum of 10 passes or an error less than 0.001%.

V. RESULTS FOR COMPARISON

Tab. II provides a comparison of five percentage values of the additional losses for each design calculated according to (23) related to DC losses listed in the second column. The values in the first two columns are obtained from the original simplified analytical equation by Turowski (13) and the newly derived analytical equation without any simplifications (22), both using a radial component of the magnetic field strength from (17). The remaining three columns are from the FEM simulations, described in Section IV-B, which portray Turowski's arrangement, part of the real windings outside the core window, and the last part of the real winding inside the window near the core yoke. The last two values can be used as references for comparison.

It is clear that there is a large discrepancy between the analytical values and, for transformers with rated power below 1000 kVA, the additional losses obtained from the simplified analytical equation (13) are very overrated owing to the reference values; only for the 1000 kVA transformer, the additional losses are comparable to the FEM model with the only core leg, but this similarity can be considered more of a coincidence. So far, the unsimplified equation (22) offers results that are between the results for the model showing analytical calculation and the FEM model with a core yoke for all designs. Based on these findings, the original simplified analytical equation (13) is not suitable for the calculation of additional losses in foil winding. Equation (22) can be used to calculate the approximate additional losses in the winding only inside the core window. An explanation for this statement can be found in deriving the magnetic field strength in (17), which is based on Rogowski's coefficient, which is determined especially for the magnetic field inside the core window.

However, most of the winding is not affected by the core yoke, and because the results for the winding outside and inside the core window, written in the last two columns of Tab. II, are very different, and losses outside the core window are almost three times higher, and also, because this part of the winding forms a larger part, calculated losses by (22) would be underestimated. Thus, this analytical method needs to undergo a deeper review to be used in practice.

VI. NEW UTILIZATION OF TUROWSKI'S METHOD

There are more reasons why Turowski's approach is inaccurate. The first has already mentioned above that the boundary conditions, steel parts near windings, are not included, although this factor significantly affects the resulting losses. Another reason is that the nonuniform current density is created only by the radial component of the magnetic field strength, but (3) has another solution, and the nonuniform current density depends on both the radial and axial components of the magnetic field strength. Thus, the same statement applies to the uniform current density component. Due to this, the radial component of magnetic field strength is not equal in each turn and gradually is getting stronger from the outer winding diameter towards the core leg. Thus, the largest nonuniformity and additional losses can be found in the very first turn of foil next to the core leg. The real divide of current density across the half height of the winding due to the magnetic field strength components H_x and H_y in the middle of the foil thickness of the first turn as shown in Fig. 3 and in the last turn of the 100 kVA transformer outside the core window in Fig. 4. Noticeably, the dependence of the axial current density on the axial magnetic field H_y increases with every other turn.

However, the course of the simulated current density distribution along the foil height is very close to the curve shape, displayed in [7], obtained from Turowski's equation or (20). Therefore, an approximation of the simulated current density in the middle of the foil thickness of each turn is offered from the obtained magnitude of the simulated current density J_m at the foil's edge, where is the maximum of density. This current density J_m value can also be calculated from the magnetic field strength known from (3). Then, for an approximate calculation using (20), the equivalent radial magnetic field strength component $H_{xm,eq}$, which would cause the simulated magnitude of the current density at the end of the winding J_m, needs to be expressed from (20) with the substitution of $y = \frac{h}{2}$ and adjustment for magnitude as:

$$J_m = \left| \frac{\sqrt{2}I}{hb} + H_{xm,eq} \left(\alpha \coth\left(\alpha \frac{h}{2}\right) - \frac{2}{h} \right) \right|, \qquad (24)$$

and because condition $h >> \frac{2}{\alpha}$ is satisfied, the equivalent

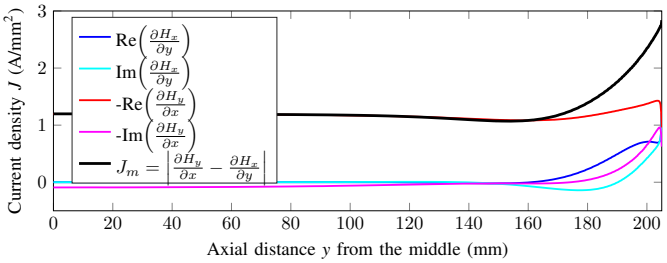


Fig. 3. Magnitude of the current density and its components in the first turn of a 100kVA transformer outside the core window obtained by FEM.

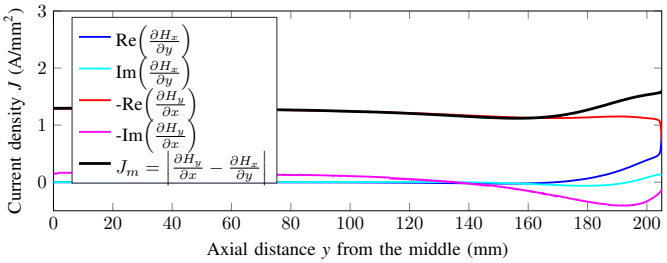


Fig. 4. Magnitude of current density and its components in the last turn of a 100kVA transformer outside the core window obtained by FEM.

magnetic field strength $H_{xm,eq}$ can be derived as:

$$H_{xm,eq} = h\frac{\frac{\sqrt{2}I}{hb}(2-hk) - \sqrt{J_m^2(2h^2k^2-4hk+4) - \frac{2I^2}{b^2}k^2}}{2h^2k^2 - 4hk + 4}, \tag{25}$$

where J_m is the (maximum) current density in the middle of the foil's edge and can be obtained from FEM for each turn.

An approximation of the current density along the half height of the 1000 kVA foil winding for all turns in the middle of the foil thickness using (20) and (25) is shown in Fig. 5. Note that the approximation is very accurate for the first half of the turns and then the fitting accuracy deteriorates. The accuracy of the approximation is evaluated by additional losses comparison, presented in Tab. III, where the results are related to the same DC losses as in Tab. II. The third column contains additional simulated losses that occur in the middle of the foil thickness, where the effect of the foil thickness on losses is negligible; therefore, it is more appropriate for comparison. The fourth column of the report showcases the results of the semi-analytical calculation. This method involved using a combination of FEM to obtain the current density at the ends of the windings, along with equations (25) and (22) for the remaining calculations. These semi-analytical values are lower than the simulated values but still very convenient.

VII. Conclusion

This paper compares additional losses in transformer foil winding calculated by the original Turowski's analytical equation, the new derivated equation, and FEM simulations of three different models. The results indicate that the original Turowski's equation is too simplified and inaccurate for any design. The new equation can be used to calculate the part of the winding inside the core window, but it is not accurate for the outer parts of the winding. Nevertheless, the new derivated equation can be used for the sufficient approximation of the current density inside each foil turn from the known magnetic field strength on the winding end.

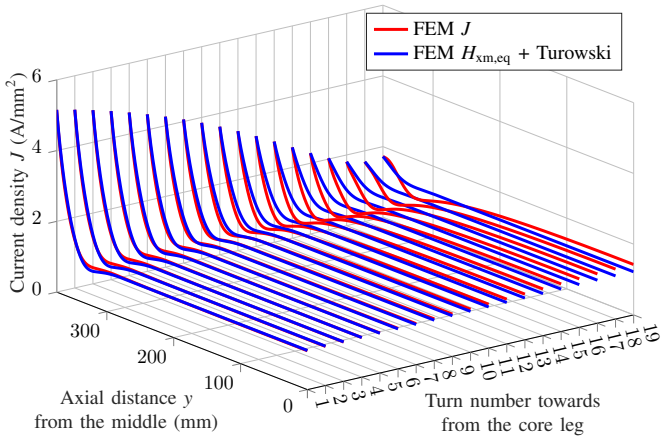


Fig. 5. Magnitude of current density at half height of the 1000 kVA foil winding.

TABLE III
Comparison of Additional Losses Obtained for the Designs Outside Core Window without Yoke

Rated power	FEM losses	FEM only axial losses	J_m from FEM and $H_{xm,eq}$ + new Turowski
100 kVA	3.88%	3.87%	2.44%
400 kVA	5.68%	5.41%	4.03%
1000 kVA	13.37%	12.14%	10.78%

Acknowledgment

The author acknowledges doc. Ing. Radoslav Cipín, Ph.D. as my research supervisor for his valuable comments and help while writing this paper.

References

[1] N. Mullineux, J. Reed, and I. Whyte, "Current distribution in sheet- and foil-wound transformers," *Proceedings of the Institution of Electrical Engineers*, vol. 116, no. 1, pp. 127–129, 1969.

[2] T. B. Gradinger and M. Mogorovic, "Foil-winding design for medium-frequency medium-voltage transformers," in *2021 23rd European Conference on Power Electronics and Applications (EPE'21 ECCE Europe)*, 2021, pp. 1–10.

[3] J. Turowski, "Additional losses in foil-wound and bar-wound transformers," *IEEE Transactions on Power Apparatus and Systems*, vol. 95, no. 3, p. 762, 1976.

[4] L. Jiang, H. Yan, J. Meng, Z. Yin, W. Wei, and Y. Wang, "Analysis of eddy current effect and loss calculation of transformer winding based on finite element algorithm," in *2017 International Conference on Computer Systems, Electronics and Control (ICCSEC)*. Dalian, China: IEEE, 2017, pp. 405–410.

[5] T. Wang, W. Yuan, and J. Yuan, "A novel semi-analytical method for foil winding losses calculation considering edge effect in medium frequency transformers," *IEEE Transactions on Magnetics*, vol. 58, no. 5, pp. 1–9, 2022.

[6] W. Knaack and H. Schwaab, "Additional inductance in transformer," *Archiv für Elektrotechnik*, vol. 32, no. 7, pp. 470–482, 1938.

[7] J. Turowski and M. Turowski, *Engineering Electrodynamics: Electric Machine, Transformer, and Power Equipment Design*, 1st ed. Lodz University of Technology, Poland: CRC Press, 2017.

Control principles of a three-phase active rectifier and a three-phase DC/AC converter

Tomáš Macík
Department of Power Electrical and Electronic Engineering
Brno University of Technology
Brno, Czech Republic
Email: tomas.macik@vut.cz

***Abstract* — This article lists control possibilities of three-phase active rectifier and three-phase DC/AC converter. Particularly, control in dq frame for active rectifier and a full-state feedback control for DC/AC converter is described in detail.**

Keywords — active rectifier, DC/AC converter, dq frame, full-state feedback control

I. INTRODUCTION

High requirements are placed on electromagnetic compatibility and the ratio of higher harmonic components in the power grid nowadays. From the point of view of the power grid, the ideal load is a resistive load because the drawn phase currents are in phase with the voltage and are sinusoidal. This desired effect can also be achieved by using an active rectifier. A three-phase active rectifier is able to transfer energy in both directions, i.e. from the grid to the load and vice versa if the load is active. This happens under the condition that the input phase currents have a sinusoidal shape. In addition, it is possible to set the phase shift between the input current and the phase voltage and thereby regulate the reactive power transferred to/drawn from the power grid.

The article describes the control principles of a three-phase active rectifier and a three phase DC/AC converter, which is connected behind the active rectifier. The described control was verified on a discrete (sampled) model created in Matlab Simulink. The three phase DC/AC converter also has its central point (neutral wire) conducted out and thus represents a power microgrid with various options for setting the output voltage parameters. At the same time, the entire system (active rectifier + DC/AC converter) must draw sinusoidal phase currents from the power grid in phase with the grid voltages. This requirement is ensured by an active rectifier with the possibility of energy recovery.

II. CONTROL OF THE THREE-PHASE ACTIVE RECTIFIER

The three-phase power supply network is fed through three input phase choke coils, the presence of which is necessary. The output DC voltage V_{DC} must be greater than the amplitude of the network line-to-line voltage because the power part together with the choke coils works in the STEP-UP mode. Transistors are switched in sinusoidal PWM mode (control signals s_u, s_v, s_w control the upper transistors and their negation controls the lower transistors of the branches). [2]

The described active rectifier can be controlled in a *dq* frame or in a mixed *dq-UVW* frame. Both control options use a phase-locked loop, which serves to determine the phase angle of the U phase of the network. This article will describe the control in a *dq* frame using PI controllers. In comparison, control in a mixed *dq-UVW* would require the use of PR (parallel-resonant) controllers. In this type of controllers, the integrator is designed to have infinite gain at the selected frequency, here 50 Hz. In contrast, in standard PI controllers, the integrator has infinite gain at zero frequency, i.e. for a DC signal. In general, PR regulators work fine for the steady state but they are not suitable for the transient states because of the fade of resonances. [1]

The topology of the power part together with the control structure of the three-phase active rectifier in the *dq* frame is shown in Fig. 1. The current control block is separately shown in Fig. 2.

The topology of the active rectifier can be supplemented with fuses at the input, a power button, starting resistors and a grid LC filter. A simulation with these added components was performed in Matlab Simulink. In the following text, the active rectifier will be also referred to as AR. [1]

A. Parameters of the active rectifier

The network is three-phase, RMS value of line-to-line voltage 400 V, frequency 50 Hz. Required AR parameters: rated power P_n = 12 kW, DC link voltage V_{DC} = 900 V, pulse-width modulation frequency f_{PWM} = 25 kHz, inductance of the input phase choke coil L_{ph} = 1,3 mH, capacity of DC link capacitor C_{DC} = 2,94 mF. The starting resistance R_{start} = 10 Ω is placed in each phase at the input to limit the current surge when charging the DC link capacitor. If the DC link voltage reaches 500 V, these resistors are bridged by the contactor. The current will start to be drawn only when the regulator ensures the DC link voltage of 900 V.

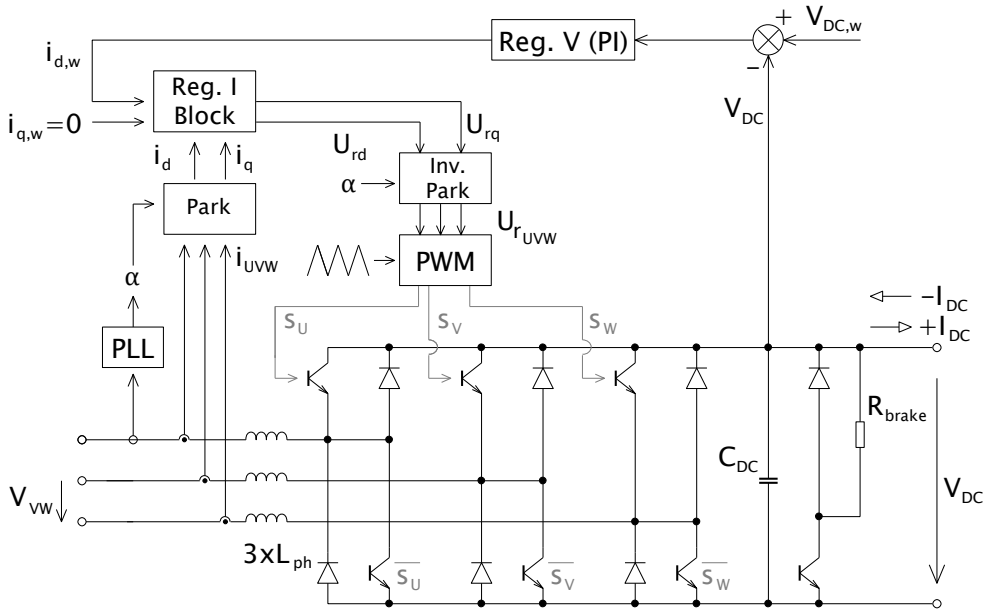


Fig. 1. Illustration of the control structure of a three-phase active rectifier in a *dq* frame and its topology (taken from [3] and modified)

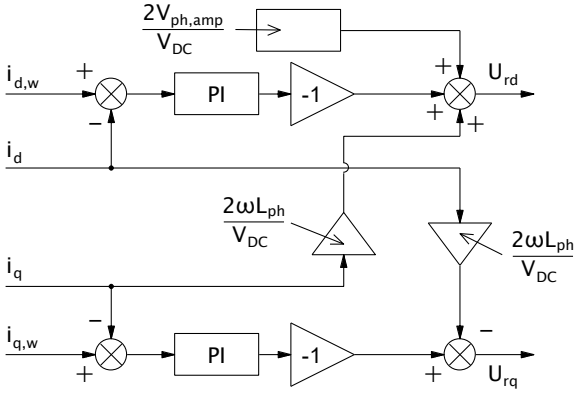


Fig. 2. Current control block (taken from [3] and modified)

B. Control in dq frame

The alternating input phase currents are transformed into the *dq* frame by the Park transformation, using the angle provided by the phase-locked loop. After converting to a *dq* frame, these alternating currents become direct currents, which allows the use of PI controllers. The output control signals from the current regulators are subsequently transferred from the *dq* frame to the three-phase *U,V,W* frame by inverse Park transformation. These three sinusoidal control signals enter the PWM modulator block where they are compared with the high-frequency triangular carrier signal. This creates six PWM control signals that control the switching of the six AR transistors. An amplitude-invariant transformation will be used. The current i_d is numerically equal to the amplitude of the active current and therefore controls the active power. The current i_q is numerically equal to the amplitude of the reactive current and therefore controls the reactive power transferred to/from the network. [1] [3] [6]

The regulators of currents i_d and i_q are identical and were designed by the optimal module method. They are of the PI type with the magnitude of the integrator gain 2,778 and the proportional component gain 0,0361. The output of the current regulators are the control signals in the *d* and *q* axes. The allowed range of their values is from -1 to 1. Therefore, in the settings of the PI controller, it is necessary to set the saturation limit to ±1. A linear pre-correction is also added to the output of the i_d current regulator, which represents the steady value of the control signal U_{rd}. Linear pre-correction reduces the current pulse when charging the capacitor at the beginning of regulation. When using pre-correction, signals U_{rd} or U_{rq} could obtain a value greater than 1. Therefore, it is advisable to limit them, taking into account the physical limitations of the converter (the maximum possible voltage is equal to the DC bus voltage). The modulus of the vector formed by the components U_{rd} and U_{rq} must not exceed the size of 1. [1]

The regulator of the DC link voltage V_{DC} was designed by the symmetrical optimum method, it is of the PI type with the size of the integrator gain 10461 and the proportional component gain 7,532. The output of the voltage regulator is the desired value of the current affecting the active power (current $i_{d,w}$). It is advisable to limit the size of this current. Therefore, the output limit values in the PI controller are set to ± 30. Current values of ± 30 A are related to the power of the converter, which is 12 kW. If the desired value of the current i_q is equal to zero, the reactive power is zero and the rectifier draws phase currents in phase with the network voltage (the power factor is 1). This mode is in demand and will be considered further. The desired value of the current i_q can generally be selected. [1]

III. CONTROL OF THE DC/AC CONVERTER

The control structures will ensure the generation of three control signals for branches A, B, C based on the required waveforms of phase voltages at the output of the DC/AC converter. From these three control signals, 6 control signals are generated for the individual transistors of branches A, B, C based on the principle of unipolar control in the PWM modulator

block. The control of the branch for the neutral wire is based on the principle of cascade regulation. The control structures of the DC/AC converter itself can be based on the principle of full-state feedback control or cascade regulation. This article will describe full-state feedback. The description of the cascade regulation of the DC/AC converter can be found in [1].

The three-phase DC/AC converter is preceded by an active rectifier. DC voltage source V_{DC} in Fig. 3. represents a DC bus of the voltage type formed by a capacitor C_{DC} with a braking resistor connected in parallel to absorb the supplied energy from the active load. Fig. 3. shows the connection diagram of the power part of the three-phase DC/AC converter with neutral wire together with the structure of the full-state feedback control.

Two output LC filters are connected to the terminals marked A, B, C, N. These filters filter the pulses from the sinusoidal PWM so that the voltages of the harmonic waveform are present at their output. In the following text, the DC/AC converter will be also referred to as DAC.

A. Parameters of the DC/AC converter

The ability to set the output phase voltage parameters for each phase separately is required. The DAC should allow setting the frequency, amplitude, phase shift and shift relative to the horizontal axis for the phase voltage of each phase separately. The adjustable frequency range is required from 45 to 55 Hz and the adjustable amplitude range from 0 to 424 V, which corresponds to the RMS value range of 0 to 300 V. The voltage of the DC bus V_{DC} = 900 V. The DAC should be able to continuously transmit a power of 12 kW (4 kW per phase) in both directions and must be able to transmit a power of 75 kW (25 kW per phase) supplied by the active load, while this power will be wasted on the braking resistor in the DC bus circuit. When supplying power of 75 kW from the active load, there is an increase in the voltage of DC bus capacitor. If the voltage exceeds 1000 V, the transistor in the branch with the braking resistor is switched on and connects the resistor in parallel to the capacitor of the DC bus. The capacitor starts to discharge. When its voltage drops below 950 V, the transistor is turned off and disconnects the braking resistor. Consumption of 75 kW is not allowed because the transistors of the active rectifier are not designed for this power in a real application. When the power of 75 kW is drawn from the output of the DAC, the voltage on the DC link capacitor starts to drop. If it drops below 850 V, the desired value of the output voltages changes to zero. If it returns above 899 V, the desired value of the output phase voltages is also restored.

Other parameters of the DAC, including circuit elements: frequency of pulse-width modulation $f_{PWM,DAC}$ = 20 kHz, braking resistance R_{brake} = 10 Ω, inductance of the filter choke coil for the N branch L_N = 112,5 μH, capacity of the filter capacitor for the N branch C_N = 100 μF, choke coil inductance of the first LC filter $L_{filt1,DAC}$ = 112,5 μH, capacitor capacity of the first LC filter $C_{filt1,DAC}$ = 60 μF, choke coil inductance of the second LC filter $L_{filt2,DAC}$ = 50 μH, capacitor capacity of the second LC filter $C_{filt2,DAC}$ = 25 μF. The resistance of the load resistor for the consumption of 4 kW at 230 V $R_{load,DAC}$ = 13,23 Ω and for the consumption of additional 21 kW at 230 V $R_{load,DAC,max}$ = 2,52 Ω.

B. Control of the neutral wire branch

The control structure is based on cascade regulation with a superior voltage loop and a subordinate current loop. In the subordinate current loop, the current through the choke coil L_N is regulated, and in the superior voltage loop, the voltage difference on the upper and lower capacitor C_N is regulated. The goal is to ensure the same voltage on both capacitors, therefore the desired value of the voltage entering the regulation block is zero.

The current regulator for N branch was designed by the optimal module method, it is of the PI type, while the gain of the I component has a value of 0,2222 and the gain of the P component has a value of 0,0025. The output of the current regulator is the control signal U_{rN}, which is compared with a triangular carrier signal with a peak-peak value from -1 to 1 in

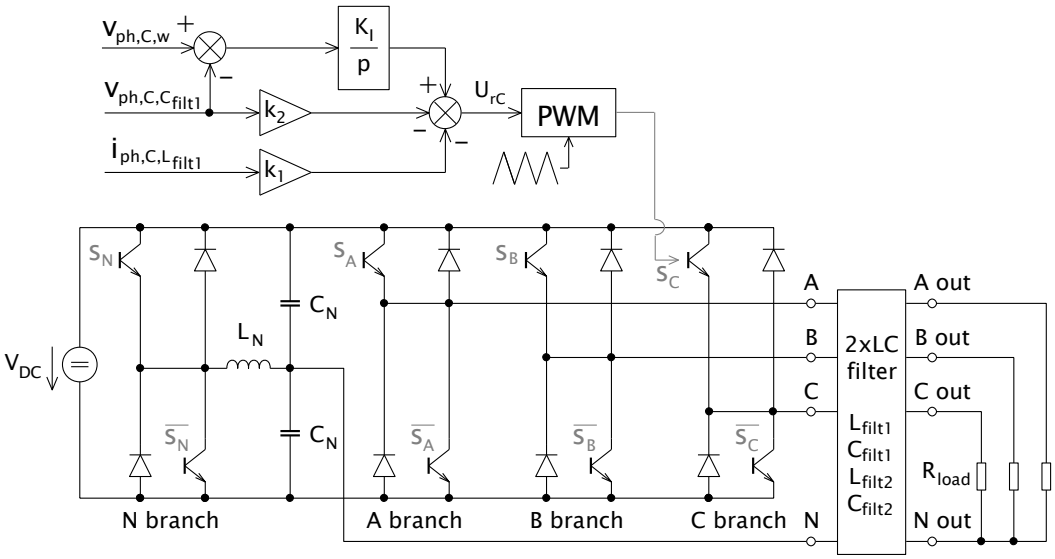


Fig. 3. Depiction of the full-state feedback control structure of a three-phase DC/AC converter with a neutral wire and its power part

the PWM modulator block. Therefore, the range of the control signal U_{rN} must also be in the range from -1 to 1. Because of that, the output limit values in the PI controller are set to ±1. By this comparison, two signals controlling the switching of the upper and lower transistor of the branch for wire N are generated in the PWM modulator block based on the principle of unipolar control. [1]

The voltage regulator for N branch was designed by the symmetrical optimum method, it is of the PI type, while the gain of the I component has a value of 1111 and the gain of the P component has a value of 0,6667. The output of the voltage regulator is the desired current through the choke coil L_N. It is advisable to limit the size of this current. Therefore, the output limit values in the PI controller block are set to ± 200. The current values of ± 200 A are related to the designing of the transistors of the N branch in a real application. [1]

C. Full-state feedback of DC/AC converter

The state model of the DAC will represent one of its phases in a simplified connection. At 50 Hz, the choke coil impedance of the second LC filter has a negligibly small value compared to the impedance of the output capacitor of this LC filter. Therefore, we proceed to the simplification and neglect this choke coil in the calculation, which reduces the order of the system from four to two (the capacitors of both LC filters are now parallel and can be replaced by one). After deduction and adjustment to matrix form, we obtain: [1] [7]

$$\boldsymbol{x} = \begin{pmatrix} i_A \\ v_A \end{pmatrix}; \tag{1}$$

$$\boldsymbol{A} = \begin{pmatrix} -\dfrac{R_{filt}}{L_{filt}} & -\dfrac{1}{L_{filt}} \\ \dfrac{1}{C_{filt} + C_{filt,2}} & 0 \end{pmatrix}; \boldsymbol{b} = \begin{pmatrix} \dfrac{V_{DC}}{2L_{filt}} \\ 0 \end{pmatrix}; \tag{2}$$

$$\boldsymbol{c}^T = (0 \quad 1);\ d = 0;\ u(t) = U_{rA}. \tag{3}$$

The described system is controllable and observable. Verification is described in more detail in [1]. The next step for the full-state feedback control design is the choice of new poles of the controlled system. There are several methods by which the desired poles of the controlled system can be selected, e.g.: selection of multiple real poles, selection of poles based on the **ITAE criterion** and selection of poles via LQR. From the listed options, in the following text the ITAE criterion will be used for the selection of new poles of the controlled system. **Integrative state control** is used, i.e. state control with an added loop with an integrator to eliminate the influence of disturbance variables. The disturbance variable in the case of the DAC is the load current. [1]

The ITAE criterion prescribes the shape of the characteristic polynomial of the controlled system. The coefficients of the required characteristic polynomial of the closed control system N_{kw} are determined as follows based on the ITAE criterion for the second-order system:

$$N_{kw} = p^2 + 1{,}505ap + a^2, \tag{4}$$

where a is a constant expressing the time scale. By adjusting it, the standard shape of the feedback characteristic polynomial according to ITAE is adapted to correspond to the real system. [4]

Practically, it is about setting the required transient characteristic of the controlled system so that it settles at a value equal to the frequency response to the DC signal for the original system. In this case, it is a value equal to the gain of the converter, that is $K_{DAC} = V_{DC}/2 = 450$. A suitable setting is for $a = 10\,226$. Then the coefficients for individual powers of p have the value $a_1^w = 15390, a_0^w = 104{,}6 \cdot 10^6$, what will be used in the calculation of feedback vector $\boldsymbol{k}^T$ with the help of **Ackermann relation**. After that, we obtain numerical expression of the feedback vector $\boldsymbol{k}^T = (k_1 \quad k_2) = (3{,}825 \cdot 10^{-3} \quad -4{,}066 \cdot 10^{-8})$. The poles according to the ITAE criterion are $p_{1,2}^w = -7695 \pm 6735j$. To compare, the poles of the original system without full-state feedback control were $p_{1,2} = -44 \pm 10226j$. [4] [1]

To complete the design of the integrative state control, it is necessary to determine the weight of the integration component K_I. The K_I parameter can be selected using the root hodograph method, which can be done through the *rltool* in Matlab. The selection of the parameter K_I takes place by simulating the open loop, while the parameter K_I is chosen so large that the phase safety of the open loop approaches 40° from above. For phase safety of 40,2°, it is necessary to set K_I to the value of 13,5. [3], [5]

The output control signals $U_{rA,B,C}$ from the state controller should also be limited to the range from -1 to 1 in order to compare them with the triangular carrier signal with the range of -1 to 1 in the PWM modulator.

IV. WAVEFORMS DEPICTION FOR ACTIVE RECTIFIER AND DC/AC CONVERTER

The waveforms in Fig. 4. can be divided into several parts. In the time from 0 s to approximately 0,16 s, the DC link capacitor is charged through the starting resistors. At the time of 0,16 s, the starting resistors are bridged. The phasing of the phase-locked loop takes place within 0,3 s. From 0,3 s to 0,4 s, the DC bus voltage is regulated to the desired value of 900 V. From 0,4 s to 0,6 s, the power of 12 kW is supplied to the DAC from active load, while this power increases to its nominal value from zero following the ramp in 20 ms (12 kW does not apply when the DAC voltage is reduced between 0,5 s and 0,575 s). In the time from 501 ms to 550 ms, the amplitude of the output phase voltages is reduced from the original 100% to 10%. The amplitude decreases from 100% to 10% in 1 ms (between 500 ms and 501 ms) and increases back to its original 100% value between 550 ms and 575 ms. The supply of power from the active load can also be seen in the waveform of grid currents to the active rectifier. Currents are in antiphase with voltages. The power factor is equal to minus one. Between 0,6 s and 0,7 s, the power of 75 kW is supplied to the DAC from the active load. However, when the DC link voltage exceeds 1000 V, this power is dissipated on the braking resistor, so the input phase currents do not exceed the permitted value. By connecting the braking resistor in parallel to the DC link capacitor, the DC link voltage also decreases towards the desired value of 900 V. When the voltage drops below 950 V, the braking resistor is disconnected. In the time from 0,7 s to 0,8 s, 12 kW of power is drawn by a

resistive load. The input power grid currents to the active rectifier are in phase with the voltages, the power factor is equal to one. From 0,8 s to 0,9 s, 75 kW of power is drawn to a resistive load. However, such a high consumption is prevented by reducing the output voltage of the DAC to zero so the input phase currents do not exceed the permitted value of 30 A in amplitude (given by limiting the currents i_d or i_q in the DC link voltage regulator in the AR part). The currents are also in phase with the network voltages. The current i_U represents the current through the input phase choke coil L_{ph} of the first phase of the network. Voltage u_U expresses the voltage of the first phase of the network. The current i_A represents the current from the output terminal *A out* of the DAC. Voltage v_A expresses the output voltage of phase A of the DAC between terminals *A out* and *N out*.

V. CONCLUSION

The added regulation structure ensures that the phase currents and DC link voltage are kept within safe limits. The requirements for the active rectifier and the three-phase DC/AC converter have been met, as evidenced by the waveforms shown in chapter IV.

The disadvantage of full-state feedback control is the impossibility of setting current limitation. Therefore, the most suitable control structure would be a parallel control structure combining a state control and a parallel current controller. If the permitted value of the choke coil current is exceeded, the parallel current regulator would take over the control and limit the current. Here, however, it would be necessary to dynamically limit the output of the integrator in the full-state control when the state control is not active. In described DC/AC converter there was no excessive increase in currents through the choke coils when the full-state feedback control was used.

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-23-8430). Author acknowledges doc. Ing. Pavel Vorel, Ph.D. as my research supervisor and Ing. Ivo Pazdera, Ph.D. for their valuable comments and help while writing this paper.

REFERENCES

[1] T. Macík, Model of the three-phase power microgrid. Brno, 2020. Also available from: https://www.vutbr.cz/studenti/zav-prace/detail/127373. Diploma thesis. Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Power Electronics and Engineering. Supervisor Ivo Pazdera.

[2] M. Patočka, Selected articles on power electronics, vol. 2. Scriptum FEKT, BUT Brno: Pulse converters without a transformer. Brno, 2005.

[3] S. Bacha, I. Munteanu, and A. I. Bratcu, Power electronic converters modeling and control: with case studies. London: Springer, 2014. Advanced textbooks in control and signal processing. ISBN 978-1-4471-5477-8.

[4] M. Vítečková and A. Víteček, Full-state feedback control. Ostrava, 2016. ISBN 978-80-248-3900-4.

[5] P. Blaha and P. Vavřín, Control and regulation I. Scriptum FEKT, BUT Brno.

[6] [online]. [cit. 2023-03-02]. Available from: https://www.mathworks.com/help/physmod/sps/powersys/ref/abctodq0dq0toabc.html

[7] J. Skalický, Full-state feedback control of electric drives. Brno, 2009.

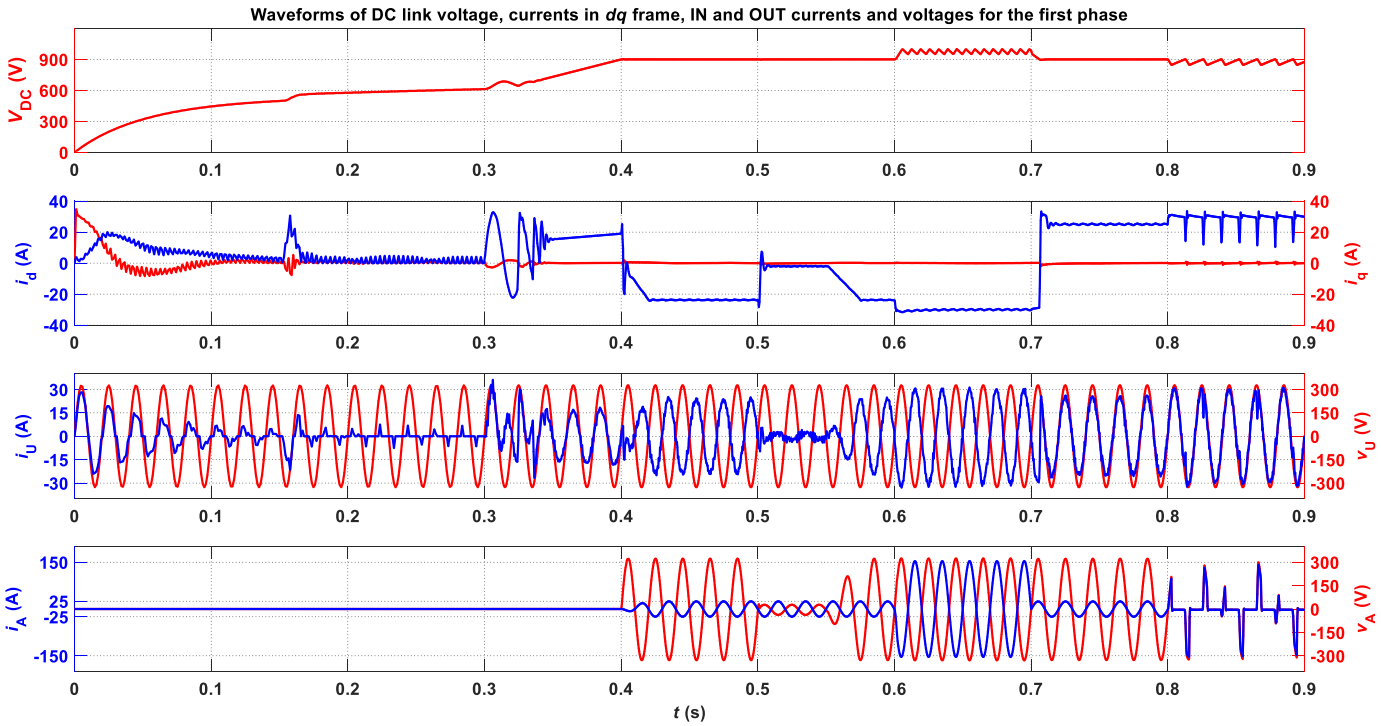


Fig. 4. Waveforms of DC link voltage, currents in *dq* frame, IN and OUT currents and voltages for the first phase

Modeling Li-ion battery pack for stationary energy storage systems

Marek Sedlařík
UETE
BUT FEEC
Brno, Czech republic
xsedla1a@vutbr.cz

Petr Vyroubal
UETE
BUT FEEC
Brno, Czech republic
vyroubal@vut.cz

***Abstract*— The use of renewable energy sources such as solar and wind power has been increasing, but their intermittent nature presents a challenge for energy storage systems. Stationary energy storage systems with Li-ion batteries are one of the proposed solution, but temperature can significantly affect their performance and lifespan. Therefore, it is necessary to understand the thermal behavior of Li-ion batteries in different situations. This paper deals with the thermal simulation of Li-ion Battery pack using computational fluid dynamic (CFD) analysis in ANSYS Fluent software.**

Keywords—Li-ion batteries, thermal simulation, ANSYS Fluent, MSMD model

I. Introduction

The European Union has set a target to achieve 40% renewable energy coverage in the overall energy mix by 2030 and therefore there is an increase in the construction of renewable energy sources such as solar and wind power. However, one of the major challenges of using these sources is their intermittent nature. This creates a problem of how to store the energy generated when it is not being used. One solution to this problem is stationary energy storage with Lithium-ion (Li-ion) batteries.

Li-ion battery is used for all kind of accumulators based on Li-ion intercalation. A Li-ion battery is composed of three main components: the anode, the cathode, and the electrolyte. The anode is typically made of graphite, while the cathode is made of a lithium metal oxide. The electrolyte is a lithium salt dissolved in an organic solvent. The specific chemistry of the cathode material plays a crucial role in determining the performance of a lithium-ion battery. Different cathode materials have different voltage profiles, energy densities, and cycle lives. Some common cathode materials include lithium cobalt oxide (LCO), lithium manganese oxide (LMO), lithium iron phosphate (LFP), and lithium nickel manganese cobalt oxide (NMC). [1, 2]

Li-ion batteries have become the dominant technology for stationary energy storage due to their high energy density, long cycle life, and low self-discharge rate. However, the performance and lifespan of Li-ion batteries can be significantly affected by temperature, which can rise during the charging and discharging cycles. In order to optimize the performance and ensure the lifespan of the system, it is essential to understand the thermal behavior of the storage device.

Temperature simulations are an important tool for understanding how temperature affects battery performance and how to design optimal storage systems. As renewable energy sources continue to grow in popularity, stationary energy storage systems and thermal simulations will become increasingly important for ensuring a stable and sustainable energy future.

These simulations can also be used to evaluate the performance of Li-ion batteries in different operating environments. For example, in hot and humid environments, Li-ion batteries can experience reduced performance and capacity due to accelerated aging. High temperatures can cause thermal runaway, a condition in which the battery generates excessive heat and can potentially explode or catch fire.

II. Numerical Modeling

ANSYS Fluent is a Computational Fluid Dynamics software using a finite volume method to solve partial differential equations. This software can simulate the complex fluid dynamics, heat transfer, and electrochemical reactions that occur within the battery, which are crucial for predicting its behavior and increase reliability of the battery.

Multi-scale multi-domain (MSMD) modeling is a tool for simulating the behavior of materials at multiple scales. MSMD is capable of simulating the behavior of materials from the atomic scale all the way up to the macroscopic scale, which enables the user to gain a deep understanding of how the material behaves under different conditions. The MSMD model is suitable for simplifying the description of complex electrochemical processes in a battery that has a complicated structure. [3, 4]

The thermal and electrical fields in the battery cell are solved using differential equations in computational fluid dynamics (CFD). [4] These equations allow the behavior of the battery to be simulated:

$$\frac{\partial \rho C_p T}{\partial t} - \nabla \cdot (k \nabla T) = \tag{1}$$

$$= \sigma_+ |\nabla \phi_+|^2 + \sigma_- |\nabla \phi_-|^2 + \dot{q}_{E\mathrm{Ch}} + \dot{q}_{short} + \dot{q}_{abuse}$$

$$\nabla \cdot (\sigma_+ \nabla \varphi_+) = -(-j_{E\mathrm{Ch}} - j_{\mathrm{short}}) \tag{2}$$

$$\nabla \cdot (\sigma_- \nabla \varphi_-) = j_{E\mathrm{Ch}} - j_{\mathrm{short}} \tag{3}$$

Where:

$\rho\ [\frac{kg}{m^3},]$: Density

$T\ [K]$: Temperature

$t\ [s]$: Time

$C_p\ [\frac{J}{kg \cdot K}]$: Specific heat capacity.

$\sigma_+, \sigma_-\ [S]$: The effective electric conductivities for the positive and negative electrodes.

$\varphi_+, \varphi_-\ [V]$: Phase potentials for the positive and negative electrodes.

$j_{E\mathrm{Ch}}, \dot{q}_{E\mathrm{ch}}\ [\frac{A}{m^3}, \frac{W}{m^3}]$: The volumetric current transfer rate and the electrochemical reaction heat due to electrochemical reactions.

$j_{short}, \dot{q}_{short}\ [\frac{A}{m^3}, \frac{W}{m^3}]$: The current transfer rate and heat generation rate due to battery internal short-circuit.

MSMD method includes of three electrochemical submodels:

- The Newman, Tiedemann, Gu, and Kim (NTGK) model.
- Equivalent Circuit Model (ECM).
- Newman's Pseudo-two-dimensional (P2D) model.

These models are used to define the macroscopic value of the $j_{E\mathrm{Ch}}$. This is described in equations 2 and 3.

A. NTGK model

The Newman, Tiedemann, Gu, and Kim (NTGK) model is a highly efficient semi-empirical electrochemical model for modeling batteries, originally proposed by Kwon et al. and widely used by others. [5] The NTGK model can accurately predict current-voltage (I-V) curves under different C-rates, which measure the charge and discharge rate of a battery. The model's formulation is based on the potential difference between the positive and negative electrodes, with the volumetric current transfer rate being a function of this difference. The two parameters in the model follow a polynomial form, and the discharge depth is a variable in this empirical function. It can experimentally obtain polarization curves for a specific battery cell and then determine the two parameters by fitting the data to the curve. Due to its accuracy and efficiency, the NTGK model is used for 3D thermal modeling of Li-ion batteries, particularly in the thermal modeling of prismatic Li-ion battery cells. [5, 6] The volumetric current transfer rate is related to the potential field by the following equation:

$$j_{E\mathrm{Ch}} = \frac{Q_{nomin}}{Q_{ref}\, V_{ol}} Y[U - (\varphi_+ - \varphi_-)] \tag{4}$$

Where:

$Q_{nominal}\ [Ah]$: Battery total capacity.

$Q_{ref}\ [Ah]$: The capacity of the battery that is used in experiments to obtain model parameters Y and U.

$V_{ol}[V]$: The active zone's volume of a single battery.

The depth of discharge (DoD) is calculated using the following equation:

$$DoD = \frac{V_{ol}}{3600 Q_{nominal}} \int_0^t j\, dt \tag{5}$$

The electrochemical reaction heat $\dot{q}_{E\mathrm{ch}}$ is calculated as

$$\dot{q}_{E\mathrm{ch}} = j_{EC} \left[U - (\varphi_+ - \varphi_-) - T \frac{dU}{dT}\right] \tag{6}$$

U and Y are functions of battery the depth of discharge (DOD) as the following set of equations:

$$U = (\sum_{n=0}^{5} a_n\, (DoD)^n) \; - \; C_2 (T - T_{ref}) \tag{7}$$

$$Y = (\sum_{n=0}^{5} b_n\, (DoD)^n)\; exp[-C_1 (\frac{1}{T} - \frac{1}{T_{ref}})] \tag{8}$$

Where:

$Y, U\ [S, V]$: GU's parameters.

$a_n, b_n\ [-]$: Fitting coefficients of the experimental data.

$C_1, C_2\ [-]$: The battery-specific NTGK model constants.

The equations can be decomposed into individual polynomials:

$$U = a_0 + \; a_1 DoD + a_2 DoD^2 + a_3 DoD^3 \; + \; a_4 DoD^4 + a_5 DoD^5 \tag{9}$$

$$Y = b_0 + \; b_1 DoD + b_2 DoD^2 + b_3 DoD^3 \; + \; b_4 DoD^4 + b_5 DoD^5 \tag{10}$$

B. Other E-chemistry models

Chen and Rincon-Mora's equivalent circuit model is capable of predicting both battery runtime and I-V performance. Their circuit comprises of three resistors and two capacitors, with the corresponding resistances, capacitances, and open-circuit voltage being assumed as functions of the state of charge (SOC).

On the other hand, Newman's group developed a physics-based model that employs porous electrode and concentrated solution theories to accurately capture Li-ion migration in the battery. This model considers physical processes such as Li-ion release from solid active particles to the electrolyte, Li-ion insertion into solid active particles, electrochemical reactions on the surface of solid active particles, as well as Li-ion transport in the electrolyte solution, based on the concentrated solution transport theory. By adopting this approach, the model provides a more comprehensive representation of the battery's electrochemical behavior. [7]

The P2D model is a complex mathematical model that can capture the electrochemical processes in Li-ion batteries. Newman's research team employed the concentrated solution theory to develop a mathematical model that can simulate the galvanostatic charge and discharge of a lithium-ion battery with a lithium anode, solid polymer separator, and composite insertion cathode. The model's versatility allows it to encompass various polymeric separator materials, lithium salts, and cathode composites. The model uses superposition to for the interpolation of lithium into the active cathode material, which simplifies numerical computations. [8]

III. Simulation

This experiment deals with the simulation of a battery pack that consists of three 78 Ah battery cells in parallel. Battery cell is Li-ion pouch NMC based. Each battery cell weighs 1073 g and its dimensions are 512 x 97 x 8 mm. From the measured discharge curves (0.1C, 0.2C and 0.5C), U and Y curves were generated using Ansys Fluent software to determine the nature of each battery cell modelled. These curves are plotted based on polynomial equations (9) and (10), whose constants "a" for the U curve and "b" for the Y curve are shown in Table 1 and the individual curves are plotted in Fig. 2 and Fig. 3.

This parallel battery pack is the basic building unit for a battery pack, which consist of eight of these battery packs with capacity of 248 Ah connected in series.

The battery pack model was created in 3D CAD Solidworks software and meshed in Ansys Mechanical. The number of elements in 3D model equals 106 456. Active material is created by tetrahedrons structure and electrodes by hex dominant meshing.

The simulated battery pack consists of an active material that has a defined density, electrical conductivity and electrodes (positive and negative tabs) that are made of aluminum. Heat passes through the walls of the active material, but electric charge does not, these walls are defined as coupled walls. Both electric charge and heat pass through the tabs. All material has a heat transfer coefficient set to 5 W/m^2k.

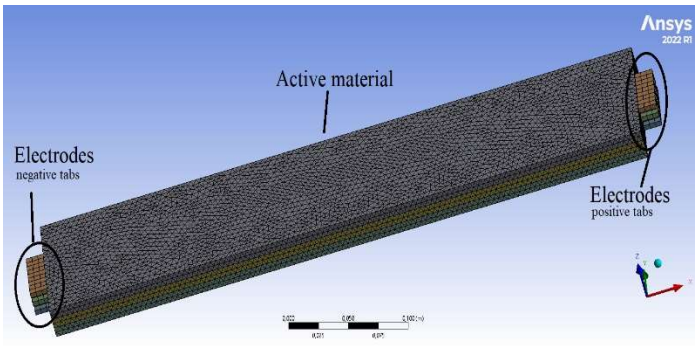


Fig. 1: Battery pack

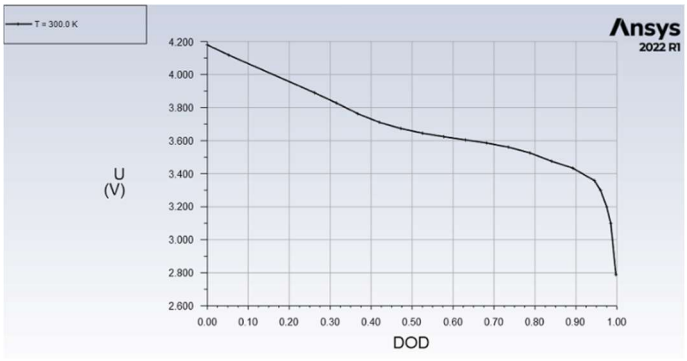


Fig. 2: U function

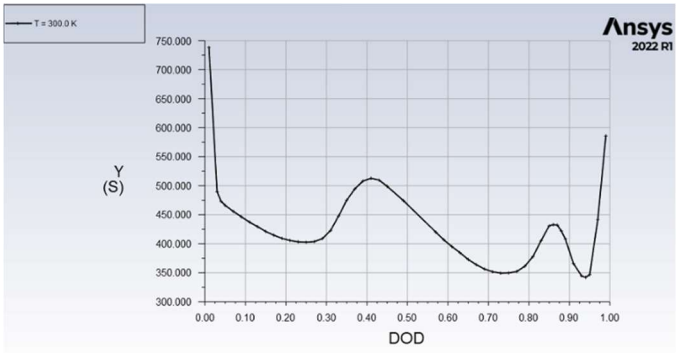


Fig. 3: Y function

IV. Results

From the following graphs, it can be seen that the discharge time of the real battery is slightly shorter than the discharge time of the simulated battery. This phenomenon could be due to the inaccurate fitting of the U and Y parameters evaluated by the ANSYS Fluent software.

The parameter fitting is directly dependent on the input measured data, if there are any anomalies in the measured data, this has a large impact on the determination of the constants a_n and b_n and therefore the overall appearance of the U and Y curves, which further invalidates the simulated models.

TABLE I. U AND Y CONSTANTS

index	0	1	2	3	4	5
a	*4.1768*	*-1.1203*	*0.47882*	*-4.0055*	*9.93097*	*-6.3301*
b	427.017	-1670.8	10890.7	-24766	23149.7	-7595.3

A. Passive zone potential

The simulation of the Li-ion battery was performed at 0.1 C, 0.2 C and 0.5 C discharge as well as the battery was measured, and the voltage drop and heating of the battery were monitored as a function of time. The batteries were simulated with a time step of 20 seconds.

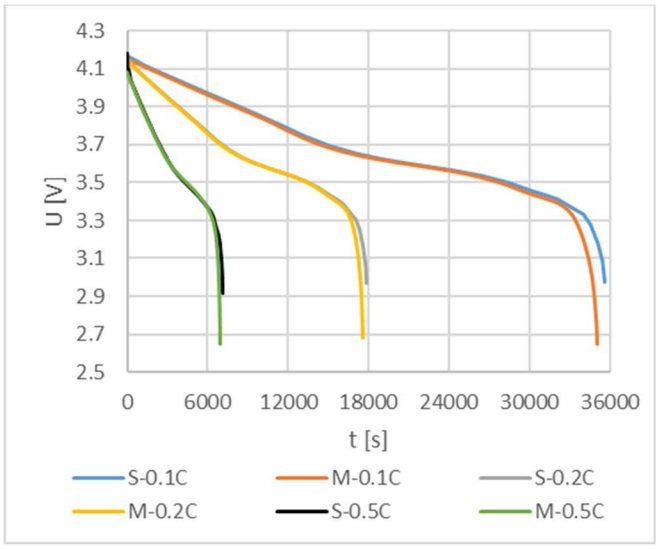


Fig. 4: Discharge curves (0.1C; 0.2C; 0.5C)

B. Static temperature

The temperature rise of the simulated battery pack depends on the discharge rate of the battery exponentially. The close packing of the cells also results in more heating of the cells inside the battery pack than at the edges. The difference in battery heating at 0.5 C and 0.1 C is approximately 18.5 K. The results show that thermal analysis is necessary at higher C-rate, when the temperature rises sharply as the battery discharges.

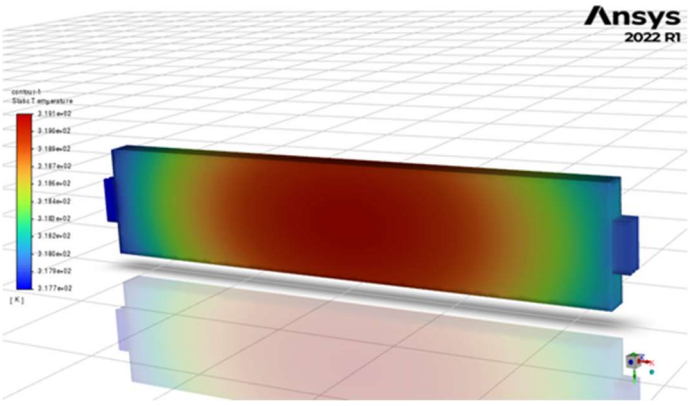


Fig. 5: Thermal contours simulation

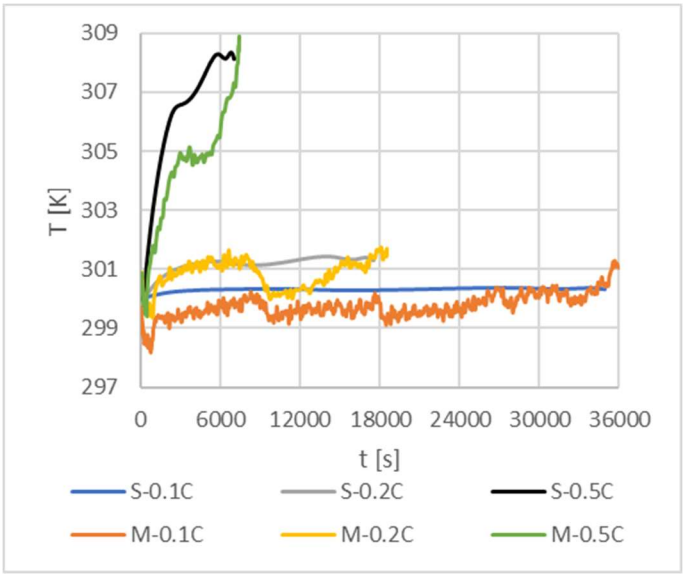


Fig. 6: Thermal simulation of battery pack

C. Current density

The following figure shows the current density in tabs and active material. Current magnitude is much higher near the tabs.

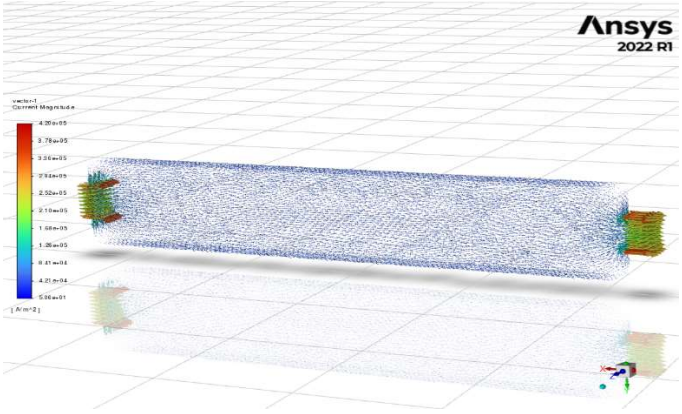


Fig. 7: Current density

V. Conclusion

A 3D model of a Battery pack with a total capacity of 234 Ah was created, consisting of three battery cells with capacity of 78 Ah. The model is based on the experimentally measured discharge curves of the battery cell at 0.1 C, 0.2 C and 0.5 C, and the polynomial functions U and Y were determined from these measured data. These functions define the nature of the active material of the battery cell. The developed model was simulated at the same C-rate as measured and the discharge characteristics were evaluated and compared with those measured in Fig. 4. The battery was also subjected to thermal analysis. The maximum error of the simulated value from the measured value is 2.29 K with a discharge curve of 0.5 C. Excluding the beginning of the measurement, the maximum error at 0.1 C discharge curve is 1.19 K and at 0.2 C it is 1.27 K. This error is due to conditions that were not ideal when measuring the temperature of the real battery. The thermal measuring is important during rapid discharge when the battery temperature rises rapidly. This can lead to accelerated degradation and in critical situations the battery can explode. Finally, a current density analysis was performed on the battery in the model volume.

Acknowledgment

This work was supported by the BUT specific research program (project No. FEKT-S-23-8286).

References

[1] GOODENOUGH, John B. a Kyu-Sung PARK. The Li-Ion Rechargeable Battery: A Perspective. Journal of the American Chemical Society. 2013, 135(4), 1167-1176. ISSN 0002-7863. Dostupné z: doi:10.1021/ja3091438

[2] DARJAZI, H., E. GONZALO, B. ACEBEDO, R. CID, M. ZARRABEITIA, F. BONILLA, M.Á. MUÑOZ-MÁRQUEZ a F. NOBILI. Improving high-voltage cycling performance of nickel-rich NMC layered oxide cathodes for rechargeable lithium–ion batteries by Mg and Zr co-doping. Materials Today Sustainability. 2022, 20. ISSN 25892347. Dostupné z: doi:10.1016/j.mtsust.2022.100236

[3] BLAZEK, Pavel, Peter WESTENBERGER, Simon ERKER, et al. Axially and radially inhomogeneous swelling in commercial 18650 Li-ion battery cells. Journal of Energy Storage. 2022, 52. ISSN 2352152X. doi:10.1016/j.est.2022.104563

[4] MADANI, Seyed Saeed, Maciej Jozef SWIERCZYNSKI a Soren Knudsen KAER. The discharge behavior of lithium-ion batteries using the Dual-Potential Multi-Scale Multi-Dimensional (MSMD) Battery Model. 2017 Twelfth International Conference on Ecological Vehicles and Renewable Energies (EVER). IEEE, 2017, 2017, 1-14. ISBN 978-1-5386-1692-5. doi:10.1109/EVER.2017.7935915

[5] SEONG KIM, Ui, Jaeshin YI, Chee Burm SHIN, Taeyoung HAN a Seongyong PARK. Modeling the Dependence of the Discharge Behavior of a Lithium-Ion Battery on the Environmental Temperature. Journal of The Electrochemical Society. 2011, 158(5). ISSN 00134651. doi:10.1149/1.3565179

[6] KWON, Ki Hyun, Chee Burm SHIN, Tae Hyuk KANG a Chi-Su KIM. A two-dimensional modeling of a lithium-polymer battery. Journal of Power Sources. 2006, 163(1), 151-157. ISSN 03787753. Dostupné z: doi:10.1016/j.jpowsour.2006.03.012

[7] CHEN, Min a G.A. RINCON-MORA. Accurate electrical battery model capable of predicting runtime and I-V performance. IEEE Transactions on Energy Conversion. 2006, 21(2), 504-511. ISSN 0885-8969. Dostupné z: doi:10.1109/TEC.2006.874229

[8] JOKAR, Ali, Barzin RAJABLOO, Martin DÉSILETS a Marcel LACROIX. Review of simplified Pseudo-two-Dimensional models of lithium-ion batteries. Journal of Power Sources. 2016, 327, 44-55. ISSN 03787753. doi:10.1016/j.jpowsour.2016.07.036 P2D

Image Reconstuction in Electrical Impedance Tomography with Total Variation and Gauss-Newton Algorithm

Serge Ayme Kouakouo Nomvussi[1], Jan Mikulka[2]
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology
Brno, Czech Republic
[1]244245@vut.cz, [2]mikulka@vut.cz

***Abstract*— In this paper, the Total Variation and Gauss-Newton algorithms are used to reconstruct the conductivity maps in Electrical Impedance Tomography. The Total Variation gives better results than the Gauss-Newton algorithm. The quality of the reconstructed conductivity maps is clear and almost identical to the image of the forward problem when using the Total Variation method on the whole dataset. In contrast, the Gauss-Newton algorithm provides a smooth image of the object without a perfect reconstruction of its edges. The correlation coefficient of the Total Variation is higher than that of the Gauss-Newton method. The drawback of these two algorithms is their inability to reconstruct the image from noisy data.**

Keywords— Total Variation; Gauss Newton; EIT; EIDORS.

I. INTRODUCTION

Electrical Impedance Tomography (EIT) is an imaging technique used to reconstruct the conductivity distribution of the target in a volume conductor using voltage and current data measured at the domain boundary [1]. EIT measurement system consists of hardware and software. The hardware is impacted by the stability of the current source. The EIT system will be more coherent if the current supply is also more stable. The number of electrodes also influences the results of the reconstructed images. It is proposed to use 16 electrodes in the EIT system to obtain a good two-dimensional image reconstruction result [2]. A pair of electrodes in the measuring equipment is used to introduce a low intensity alternating current into the domain to be measured. The resulting voltage on the other pairs of electrodes is then measured as a potential difference. This measured boundary voltage will be used to reconstruct the conductivity distribution map [3].

The software of a EIT system consists of the forward model and the inverse model. The forward model is a difficult part that involves a design process of a phantom and electrode geometry through a data collection technique. The data acquisition that will be implemented involves a forward solving function of the problem. This problem is to find the voltage distribution of the current injection from the conductivity distribution of the object [4, 5]. The inverse model is a part that completes the image reconstruction process. The model applies a different type of image reconstruction by comparing two data of the same object. The first data is called homogeneous data and is determined when the phantom medium is water. The second data is called inhomogeneous data and is determined when the phantom contains water and an object. This part contains the algorithm that is applied to complete the reconstruction process and the value of the hyperparameter that affects the image quality [6].

The reconstruction of the conductivity distribution map in EIT is an ill-posed problem [7, 8]. Despite that, the resolution of the inverse problem can be solved by using a multitude of algorithm types. The Gauss-Newton method gives an accurate detection position, features and a smooth image [9, 10]. However, this method does not display the edge of the object side in the reconstruction. For lung cancer monitoring, the Thiknov regularization algorithm can be used to recover the conductivity maps of the body [11]. The drawback of Tikhonov regularization is that it does not preserve sharp discontinuities in the image [12]. The Total Variation (TV) permits the regularization of the image without imposing smoothing. The Total Variation function plays an important role in the regularization of inverse problems in many disciplines, due to its ability to preserve discontinuities in the reconstructed profiles [10].

This study will examine the Total Variation algorithm and compare it to the Gauss-Newton method on a synthetic dataset to determine which algorithm performs better.

II. METHODS

For this study, a synthetic dataset is generated with the framework EIDORS, which is designed to solve the forward and inverse modelling in EIT. The only requirement to use EIDORS is to have MATLAB installed on your computer. The dataset generation Script has a main program that call three different functions where the relevant parameters for a sample are defined and automatically modified during the process. In contrast to other research, the samples are created using circular and rectangular targets with varying conductivity and size in different positions, placed in a bucket containing water. This water has a conductivity of 0.04 S/m and the target conductivity varies between 0.01 S/m and 0.12 S/m. All around the bucket, there are electrodes that are needed to inject current in the system and to measure the voltage between two electrodes according to the chosen injection and voltage measurement pattern as illustrated in the Fig. 1. The bucket has a diameter of 100 mm as presented in the Fig. 2 and the circular target has

diameter in a range of 6 to 10 mm. The length of the rectangular target is 5 mm, and its width is between 2 and 5 mm. For our study, we have opted for an adjacent injection and voltage measurement pattern. A second case where the injection and voltage measurement are separated by two electrodes is also considered. The finite element method (FEM) is based on a homogeneous cylindrical object 'd2c', which has 1024 pixels with a pattern of 16 complete electrodes. The target for a sample is either circular or rectangular, or both shapes for the whole dataset. The measured data were produced with an injected current of 1 mA at 20 kHz. The measured voltage of the dataset is in a range of 0.43 to 28.26 mV and the value of the conductivity of the reconstructed image is between 0.04 and 0.29 S/m. The dataset has 4056 samples. During the voltage measurement process, the noise with several levels is added to the input vectors to achieve signal to noise ratio of 50 to 100 dB.

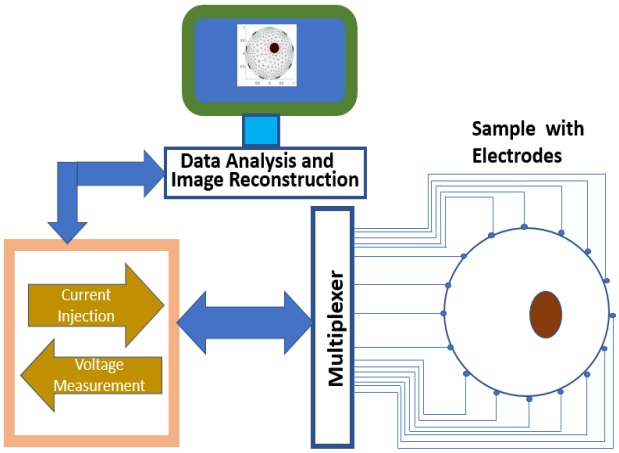


Fig. 1 EIT System with electrodes

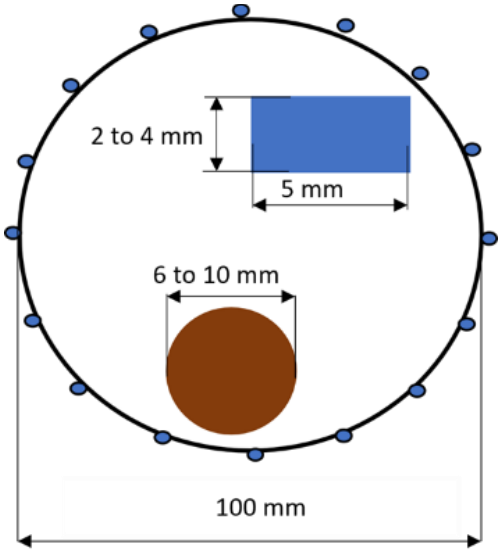


Fig. 2 Bucket and target size

The standard forward problem in EIT is, given the conductivity distribution σ and the low current J injected via the boundary electrodes, to find the potential distribution ϕ within Ω and in particular the resulting potentials at the measuring electrodes ∂_m. The quasi-static approximation is valid and the capacitive and inductive effects are ignored since the frequencies used in the EIT are sufficiently low [13, 14]. Under these quasi-static conditions, the solution of the direct problem is quite simple, since it is sufficient to solve the following Laplace equation:

$$\nabla.(\sigma\nabla\phi) = 0 \quad (1)$$

At the boundary, currents are injected through electrodes; the current density across the surface of the l-th electrodes J_l is therefore given:

$$\sigma\frac{\partial\phi}{\partial\hat{n}} = J_l \quad (2)$$

Where $\hat{n}$ is the outward unit normal vector on electrode surface

It is necessary to mention that the solution of the forward problem is a very important step for image reconstruction in EIT which is achieved by solving the inverse problem. For the real-world problems, there are many numerical methods that can be used to solve the forward problem such as the boundary element method (BEM), the finite element method (FEM), the finite difference method (FDM) and the finite volume method (FVM) [15, 16]. For this study, the finite element method is considered to solve the problem because of its ability to build a model for an object of any shape and to handle inhomogeneous material properties [17, 18].

A. *Total Variation Algorithm*

In almost all areas of EIT application, there are situations where the imaged conductivity shows discontinuities. In the medical domain, there are examples of inter-organ boundaries where each organ has its specific electrical properties. In the field of archaeology, a buried wall causes a sudden change in conductivity, and in process tomography, a multiphase fluid causes discontinuities at each phase interface. Therefore, it is crucial to reconstruct these situations correctly, even if such conductivities are difficult to handle by traditional algorithms. Many approaches have been explored to overcome these limitations. They can often be seen to introduce prior information. One example is anisotropic regularization in which the structure of the expected sudden changes is supposed to be roughly known [19, 20]. The smoothing constraints are consequently relaxed in the normal direction of the discontinuities. In this way, the algorithm can better describe the fast variations of the object, but prior structural information is required before such methods can be adopted.

Many regularization matrices are discretized representations of differential operators and are used in combination with the L2 -norm. A distinct approach is illustrated by the choice of the Total Variation functional, which is still a differential operator but leads to a L1-regularization [21]. The Total Variation algorithm maintains the value of the discontinuity on the resulting reconstruction so that the images obtained have sharp edges at the discontinuity. In the medical sector, the difference

in discontinuity comes from the difference in the conductivity value of the tissue. Therefore, it generates different colors which are given by the following equation:

$$TV(\theta) = \int |\nabla\theta| d\Omega \quad (3)$$

where TV is the Total Variation which depends on the conductivity image. In addition, θ is the conductivity and Ω is the area where the images are produced.

The Total Variation function was first used to regularise the restoration of images with noise [22]. This method is very effective for recovering "bulk" images and has been widely accepted by the image restoration community. The effectiveness of the approach in recovering discontinuous images can be appreciated by considering the following one-dimensional situation.

Let us assume that the points A and B in Fig. 3 are connected by a path. Four possible functions f(x) connecting them are represented in the Fig. 3.

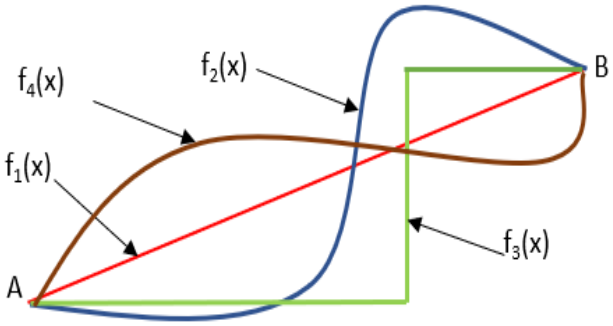


Fig. 3 Points A and B connected via several different paths.

Since the functions are monotonically increasing, the TV of each function is describing in the following equation:

$$TV(f) = \int_A^B f'(x)dx = f(B) - f(A) \quad (4)$$

Total Variation treats the functions f_1, f_2 and f_3 identically and, by using it as a penalty term in a regularised inverse Tikhonov problem, it will not bias the result towards the smooth solution. In contrast, the L2-norm takes on different values for f_1, f_2 and f_3. When used as a penalty term, the L2- norm biases the solution towards smoother functions, for which the L2-norm takes smaller values. In the example given in Fig.3, f_3 is not a quadratic solution because its L2- norm is infinite. The use of Total Variation as a regularisation penalty term therefore allows a much larger class of functions to be used to solve the inverse problem, including functions with discontinuities. Assuming that the conductivity is described by piecewise constant elements, the Total Variation of a 2D image is expressible as the sum of the Total Variation of each of the k edges, where each edge is weighted by its length.

$$TV(\sigma) = \sum_k l_k |\sigma_{m(k)} - \sigma_{n(k)} \quad (5)$$

where l_k is the length of the k_{th} edge of the mesh, m(k) and n(k) are the indices of the two elements on either side of the kth edge, and the index k extends to all edges. Equation (5) can be expressed in terms of matrices as:

$$TV(\sigma) = \sum_k |L_k\sigma| \quad (6)$$

where L is a sparse matrix, with one row for each edge of the mesh. Each row Lk has two non-zero elements in the columns m(k) and n (k).

B. Gauss-Newton Algorithm

The Gauss-Newton algorithm is useful for solving non-linear least squares problems, which is equivalent to minimizing the sum of the values of a squared function. This algorithm is an extension of Newton's method for determining the minimum of a non-linear function. Since a sum of squares must be non-negative, the approach can be considered as using Newton's method to approximate iteratively the zeros of the components of the sum, and consequently minimize the sum. The algorithm provides an efficient method for solving systems of overdetermined equations and does not require second derivatives, which can be difficult to solve.

EIT System is a nonlinear and ill-posed inverse problem in which the approximate solution is obtained with the minimization technique. The minimization of the objective function is achieved by the difference between the measured and predicted data. A solution $\hat{x}$ to the inverse problem in EIT is specified as the minimum of a sum of quadratic norms when the Gauss-Newton algorithm is used [22].

$$\hat{x} = ||z - Jx||^2 + ||x - x_0||^2 \quad (7)$$

Where x represents the change of conductivity, x_0 the expected value of conductivity changes, z the voltage changes and J Jacobi Matrix. By solving equation (7), a one-step linear inverse problem can be obtained and calculated using the following equation.

$$\hat{x} = (J^T W J + \lambda^2 R)^{-1} J^T W z \quad (8)$$

Where R is the matrix regularization, **λ** the hyperparameter or regularization parameter and W the weighted matrix.

The correlation (r) calculated in equation 9 is used to evaluate the degree of correlation between the recovered image (Y^*_{mn}) and the original image (Y_{mn}). The greater the value of the correlation coefficient, the stronger the correlation and the higher the quality of the reconstructed conductivity map. $\overline{Y^*_{mn}}$ and $\overline{Y_{mn}}$ are respectively the mean value of the recovered and the original image.

$$r = \frac{\sum_m \sum_n (Y^*_{mn} - \overline{Y^*_{mn}})(Y_{mn} - \overline{Y_{mn}})}{\sqrt{(\sum_m \sum_n (Y^*_{mn} - \overline{Y^*_{mn}})^2)(\sum_m \sum_n (Y_{mn} - \overline{Y_{mn}})^2)}} \quad (9)$$

III. Results

Table 1 shows that both algorithms can reconstruct conductivity maps of phantom with one or two targets without

noise very well. The detection of the object in the exact position with the Gauss-Newton method is possible, but the boundaries of the object are not clearly visible. Consequently, the size of the object cannot be measured as well, although its correlation coefficient for the case of detecting one or two objects with noise-free input dataset is very high. All around the object reconstructed with the Gauss-Newton method, there are colored representations. The Gauss-Newton algorithm is not able to reconstruct the image of the noisy input data.

The Total Variation method produces a clear image on the edge of the object pattern and the resulting color contrast is good. But a very bright image on the side of the object may be the result of a less smooth image. The edges of the reconstructed images with the Total Variation algorithm are not exact when comparing them with the original images. The correlation coefficient of the Total Variation algorithm is better than the correlation coefficient of the Gauss-Newton algorithm on the whole dataset. The Total Variation algorithm is not also able to reconstruct the image of the noisy input data.

TABLE I. RECONSTRUCTED IMAGE AND CORRELATION COEFFICIENT

Position / Noise	Simula-tion model	Reconstructed image Gauss-Newton algorithm	Reconstructed image With Total Variation
1/No		 r = 0.9840	 r = 0.9920
2/No		 r= 0.9658	 r = 0.9816
3/No		 r = 0.9699	 r = 0.9838
4/No		 r = 0.9401	 r = 0.9553
5/No		 r = 0.9581	 r = 0.9652
6/Yes 50 dB		 r = 0.5683	 r = 0.5650
7/Yes 70 dB		 r = 0.5682	 r =0.5650

IV. DISCUSSION

Reconstruction of conductivity maps of a phantom with one or more targets in Electrical Impedance Tomography with the Total Variation and Gauss-Newton algorithm are good solutions for a noise-free dataset. This paper highlights that the Total Variation can reconstruct the image of any data without noise. Therefore, it should be used in medical and process imaging applications of EIT where the non-smooth reconstruction technique is needed. Furthermore, the Gauss-Newton algorithm gives a smooth image of the object without a perfect reconstruction of its edges. The edges of the reconstructed image with the Gauss-Newton algorithm are not clear, whereas they are clear with the Total Variation algorithm. It will therefore be necessary to combine the two algorithms and study their effects on this complex dataset in order to improve the quality of the reconstructed image. Furthermore, the artificial neural network is not sensitive to noise and should be used in further work to train this complex synthetic dataset.

ACKNOWLEDGMENT

The author S. Kouakouo thanks the Department of Theoretical and Experimental Electrical Engineering at Brno University of Technology. The preparation of this paper was assisted by the project No. FEKT-S-23-8425 of the BUT internal grant office.

REFERENCES

[1] Michalikova, Marketa, Rawia Abed, Michal Prauzek, and Jiri Koziorek. 'Image Reconstruction in Electrical Impedance Tomography Using Neural Network'. In *2014 Cairo International Biomedical Engineering Conference (CIBEC)*, 39–42, 2014. https://doi.org/10.1109/CIBEC.2014.7020959.

[2] Khalighi, M., et al. "Design and implementation of precise hardware for electrical impedance tomography (eit)*." *Iran J Sci Technol Trans Electr Eng* 38 (2014): 1.

[3] Khan, Talha Ali, and Sai Ho Ling. 'Review on Electrical Impedance Tomography: Artificial Intelligence Methods and Its Applications'. *Algorithms* 12, no. 5 (26 April 2019): 88. https://doi.org/10.3390/a12050088.

[4] Harikumar, R., R. Prabu, and S. Raghavan. "Electrical impedance tomography (EIT) and its medical applications: a review." *Int. J. Soft Comput. Eng* 3.4 (2013): 193-198.

[5] Adler, Andy, and David Holder, eds. *Electrical impedance tomography: methods, history and applications*. CRC Press, 2021.

[6] Adler, Andy, and Robert Guardo. "Electrical impedance tomography: regularized imaging and contrast detection." *IEEE transactions on medical imaging* 15.2 (1996): 170-179.

[7] 'Xu et al.-2010 -Electrical Resistance Tomography for Locating Incl.Pdf'. Accessed 9 January 2022. https://ir.nsfc.gov.cn//paperDownload/1000001066642.pdf.

[8] Liu, Shengheng, Jiabin Jia, Yimin D. Zhang, and Yunjie Yang. 'Image Reconstruction in Electrical Impedance Tomography Based on Structure-Aware Sparse Bayesian Learning'. *IEEE Transactions on Medical Imaging* 37, no. 9 (September 2018): 2090–2102. https://doi.org/10.1109/TMI.2018.2816739.

[9] Islam, Md Rabiul, and Md Adnan Kiber. "Electrical impedance tomography imaging using Gauss-Newton algorithm." *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*. IEEE, 2014.

[10] Farha, Maksalmina. "Combined algorithm of total variation and Gauss-newton for image reconstruction in two-dimensional Electrical Impedance Tomography (EIT)." *2017 International Seminar on Sensors, Instrumentation, Measurement and Metrology (ISSIMM)*. IEEE, 2017.

[11] Sun, Benyuan, et al. "An improved Tikhonov regularization method for lung cancer monitoring using electrical impedance tomography." *IEEE sensors journal* 19.8 (2019): 3049-3057.

[12] Vauhkonen, Marko, et al. "Tikhonov regularization and prior information in electrical impedance tomography." *IEEE transactions on medical imaging* 17.2 (1998): 285-293.

[13] Rudin, Leonid I., Stanley Osher, and Emad Fatemi. "Nonlinear total variation based noise removal algorithms." *Physica D: nonlinear phenomena* 60.1-4 (1992): 259-268.

[14] Tavares, Renato S., et al. "Discretization error and the EIT forward problem." *IFAC Proceedings Volumes* 47.3 (2014): 7535-7540.

[15] Dong, Guoya, et al. "The comparison between FVM and FEM for EIT forward problem." *IEEE Transactions on Magnetics* 41.5 (2005): 1468-1471.

[16] Xu, Guizhi, et al. "3-D electrical impedance tomography forward problem with finite element method." *IEEE transactions on magnetics* 41.5 (2005): 1832-1835.

[17] Bera, Tushar Kanti, and J. Nagaraju. "A FEM-based forward solver for studying the forward problem of electrical impedance tomography (EIT) with a practical biological phantom." *2009 IEEE International Advance Computing Conference*. IEEE, 2009.

[18] Sarode, Vidya, Priya M. Chimurkar, and Alice N. Cheeran. "Electrical impedance tomography (EIT) based medical imaging using finite element method (FEM)." *International Journal of Engineering Sciences & Emerging Technologies* 1.2 (2012): 83-89.

[19] Borsic, Andrea, William RB Lionheart, and Christoher N. McLeod. "Generation of anisotropic-smoothness regularization filters for EIT." *IEEE transactions on medical imaging* 21.6 (2002): 579-587.

[20] Kaipio, J. P., et al. "Construction of nonstandard smoothness priors." *Inverse Problems* 15 (1999): 713-729.

[21] Borsic, Andrea, et al. "Total variation regularization in electrical impedance tomography." (2007).

[22] Islam, Md Rabiul, and Md Adnan Kiber. "Electrical impedance tomography imaging using Gauss-Newton algorithm." *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*. IEEE, 2014.

Machine Learning for Antenna Design: Combining CST Studio Suite and Python

Vojtěch Bednarský
Author, Department of Radio electronics
Brno University of Technology
Brno, Czech Republic
ORCID 0000-0002-5910-4379
E-mail: *203399@vutbr.cz*

Zbyněk Raida
Supervisor, Department of Radio electronics
Brno University of Technology
Brno, Czech Republic
ORCID 0000-0003-4466-5496
E-mail: *raida@vutbr.cz*

***Abstract*—The design and optimization of antennas is a complex and time-consuming process which combines an electromagnetic analysis to evaluate cost functions and machine learning to consequently improve designs. In this paper, CST Studio Suite performs the numerical analysis, and Python scripts implement other steps. Python executes numerical operations, automatically generates models, and supports the CST analyses without requiring user's interaction. Ultimately, the approach is aimed to utilize Python's libraries PyTochr and TensorFlow to automate antenna designs, which can be leveraged by artificial intelligence, at a later stage.**

Keywords—CST Studio Suite, Python, PyTochr, TensorFlow, particle swarm optimization (PSO), canonical antenna

I. INTRODUCTION

The antenna design and optimization are critical processes in the development of modern communication systems. When designing a model of an antenna, an appropriate geometry has to be selected to obtain resonant frequencies in the requested operational bands. The selected geometry is then modelled in CST and optimized to meet the required parameters as accurately as possible. For the optimization, CST optimisation tools can be used which might be of a slower convergence and can require a proper setup to produce reliable results.

Objective functions, are formulated to optimize electromagnetic problems, are nonlinear and nondifferentiable, contain several minims and are computationally expensive to evaluate [8].

In order to address these challenges, we use Python scripts and CST. Scripts implement artificial neural networks (ANN), and CST performs an electromagnetic analysis. When connected, the designer's experience can be trained well. ANN are trained on a set of different antenna designs that produce approximate training patterns efficiently.

For instance, the *Tera* library in Python [1] can be utilized to implement an ANN and obtain precise outcomes in tasks such as selecting the appropriate input parameters for the initial iteration [12].

The paper is focused on the integration of CST Microwave Studio (CST) with Python scripts. In the initial stage, a canonical antenna structure (described in Section 3) is analysed by CST and modified by Particle Swarm Optimization (PSO).

The studied design methodology can be subdivided into three main parts:

1. A numerical model of an antenna is implemented in CST using a Python script.
2. A PSO script is programmed in Python to evaluate the implemented antennas from the viewpoint of optimal impedance matching.
3. The optimized numerical model is compared with a manufactured antenna.

The exploitation of PSO for a global optimization of antenna structures is quite popular [2]-[5], and the design approaches exploiting CST are on extensive use [6]. Nevertheless, the scripting is mostly based on Visual Basic for Applications (VBA) [7], and the CST-Python association is quite rare [8]-[9]. The ambition of this paper is therefore to contribute to this development.

In Section II, attention is turned to the Python-CST interface. Section III presents PSO outputs and Section IV is devoted to experimental verifications. Section V closes the paper.

II. PYTHON-CST INTERFACE

The Python interface for CST allows users to automate repetitive tasks, integrate CST with other software tools, and develop custom applications that take advantage of the power of CST. The interface provides access to all of the functionality of CST, including the ability to create and modify models, define simulations, and analyse results.

A. Settings

To use the Python interface for CST, users need to have both CST and the Python programming language installed on their computer. The interface is based on the COM (Component Object Model) technology, which is a Microsoft technology for software component interoperation [10].

In order to setup, the sequence of six steps should be followed:

- CST Microwave Studio (not the student version) and Python 3.6.12 or newer are installed. It is recommended to install the Anaconda software, which includes the Spyder development environment.
- To use the Simulia Python API, the CST installation directory is important to be in the system PATH environment variable so that Python can find the API, e.g.: `C:\ProgramFiles(x86)\CSTStudioSuite202x\AMD64\python_cst_libraries\cst`

- All packages and libraries using the Anaconda Prompt should be updated by running the commands:
```
conda update conda
conda update anaconda
conda update spyder
```
- Spyder in Anaconda should be open
- Python script should be written in Spyder
- The following code should be run to test settings:
```
import sys
import cst
sys.path.append(r"C:\Program Files
(x86)\CST Studio Suite 2022\
AMD64\python_cst_libraries")
print(cst.__file__)
```
This should print the path to your Python CST libraries.
- The Python script should be run.

B. Types of the syntax

Simulia has started to release the Python API with CST 2020. The API is built on top of Python programming language. The users can create scripts to automate repetitive tasks, perform parameter sweeps and optimisation studies, and interact with external data sources. This saves time and increases productivity of complex simulations [6].

An example of using the Simulia Python API to open a new Microwave Studio project in the quiet mode follows [11]:

- We create a new instance of the DesignEnvironment class, which is part of the CST Design Studio API, by calling the function `cst.interface.DesignEnvironment()`.
- We call the method `new_mws()` on the DesignEnvironment object to create a new CST Microwave Studio project and return a reference to it.
- We set the quiet_mode parameter to None using the method `set_quiet_mode(None)` on the DesignEnvironment object. This step is optional and can be omitted if CST Microwave Studio is run in quiet mode.
- To run simulations through syntax, we call the method `run_solver()` on the modeler property of the project reference returned by `new_mws()`.

The Simulia Python API is still relatively new and may not yet have the same level of functionality as the VBA interface in some areas. One way to use the VBA code in Python is to convert the VBA code to Python code manually. This involves understanding the logic of the VBA code and translating it into equivalent Python code. While this can be time-consuming and error-prone, it is a good option if the VBA code is relatively simple and straightforward. Another option is to use the win32com library in Python to interact with the Microsoft Office applications. This library allows you to automate these applications using Python, and it can be used to call VBA functions and macros from Python. To use the win32com library, you need to install the Python pywin32 package, which provides the necessary interfaces to interact with the Windows COM system.

As previously mentioned, not all functions are yet available in the Simulia Python API. If there is a need to modify the dimensions of the parameters, it will be necessary to use the VBA code [11]. In CST, it is possible to edit each step that was performed and copy the function text by opening the history list and double-clicking on the relevant entry. This text can also be used in Python, but it must be preceded by a header with the format `name_of_function = """`, followed by the code from the history list, and concluded with `"""`. If it is desired to add this step to the history list, the syntax `"mycst1.modeler.add_to_history("optional_name ", name_of_function)"` must be included.

Below is an example of how VBA code could be implemented in Python to communicate with CST in particular case:

```
#Change frequency range and Solver Type
Full_Historyy = """
ChangeSolverType "HF Frequency Domain"
Solver.FrequencyRange "fmin ", "fmax "
"""
mycst1.modeler.add_to_history("Fsolver",
Full_Historyy)
```

III. PARTICLE SWARM OPTIMIZATION

Particle Swarm Optimization (PSO) is a population-based algorithm that mimics the social behaviour of birds flocking or fish schooling. It starts with a set of candidate solutions, and iteratively updates their positions based on the best solution found so far. PSO aims to find the optimal solution by balancing exploration and exploitation of the search space [2]-[5].

In antenna design, PSO can be used to optimise the antenna geometry (dimensions, location of elements), to achieve specific performance goals. Using PSO, antenna designers can quickly explore a large design space and find solutions that are difficult to obtain by manual optimization.

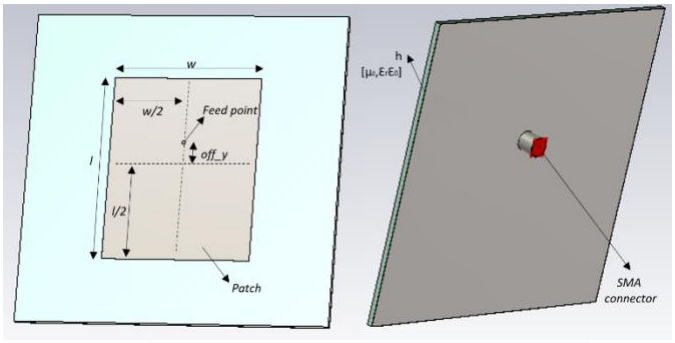


Fig. 1. Top and bottom views of a probe-fed patch antenna

In order to demonstrate the optimization of a CST antenna model by a PSO Python script, a patch element fed by a coaxial probe is optimized from the viewpoint of impedance matching at the ISM frequency f = 2.4 GHz

$$\min |S11(\mathbf{x})|$$

where $0 < |S11| < 1$ is the magnitude of the reflection coefficient at the antenna input and $\mathbf{x}$ is the state vector comprising antenna parameters shown in Figure 1.

Figure 2 shows the PSO flowchart in Python. During initialization, the number of iterations G and agents I is selected. The coefficients c_1 and c_2 represent the contribution of the "thrust" towards p_{BEST} and g_{BEST}, respectively, as shown in

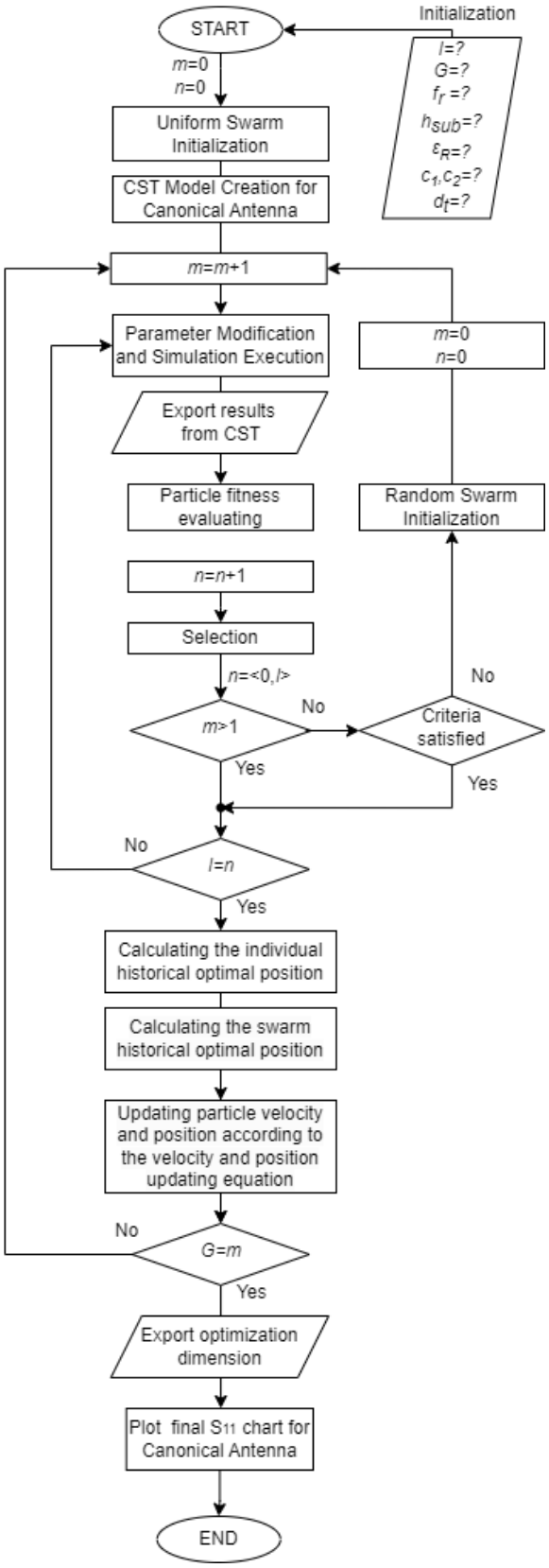


Fig. 2. Flowchart of the particle swarm optimization algorithm

$$v_n = w \cdot v_n + c_1 \cdot rand() \cdot (p_{BEST,n} - x_n) + c_2 \cdot rand() \cdot (g_{BEST,n} - x_n)$$

Updating particle velocity and position is carried out using

$$x_n = x_n + \Delta t \cdot v_n$$

Here, *w* represents the coefficient for the weight factor, which determines the amount of the original velocity that the particle will assume in the next time step. The function *rand*() generates a random number from 0 to 1 and the parameter Δt corresponds to the length of the time-step per one iteration

In order to optimise the desired parameters, it is important to determine the range of values for each parameter. In this case, the parameters *L* and *off_y* were modified while the parameter *w* was fixed to the same value.

The number of particles *I*, *G* in the PSO swarm affects the speed and accuracy of the optimisation. Particle positions should be initialized randomly and uniformly. The difference between these two types of initialisations for the first iteration (I_m = 1) is shown in Table I and Table II. For the first iteration, particles should be initialised uniformly, while for subsequent iterations, they should be initialized randomly. If the criteria are not satisfied during the first iteration, the initialisation will also be performed randomly.

TABLE I. PSO OPTIMIZATION WITH UNIFORM DISTRIBUTION FOR *FREQ*=2,4 GHZ

G_m	Number of iterations	10
I_n	Number of agents	25
time	Time consumption	153 min
f_r	Found f_r by PSO	2.408 GHz
w	Patch width	38.370 mm
L	Patch length	28.950 mm
off_y	Offset from the centre	6.802 mm

TABLE II. PSO OPTIMIZATION WITHOUT UNIFORM DISTRIBUTION FOR *FREQ*=2,4 GHZ

G_m	Number of iterations	8
I_n	Number of agents	25
time	Time consumption	118 min
f_r	Found f_r by PSO	2.372 GHz
w	Patch width	38.370 mm
L	Patch length	90.950 mm
off_y	Offset from the centre	8.480 mm

Next, particle velocities are initialized. Finally, the best personal and global positions are set. For each particle, we assign the current position as the personal best position and the best position among all particles as the global best position.

When the uniform input distribution of the parameter ***w*** was used during the initial iteration, the agents began to accumulate around the correct initial resonant frequency starting from the fifth iteration, as demonstrated in Fig. 4 and Fig. 5.

In Figures 3 to 5, optimization iterations (*I* = 1 to 10) are shown for 25 agents (*G* = 25). The final dimensions (g_{BEST}) for the patch antenna are given in Table I and shown in Fig. 6.

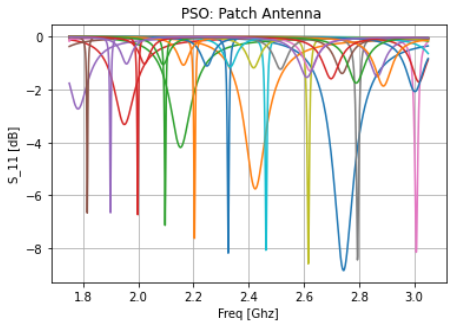


Fig.3. Initial iteration including 25 agents (I_n =25, G_m=1)

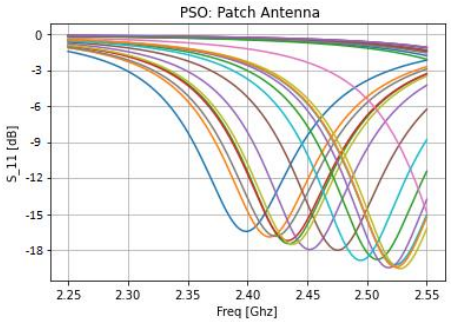


Fig.4. Third iteration including 25 agents (I_n =25, G_m=3)

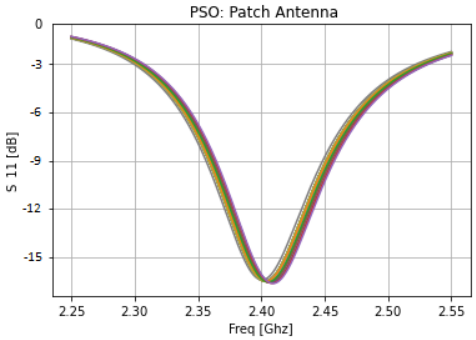


Fig.5. Eighth iteration including 25 agents (I_n =25, G_m=8)

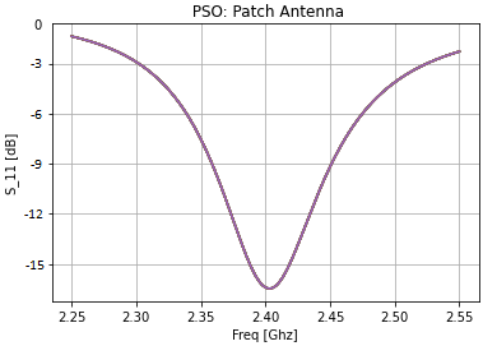


Fig.6. Tenth iteration including 25 agents (I_n =25, G_m=10)

The optimization iterations (G=1 to 8) for 25 agents (I=25) are presented in Figs. 7 to 8, and the final dimensions (g_{BEST}) for the patch antenna can be found in Table II. In this case, the optimisation process began with random initialisation, whereas in the first case (Table I), uniform initialization was used for particle positions, the input frequency was set to 2.4 GHz, and the PSO algorithm converged to this frequency. However, when PSO was applied without uniform initialization, and poorer particle positions were selected (Fig. 8).

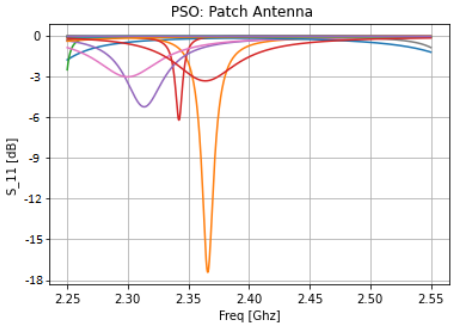


Fig.7. Initial iteration including 25 agents (I_n =25, G_m=1)

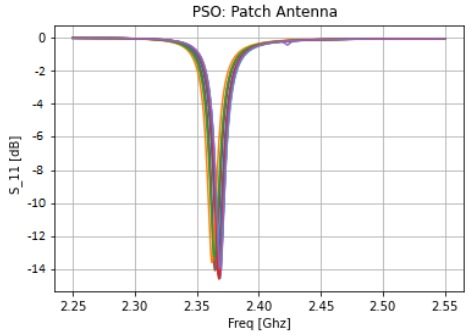


Fig.8. Eighth iteration including 25 agents (I_n =25, G_m=8)

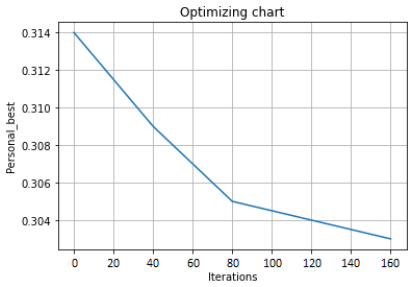


Fig.9. Optimizing chart shows performance of PSO (I_n =25, G_m=8)

Fig. 9 shows the convergence plot, where the y-axis represents the pBEST as a function of the number of iterations.

IV. Verification of PSO Functionality Using Fabricated Patch Antenna

The optimisation performance is verified and compared with a fabricated antenna. The resonance of the fabricated patch antenna was measured at 1.92 GHz, which was then set as the input frequency for the PSO code in Python. The initial values of I = 12 and G = 6 were used, and after 23 minutes, the parameters w, L, and off_y were determined (see Table 3):

TABLE III. Comparasion

Dimension	PSO	Fabricated
w [mm]	41.190	40.40
L [mm]	50.910	50.70
off_y [mm]	9.740	7.80
f_{res} [GHz]	1.918	1.919
Δf [MHz]	65	33

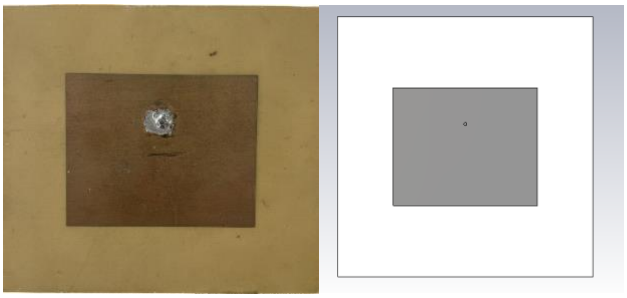


Fig.10. Comparison of fabricated antenna (on the right) with antenna simulated in CST from PSO optimizing (on the left)

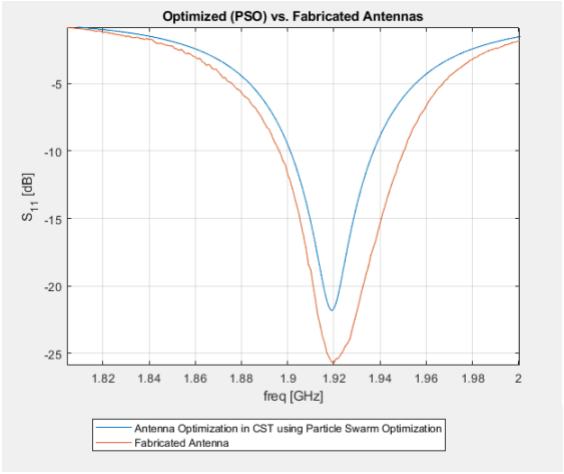


Fig.11. S11 plots comparing the results obtained from the PSO with those obtained from the fabricated model

V. Conclusion

In general, the Python connection with CST is a great tool for those working in the field of microwaves and it provides users with the capability to harness the power of Python to automate tasks, integrate with other tools, and develop customised applications that can save time and boost productivity.

This paper outlined the usage of Python with CST, and an example was presented in which the model of a canonical structure was optimised using PSO. In the final section, the optimisation performance was compared with a fabricated patch antenna model with a well-known resonant frequency, height, and permittivity of the substrate. The results were obtained with a high level of accuracy, although it was necessary to increase the number of particles for reliable results, which took more time.

All of this work will be used in the future for machine learning tasks. Without a connection between CST and Python, it would be difficult and complicated to use Python libraries such as PyTorch and TensorFlow to automate antenna designs via machine learning.

Acknowledgment

This work was supported by the Internal Grant Agency of the Brno University of Technology under project no. FEKT-S-23-8191.

References

[1] *Tera: documentations*. Tera [online]. [cit. 2023-04-02]. Dostupné z: https://tera.netlify.app/

[2] M. T. Islam, N. Misran, T. C. Take and M. Moniruzzaman, "Optimization of microstrip patch antenna using Particle swarm optimization with curve fitting," 2009 International Conference on Electrical Engineering and Informatics, Bangi, Malaysia, 2009, pp. 711-714, doi: 10.1109/ICEEI.2009.5254724.

[3] ROBINSON, J., RAHMAT-SAMII, Y. Particle swarm optimization in electromagnetics. IEEE Transactions on Antennas and Propagation, 2004, vol. 52, no. 2, p. 397–407.

[4] ROBINSON, J., RAHMAT-SAMII, Y. Particle swarm optimization in electromagnetics, IEEE Transactions on Antennas and Propagation. 2004, vol. 52, no. 2, p. 397–407.

[5] Weile, D.S.; Michielssen, E. Genetic algorithm optimization applied to electromagnetics: a review. *IEEE Trans. Antennas Prop.*, 1997, 45, 343-353. DOI: 10.1109/8.558650

[6] Design and engineering simulation: Simulia – Dassault Systèmes (2002) Design and Engineering Simulation | SIMULIA – Dassault Systèmes. Dassault Systemes. Available at: https://www.3ds.com/products-services/simulia/ (Accessed: March 11, 2023).

[7] *Visual basic (VBA) language* (2011) *VBA Macro Language*. CST Microwave Studio. Available at: http://www.mweda.com/cst/cst2013/vba/vba_macro_language_overview.htm (Accessed: March 11, 2023).

[8] Dassault Systemes Deutschland. (n.d.). *CST Python libraries*¶. CST Python Libraries - CST Python Libraries documentation. Retrieved March 11, 2023, from https://space.mit.edu/RADIO/CST_online/Python/main.html

[9] Jeanne Pages-Mounic; Alessandro De Oliveira Cabral Junior; André Barka; Hamza Kaouach; Hybrid Numerical Methodology for Efficient Design and Optimization of Transmit-Array Antennas, X-Band Application. IEEE Access, 2021, vol. 9, p. 148302-148314. DOI: 10.1109/ACCESS.2021.3124287

[10] Stevewhims (2023) *Com technical overview, Win32 apps | Microsoft Learn*. Microsoft2023. Available at: https://learn.microsoft.com/en-us/windows/win32/com/com-technical-overview#introduction-to-com (Accessed: March 11, 2023).

[11] NASCIMENTO, D.C.; LACAVA, J.C.S. Probe-fed linearlypolarized electrically-equivalent microstrip antennas on FR4 substrates, Journal of Microwaves, Optoelectronics and Electromagnetic Applications, 2014, https://doi.org/10.1590/S2179-10742014000100005

ŠKODA Kariéra



Stáže a závěrečné práce

Studuješ vysokou školu a zároveň chceš získat pracovní zkušenosti? Zkus odbornou stáž ve ŠKODA AUTO. Pod vedením našich odborníků se zapojíš do inovativních projektů – třeba těch v oblasti elektromobility. Získané znalosti pak můžeš zúročit i ve své závěrečné práci.

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 možnost spolupracovat s profesionály ve svém oboru.

Trainee program

Dokončil jsi studia a přemýšlíš kam dál? Trainee program pro absolventy ti pomůže najít obor, který tě bude opravdu bavit a naplňovat. V průběhu jednoho roku budeš na plný úvazek pracovat v mezinárodním týmu, poznáš různá oddělení a vycestuješ na zahraniční stáž.

Volné 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

Stereo Camera as a complementary sensor to LIDAR in mobile robot mapping

Jana Vyroubalová
Department of Mathematics
Brno University of Technology
Brno, Czech Republic
xvyrou04@vut.cz

***Abstract*—LIDAR (2D) has been widely used for mapping and navigation in mobile robotics. However, its usage is limited to simple environments. This problem can be solved by adding more sensors and processing these data together. This paper explores a method of how measurements from a stereo camera and LIDAR complement each other in dynamical mapping. An occupancy grid map from LIDAR data is used as a prerequisite and we describe how it could be extended by a 2D grid map from a stereo camera. This approach is based on the ground plane estimation in disparity map acquired from the stereo vision. For ground plane detection, RANSAC and Least Squares methods are used. After obstacles determination, 2D occupancy map is generated for future fusion with complementary 2D LIDAR map. Experimental results obtained from Willow Garage Stereo and Hokuyo UTM-30LX Laser are good enough to determine that this method is a benefit, although my implementation is still a prototype. In this paper, we present the applied methods, analyze the results and discuss the modifications and possible extensions to get better results.**

***Index Terms*—LIDAR, stereo camera, mapping, least squares plane, sensors fusion, disparity map, ground plane detection**

I. Introduction

Environment mapping is one of the most important parts of mobile robotics. It is the basis for localization and navigation - especially in real time in an unknown environment. There is plenty of research about mapping by 2D LIDAR and stereo camera separately (not only because of their frequency of use). LIDAR provides very accurate information and its frequency is high. However, this sensor provides information only in one plane. For example, a LIDAR is able to see the legs of a table in the right position but not the table surface which presents an obstacle to the robot. On the other hand, a stereo camera can provide 3D structural information with the color and texture of any object. However, the camera is less accurate compared to a laser range detector and the range of view is thus limited.

The problem we address in this paper is primarily which data from LIDAR and stereo camera should be fused and next how to fuze the data to get any new or better information for mapping than just from the data of one sensor only. Furthermore, we explore which methods we should use to achieve the fusion with a focus on stereo camera (shortly stereo) data processing. The objective of this work is to find out if our method can be useful. We do not provide the best possible results of this approach. Of course, we can achieve better results with a superior camera and with more suitable data.

There are some presented techniques which could solve some parts of our problem. For example, there is a 3D object detection technique presented in [1]. However, this approach is performed by using three cameras and they do not focus on environment mapping, but only on object detection. In these works [2], [3] there are obstacles and road detected. Their method is based on ”v-disparity”. However, this method will fail in an environment where obstacles occupy most part of the image (mostly indoors) or in an outdoor environment where, in the presence of a slope variation, the 2D line representing the ground plane (road) is not a straight line. There is also paper [2], which prefers the transformation of disparity map to 3D point cloud and then uses depth information to extract ground plane and obstacles before raw extraction from disparity map [3], [4]. However, working with a 3D point cloud is not reliable due to the nonlinear transformation involved in range estimation, where the inaccurate parameters of the camera are used.

Our method is based on ground plane detection in a disparity map from a stereo camera. Then we determine obstacles in the environment and project this scene into a 3D point cloud, which is used for 2D map construction. Finally, we complement the 2D map from LIDAR with this 2D map from the stereo. Basic information and problems about both sensors used in mapping are presented in Section II. Details about our approach in stereo mapping based on ground plane detection are described in Section III. In Section IV, we present the idea of data fusion and finally, we show and discuss the results of our approach in Section V. All the conclusions are summarized in Section VI.

II. Sensors in mapping

A LIDAR and a camera are the most frequently used sensors in mobile robotics, although a laser range finder is relatively expensive. Conversion from data to information is on high level.

A. LIDAR

A 2D LIDAR is a non-contact optical sensor, which uses laser beams to scan an environment in a two-dimensional area. It scans from small angles to full 360 degrees and a scanning

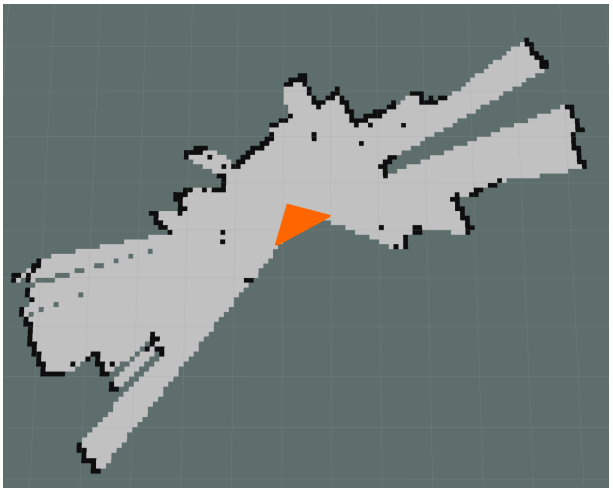


Fig. 1. Occupancy grid map with robot pose (the orange triangle).

frequency can be from 5 Hz to 50 Hz for example of the most used LIDAR sensors.

For many years, the most used approach to get a map for localization and navigation from LIDAR is *Occupancy grid mapping* (see [5] for a specific useful model of occupancy grid). This method has been investigated for a long time and therefore we use it as a prerequisite for our method. For a better idea, see Figure 1. The map is a composition of small cells - you can see the cells as rasterization or pixelization. The value of each cell represents the probability of whether the cell is occupied by an obstacle or not (in grayscale - black means 100% occupied, white means free). The green cells represent an unknown environment. As we can see, in the area in front of the robot in Figure 1 there are many objects and on the back side, the area is not investigated. For details about implementation of this method, see [6]–[8].

The main disadvantage of LIDAR mapping is that the sensor sees only in two dimensions, so some obstacles may not be detected by it. However, this sensor is very accurate, therefore this is the basis for localization and navigation. Due to these advantages, we try to extend this method with information from stereo camera.

B. Stereo camera

A stereo camera provides important information about the scene: color, texture, intensity and mainly 3D structure. Moreover, the camera is able to see behind the obstacles. However, the narrow field of view and limited range are disadvantages.

If we want to work with stereo camera properly, we need to calibrate the camera to get the camera intrinsic and extrinsic parameters to repair the images (so our images are no longer deformed). Based on the parameters from camera calibration and epipolar geometry, we are able to *rectify* the stereo images and find the pixel's correspondences [9].

Finally, we can compute the *binocular disparity* between two points, which we need for information about depth. This

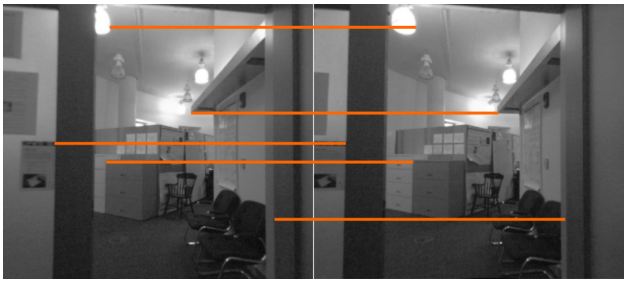


Fig. 2. Pixel correspondences in left and right rectified images.

principle is based on a difference between the indexes of corresponding pixels (see Figure 2):

$$d = (x_l - x_{cl}) - (x_r - x_{cr}), \tag{1}$$

where $[x_l, y_l]$ and $[x_r, y_r]$ are the corresponding points (endpoints of each line in Figure 2) in left and right images. Coordinate y is the same in pair and x_{cl} and x_{cr} are indexes of the center in the left and right images.

The 3D scene is based on a disparity map, from which the scene is backward reconstructed. For this backward projection into the camera frame, we need the parameters from camera calibration:

$$z = \frac{fB}{d}, \tag{2}$$

$$x = \frac{(u - u_c)z}{f}, \tag{3}$$

$$y = \frac{(v - v_c)z}{f}, \tag{4}$$

where f is the focal length, B is the baseline, $[u, v]$ is the computed pixel position (index) in the disparity map and $[u_c, v_c]$ is the position of a pixel in the center of the image. For more details see [9], [10].

We can see that there are many steps, where the stereo camera cumulates software accuracy errors:

1) estimation of calibration parameters,
2) images rectification,
3) finding the pixel correspondences for disparity algorithm,
4) backward projection from disparity map to 3D coordinates.

Notice that from the principle of disparity computation (Figure 2 and equation 1), there can be a problem with large homogeneous regions.

III. Stereo Mapping

We work with ground robots, so we consider robot motion in two dimensions. Although the stereo camera provides 3D information, we transform it into 2D grid map similar to the LIDAR map in Section II-A. This map complements the LIDAR map (both in 2D), which is a suitable approach for localization and navigation for ground robots.

In our approach, we determine obstacles and free regions from the disparity map directly - notice that the backward projection from 2D perspective to 3D orthogonal system is on the end of the chain. The scheme of our method is depicted in Figure 3:

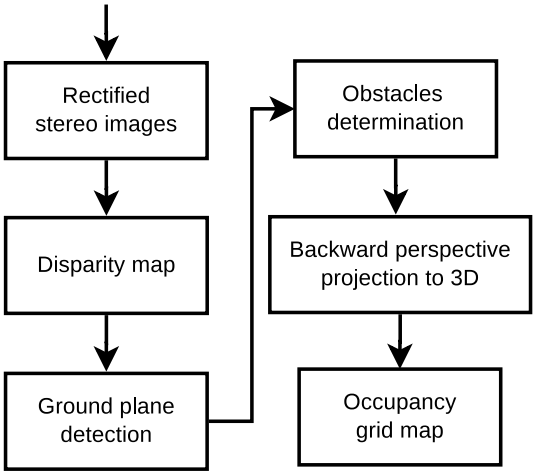


Fig. 3. Scheme of creating the map from stereo camera.

A. Disparity map

A disparity map allows us to work with raw data from the camera sensor, obviating all needless transformations based on some estimated parameters. We recommend some preprocessing such as smoothing images by the bilateral filter. Then pure disparity computation is not enough, because we lose information about exact edge and the objects are incomplete. We can solve this problem by post-filtering of raw disparity map on base of *Weighted Least Squares* method [11].

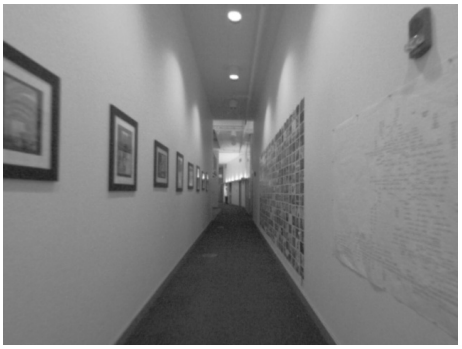


Fig. 4. Source image for filtered disparity.

Our result of the disparity map is shown in Figure 5 which was computed from source Figure 4. Some information can be better seen in Figure 6 or Figure 7. The disparity of the ground plane looks relatively fine, although we notice the fact that this ground plane is a homogeneous area. There are some papers [12], which explore disparity map in homogeneous areas. We can also upgrade the ground plane disparity map by image segmentation or hypothesis about the ground plane such as



Fig. 5. Example of our filtered disparity map.

flatness and parallelism. However, it is not the aim of this paper.

B. Ground plane detection

We use a geometric model of the plane to find the ground plane in front of the robot. We could simply assume that the ground plane is of constant height. It can be useful for indoor applications, where there is no change in the position of the ground plane relative to the camera. However, the dynamic approach (which we use) works in more cases. The geometric definition of the plane is:

$$ax + by + cz = d, \tag{5}$$

where $n = (a, b, c)$ is normal vector. Distance R_i of point i (coordinates of i are $[x_i, y_i, z_i]$) from the plane (5) is

$$R_i = ax_i + by_i + cz_i - d, \tag{6}$$

assuming that the plane is normalized, so

$$a^2 + b^2 + c^2 = 1. \tag{7}$$

To get the concrete definition of our plane, we need for example 3 points, which fit the plane. For this purpose, *Random Sample Consensus* (RANSAC) [13] is an effective technique developed from within the computer vision community. Basically, we iteratively and randomly generate 3 points in a region of interest, then fit a plane and compute the ratio of inliers and outliers. The ground plane mostly covers the bottom half of an image, so we can use only this part as the region of interest for better efficiency.

Then we do *Least Squares Plane* (LSQP) fitting to find the best solution. We need to minimize Q:

$$Q = \sum_{n=1}^{N} R_i^2, \tag{8}$$

where N is the set of inliers from RANSAC. Using (6) we can rewrite (8) as:

$$Q = \sum_{n=1}^{N} (ax_i + by_i + cz_i - d)^2. \tag{9}$$

Then from (9):

$$\frac{\partial Q}{\partial a} = \sum_{n=1}^{N} 2x_i(ax_i + by_i + cz_i - d) = 0, \qquad (10)$$

and similarly the equations $\frac{\partial Q}{\partial b} = \frac{\partial Q}{\partial c} = \frac{\partial Q}{\partial d} = 0$. To avoid the trivial solution $a = b = c = 0$, we define the condition (7). With this condition, the solution of these equations becomes *Eigenvalue problem*. You can find more information about this LSQP approach in [14]–[16].

Our result of the ground plane detection is shown in Figure 6 and Figure 7. The green points (the biggest area) are the disparity map in 3D space. The three red points define the plane. The black plane which fits the red points is our ground plane. We can also see the tilt of the black plane, which copies the decreasing disparity values to the center of the image.

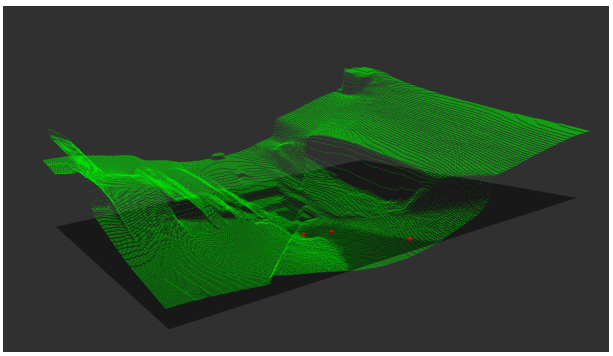


Fig. 6. Disparity map visualized in 3D space with detected ground plane.

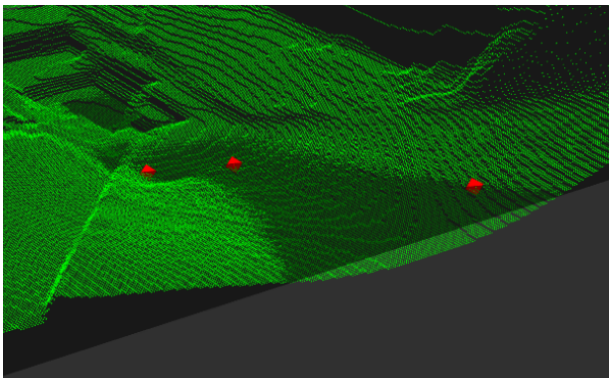


Fig. 7. Detail of ground plane from Figure 6.

C. Obstacles determination

For determination between obstacles and free areas for robot movement, we use a simple decision rule:

$$p_d(x,y) = \begin{cases} free, \Delta d < low \\ occupied, low \leq \Delta d \leq high \\ free, \Delta d > high \end{cases} \qquad (11)$$

where $p_d(x,y)$ is the pixel in disparity map, Δd is the difference between estimated disparity (from the ground plane) and measured disparity, and low and $high$ are two thresholds which are obtained by our experiments. Although the lower part of an obstacle can be resolved as free, it does not affect our mapping, because the lower part is covered by upper part in the grid map (see Section III-D).

You can see our result in Figure 8. The green area is free and the red is occupied. Notice that the estimator determined the area at the bottom of the walls as free. The reason is that we have to set the $high$ parameter high to detect the whole ground plane because of inaccuracy in the disparity map.

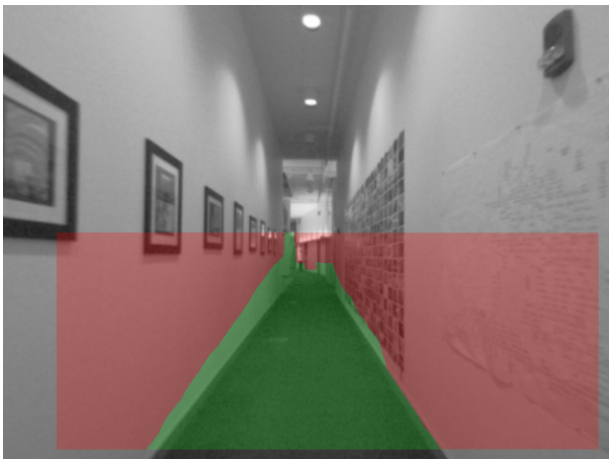


Fig. 8. Result of determination between free and occupied area.

D. Occupancy grid map

To represent the result in the occupancy grid map, we need to transform our information from perspective (disparity map) to orthogonal 3D. For this purpose, we use the equations mentioned in Section II. When we have orthogonal coordinates of points, we can project these into the 2D space. This is the step, where the wrong-determined bottom parts of obstacles are covered properly by the top parts.

We do not discuss the occupancy probabilities of cells in the map, because it is not the aim of this paper. We simplify our approach only to 2 values in the grid - occupied and unknown (see Section V), because this is enough for our research. Moreover, the omission of free cells from stereo camera is justified by its inaccuracy. For a complete implementation of the correct disparity map see Section II-A or directly [6].

IV. Data fusion

The aim of our fusion is to bring new information to mapping, which is for example used in path planning. Because of the great accuracy of LIDAR (see II-A), the localization should be computed mainly from LIDAR map. To include the camera in localization, the priority of this sensor should be set to the right value. On the other hand, the path planner definitely benefits from having information from our approach.

We do not include the information about obstacles from camera to localization (for this purpose, we can use something like camera SLAM [17]). However, we recommend using this method to upgrade the path planner - it should know about

obstacles, which are invisible for LIDAR, although the camera is a little bit inaccurate. Advanced implementation of the path planning, such as *Vector Field Histogram* [18] based on a local cost map, can be used.

We implemented only simple fusion, which is suitable for our approach. The occupied cells from camera are mapped into the LIDAR occupancy grid map. We use something like OR operation:

$$M_i = \begin{cases} free, (L_i = free) \wedge (C_i = unknown) \\ unknown, (L_i = unknown) \wedge (C_i = unknown) \\ occupied, (L_i = occupied) \vee (C_i = occupied) \end{cases} \quad (12)$$

where M_i is cell in the fused map M on index i, L_i is cell in the LIDAR map on index i and finally C_i is cell in the stereo camera map on index i. This map is then given to path planner.

However, for better results, these two sensors should be rigorously calibrated each other, moreover one can say that there exist more sophisticated methods to fuse this information together. But that goes beyond the scope of this work

V. Implementation and Experiments

We implemented our method in C++ with *Robotic Operation System* (ROS) framework [19]. Especially for LIDAR mapping, we used the GMapping module from ROS. To process some graphical calculations, the *OpenCV* library [20] was used. The complexity of our implementation is

$$O(n^2), \quad (13)$$

where n is the input image resolution. The most critical section is a computation of the disparity map. However, some optimizations for finding disparity exists, but it is not the topic of this work.

Experimental results were obtained from MIT Stata Center Data Set [21], where they have Hokuyo UTM-30LX laser range finder and Willow Garage Stereo camera. Partial results (of each significant step) are shown in this paper in the appropriate section. Our final results are shown in the next several figures, where are illustrated different situations. In every example, we show a source image and next we show corresponding occupancy maps from LIDAR sensor (on the left side) and stereo camera sensor (on the right).

As is shown in Figure 9, the laser scanner is not able to distinguish the chair (in the orange ellipse in the source image), because its height is more than the plane scanned by the LIDAR. LIDAR is able to see only the chair leg, which is much smaller than the real obstacle for a mobile robot. The camera detected this chair, although stereo vision data is noisy and inaccurate in the range measurement, especially over a large distance.

The appropriate example of the problem with large distance detection by stereo camera is shown in Figure 10. There we can see a narrow and straight passage. The length of this passage is more than 25 meters. The LIDAR is able to see the whole range without any problem. However, the stereo

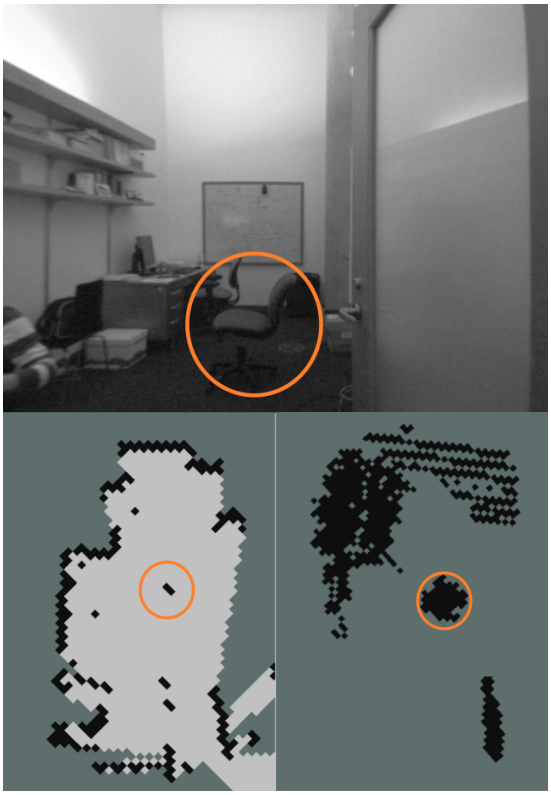


Fig. 9. Chair scene - source image on the top, LIDAR map at the bottom left, camera map at the bottom right.

camera is lost in this range - it sees very inaccurately (the orange ellipse). The reason is evident from Figure 5, where the disparity map is wrong in the rear part of the passage.

Figure 11 shows a situation, where are many chairs with thin and shiny legs around the table. LIDAR is only able to see table legs. On the opposite side, the camera sees all the chairs with a table (the orange ellipse). However, we can see a failure of the camera detection in the red ellipse. The reason for this failure is similar to Figure 7. The disparity map of this ground plane is cambered (because it is a large homogeneous area), therefore the ground plane detection is a little bit inaccurate (especially at the wrong tilt). So there is a possibility to detect part of the ground plane as an obstacle.

The last shown experiment is in Figure 12, which is a great illustration of an invisible obstacle for LIDAR, but the stereo camera knows about this obstacle. A vision of LIDAR is overshadowed.

Based on all these experiments, we can see many facts about our approach. Camera vision is very noisy. However, with the LIDAR background, it can be helpful in many situations. The problem with inaccurate disparity map occurs very often, which can lead to the situation in Figure 11. We can eliminate it by improving the disparity map extraction, which is not a simple topic for homogeneous areas. We have already opened this theme in Section III. Moreover, data from a camera in a large distance is so sparse. The reason is that the perspective projection shrinks real information into a small

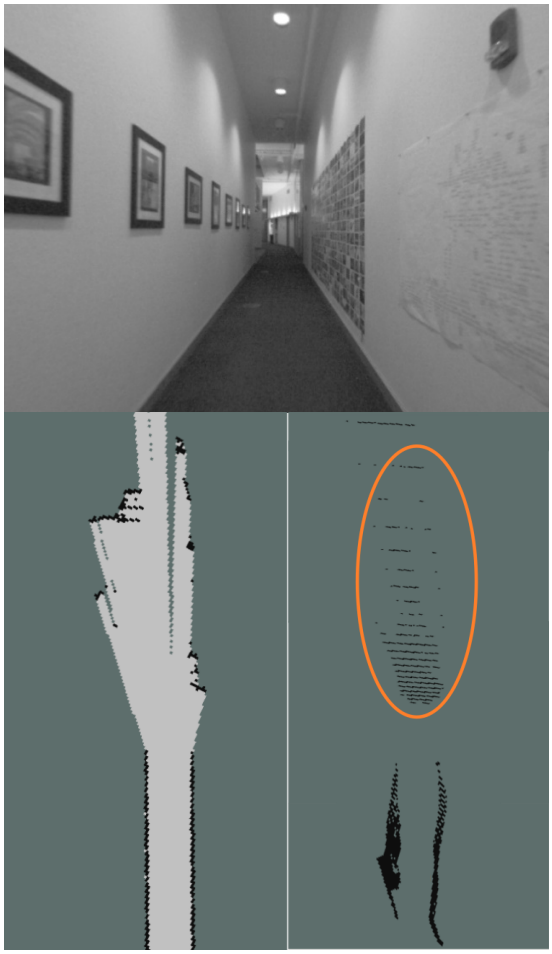


Fig. 10. Passage scene - source image on the top, LIDAR map at the bottom left, camera map at the bottom right.

part of the camera image - so there are only few pixels with information in large distances, which are reprojected back into 3D orthogonal system.

VI. Conclusions

In this paper, we have presented the method for LIDAR and stereo camera data fusion to reach better mapping and navigating. Particularly, we focused on new information extraction from stereo camera compared to LIDAR mapping. This approach includes techniques such as the Least Square (Plane) method, Random Sample Consensus, and extraction of disparity map from raw stereo images.

We evaluated every step of this method and we find out, that there are big reserves in disparity map detection. However as our tests show, we reached what we wanted to - we shew that this approach is useful and it brings new information to the mapping of mobile robots. Also we discussed the modifications of this method leading to better results.

One possible way to further improve the accuracy of the presented method is to investigate the side of disparity map extraction because we did not focus on it so much. We only used some implemented method. Next, there are some topics



Fig. 11. Table scene - source image on the top, LIDAR map at the bottom left, camera map at the bottom right.

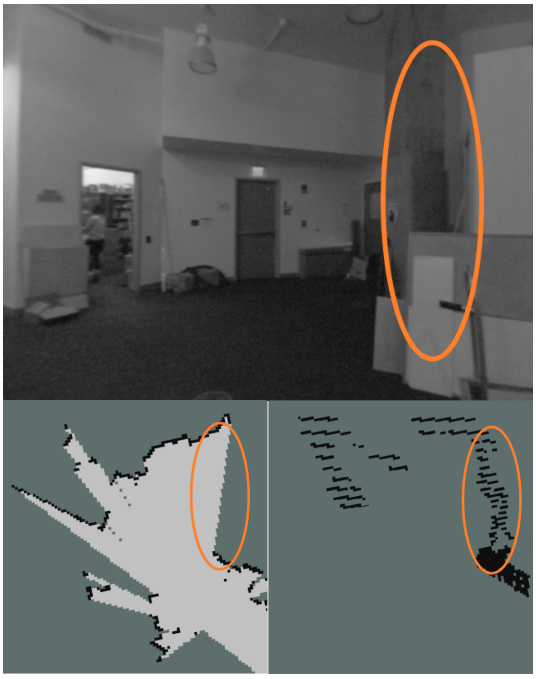


Fig. 12. Wall scene - source image on the top, LIDAR map at the bottom left, camera map at the bottom right.

to continue this method. For example, how to assign weights of the detected object from stereo camera to navigation or localization compared to the LIDAR. And next one is how to adapt the path planner to fused map [18], [22].

REFERENCES

[1] L. Romero, A. Núñez, S. Bravo, and L.E. Gamboa. *Fusing a Laser Range Finder and a Stereo Vision System to Detect Obstacles in 3D*. Advances in Artificial Intelligence – IBERAMIA 2004, 2004. ISBN: 978-3-540-30498-2, pages 555-561.

[2] L. Matthies and P. Grandjean. *Stochastic Performance Modeling and Evaluation of Obstacle Detectability with Imaging Range Sensors*. IEEE Transactions on Robotics and Automation, 1994. ISSN: 1042-296X, pages 783-792.

[3] R. Labayrade, D. Aubert, and J.-P. Tarel. *Real Time Obstacle Detection in Stereovision on Non Flat Road Geometry Through "V-disparity" Representation*. IEEE Intelligent Vehicle Symposium, 2002. Pages 646-651.

[4] F. Ferrari, E. Grosso, G. Sandini, and M. Magrassi. *A Stereo Vision System for Real Time Obstacle Avoidance in Unknown Environment*. IEEE International Workshop on Intelligent Robots and Systems, Towards a New Frontier of Applications, 1990. Pages 703-708.

[5] Š. Křehlík and J. Vyroubalová. *The Symmetry of Lower and Upper Approximations, Determined by a Cyclic Hypergroup, Applicable in Control Theory*. Symmetry, 2020. Volume 12,ISSN: 2073-8994.

[6] T.-D. Vu, O. Aycard, and N. Appenrodt. *Online Localization and Mapping with Moving Object Tracking in Dynamic Outdoor Environments*. IEEE Intelligent Vehicles Symposium, 2007. ISBN: 1-4244-1067-3, pages 190-195.

[7] G. Grisetti, C. Stachniss, and W. Burgard. *Improved Techniques for Grid Mapping with Rao-Blackwellized Particle Filters*. IEEE Transactions on Robotics, 2007. Volume 23, pages 34-46.

[8] G. Grisetti, C. Stachniss, and W. Burgard. *Improving Grid-based SLAM with Rao-Blackwellized Particle Filters by Adaptive Proposals and Selective Resampling*. In Proc. of the IEEE International Conference on Robotics and Automation (ICRA), 2005. ISBN: 0-7803-8914-X, pages 2432-2437.

[9] M. S. H. Achmad, W. S. Findari, N. Q. Ann, D. Pebrianti, and M. R. Daud. *Stereo camera — Based 3D object reconstruction utilizing Semi-Global Matching Algorithm*. 2nd International Conference on Science and Technology-Computer (ICST), 2016. ISBN: 978-1-5090-4357-6, pages 194-199.

[10] L. Yan, X. Zhao, and H. Du. *Research on 3D measuring based binocular vision*. 2014 IEEE International Conference on Control Science and Systems, 2014. ISBN: 978-1-4799-6397-3, pages 18-22.

[11] D. Min, S. Choi, J. Lu, B. Ham, K. Sohn, and M. N. Do. *Fast Global Image Smoothing Based on Weighted Least Squares*. IEEE Transactions on Image Processing, 2014. ISSN: 1941-0042, pages 5638-5653.

[12] J.-E. Kim, K.-D. Kim, and K.-H. Jung. *Reliable estimation of disparity map in textureless region of roadway*. 2017 19th International Conference on Advanced Communication Technology (ICACT), 2017. ISBN: 978-89-968650-9-4, pages 399-402.

[13] M.A. Fischler and R.C. Bolles. *Random consensus: A paradigm for model fitting with applications to image analysis and automated cartography*. Communications of the ACM, 1981. Volume 24, pages 381-395.

[14] K. Pearson. *On lines and planes of closest fit to systems of points in space*. Philosophical Magazine, 1901. Volume 2, pages 559-572.

[15] V. Schomaker, J. Waser, R.E. Marsh, and G. Bergman. *To Fit a Plane or a Line to a Set of Points by Least Squares*. Acta Cryst, 1959. Volume 12, pages 600-604.

[16] D.M. Blow. *To fit a plane to a set of points by least squares*. Acta Cryst, 1960. Volume 13, page 168.

[17] A.J. Davison, I.D. Reid, N.D. Molton, and O. Stasse. *MonoSLAM: Real-Time Single Camera SLAM*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007. Volume 29(6): pages 1052-1067, ISSN: 0162-8828.

[18] J. Borenstein and Y. Koren. *The vector field histogram-fast obstacle avoidance for mobile robots*. IEEE Transactions on Robotics and Automation, 1991. Volume 7(3): pages 278 - 288, ISSN: 1042-296X.

[19] Group of authors. *ROS: Robot Operating System*. Cit. [2023-02-10] URL http://www.ros.org/.

[20] G. Bradski and A. Kaehler. *Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library*. O'Reilly Media, 2017. 1st edition, ISBN: 1491937998.

[21] M. Fallon, H. Johannsson, M. Kaess, and J. Leonard. *The MIT Stata Center Dataset*. IJRR Dataset Paper, 2013. Volume 32(14): pages 1695-1699.

[22] I. Ulrich and J. Borenstein. *VFH*: local obstacle avoidance with look-ahead verification*. IEEE International Conference on Robotics and Automation, 2000. Volume 3, pages 2505-2511.

Performance Evaluation of Post-Quantum Schemes on Mobile Platforms

1st Patrik Dobiáš
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
xdobia13@vut.cz

2nd Lukáš Malina
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
malina@vut.cz

***Abstract*—In this work, we present the performance assessment of Post-Quantum Cryptography (PQC) schemes on iOS and Android platforms using the liboqs library. The experimental setup for both platforms is outlined, with details on how liboqs was incorporated into the applications. We focus on the schemes that are chosen to be standardized by the National Institute of Standards and Technologies (NIST). Our findings indicate that Dilithium is the most suitable signature scheme for all phases, while the Falcon signature scheme is a better choice for use cases requiring a fast verification phase. With regard to Key Encapsulation Mechanisms (KEMs), the Kyber scheme outperforms other KEM schemes in all aspects. We believe that our results could be beneficial for engineers and developers who consider substituting current asymmetric schemes by quantum-resistant schemes.**

***Index Terms*—Post-Quantum Cryptography, liboqs, iOS, Android**

I. Introduction

With the advent of quantum computers, many traditional cryptographic algorithms are at risk of being broken. To mitigate this threat, there has been significant research into post-quantum cryptography, which aims to develop new cryptographic schemes that are secure against attacks by quantum computers. To support this research, NIST started a post-quantum scheme standardization process in 2016 [1].

As research continued, many publicly available implementations of the proposed schemes were made. To unify this progress, the Open Quantum safe project arose as an open-source initiative. This community developed the liboqs library, which is a collection of post-quantum cryptography algorithms selected as part of the NIST post-quantum cryptography standardization process. They aim to use a unified API interface to easily incorporate well-known cryptographic libraries such as OpenSSL [2].

In this paper, we evaluate the performance of the post-quantum schemes based on the liboqs library on mobile platforms. We focus on several key metrics, including encryption and decryption speed, key generation speed, and memory usage. Our results demonstrate the feasibility of using post-quantum cryptography on mobile devices and provide insights into the trade-offs between security and performance on these platforms.

A. Related Work

In the last few years, several works have already dealt with benchmarks and assessments of PQC. For example, Malina *et al.* [3] investigated the feasibility of implementing post-quantum cryptography on resource-constrained devices such as those used for mobile systems and Internet of Things (IoT) networks. They compared the performance of different post-quantum key exchange schemes and their message sizes across selected ARM devices using implementations from the Open Quantum Safe project. The study concluded that lattice-based schemes are the most promising due to their efficiency and relatively reasonable key sizes.

Paquin *et al.* [4] benchmarked the performance of post-quantum key exchange algorithms under realistic network conditions using the networking features of the Linux kernel. Their study focused on the impact of post-quantum primitives on TLS connection establishment, specifically hybrid elliptic curve/post-quantum key exchange and post-quantum digital signatures based on implementations from the Open Quantum Safe project. Their study revealed that packet loss above 3-5% starts to have a significant impact on post-quantum schemes that fragment across multiple packets. They also measured the latency of loading TLS-enabled websites, showing that network latency hides most of the impact of slower algorithms.

Another relevant work was done by Saarinen [5] who analyzed the energy requirements of post-quantum cryptography schemes on mobile and embedded systems. The study established a relation between computational costs and data transmission costs, proposing metrics and guidelines for the use of post-quantum cryptography algorithms in IoT and mobile systems. The results indicate that fast-structured lattices are a preferred option. In our work, we focus solely on mobile platforms and NIST winners/4th round schemes.

B. Contribution

The main contribution of this work is the performance evaluation of PQC schemes that are to be standardized by NIST on mobile devices with different operating systems (Android, iOS).

II. Background

The NIST standardization process for post-quantum cryptography has resulted in the selection of four schemes for standardization: three for signature generation and one for key encapsulation. In addition, three more key encapsulation schemes are currently being reviewed in the fourth round of the standardization process as potential alternatives to the selected scheme.

In the following sections, we provide a brief overview of the signature generation and key encapsulation schemes that have been selected for standardization. We also describe the Open Quantum Safe project, which provides implementations of these schemes at different security levels through the liboqs library.

A. Signature Schemes

This section provides a short description of individual digital signature schemes that have been selected for standardization by NIST. Table I compares these schemes based on their key and signature sizes. The table shows that the lattice-based schemes offers relatively compact signatures, while the hash-based SPHINCS+ scheme has smaller key sizes.

1) Dilithium [6]: is a signature scheme based on Module Learning with Errors (M-LWE) and Module Short Integer Solution (M-SIS) problems. It provides strong security guarantees and has relatively low computational time and memory complexity compared to other post-quantum signatures. Dilithium is designed to be highly flexible, allowing for a variety of security and performance trade-offs. With its strong theoretical foundation and cryptanalytic history, Dilithium is an ideal choice for a wide range of cryptographic applications. NIST has selected Dilithium as the primary signature algorithm for standardization due to its high efficiency, simple implementation, and strong security [7].

2) SPHINCS+ [8]: is a stateless hash-based signature scheme that provides security against quantum computers. It is based on the Merkle tree construction and has been designed to provide high security against both classical and quantum attacks. Unlike stateful hash-based signature schemes, that require user to keep track of state between signatures with disastrous consequences if the state is mismanaged, this scheme does not need to maintain any state information. However, it is also slower and produces larger signatures compared to stateful schemes. SPHINCS+ has a relatively high time and memory complexity compared to other post-quantum signatures [7].

3) Falcon [9]: is a post-quantum signature scheme based on the SIS problem. It provides strong security guarantees and has relatively low computational time and memory complexity compared to other post-quantum signatures. Falcon is designed to be efficient, especially in resource-constrained environments such as embedded systems and IoT devices. Despite its complexity and potential implementation challenges, it was chosen for standardization by NIST due to its strong security assurance and its narrow bandwidth, which may be essential in certain applications [7].

TABLE I
Memory requirements comparison of NIST PQC digital signature schemes.

Scheme	NIST Level	Private Key $[B]$	Public Key $[B]$	Signature $[B]$
Dilithium	2	2528	1312	2420
Dilithium	3	4000	1952	3293
Dilithium	5	4864	2592	4595
SPHINCS+	1	64	32	17088/7856
SPHINCS+	3	96	48	35664/16224
SPHINCS+	5	128	64	49856/29792
Falcon	1	1281	897	690
Falcon	5	2305	1793	1330

NIST Level 2 is equivalent to SHA3-256, level 3 is equivalent to AES192, level 5 is equivalent to AES256 [10].

B. Key Encapsulation Schemes

One key encapsulation scheme was chosen for standardization, which relies on lattices, while three others are currently in their fourth round of selection and are based on code. Table II provides a comparison of these schemes in terms of their key and ciphertext sizes. As can be seen in the table, the lattice-based scheme offers a smaller key and ciphertext size.

1) Kyber [11]: is a lattice-based post-quantum encryption scheme that is designed for use in key encapsulation mechanisms. It is built on the M-LWE problem and offers high flexibility with regard to security, allowing for changes in parameters without affecting the underlying operations. Kyber has excellent performance in software, hardware, and many hybrid settings [7].

2) Classic McEliece [12]: is a KEM scheme that is based on error-correcting codes. It has been around since the 1970s and is considered one of the oldest KEM scheme. Due to its long history, the security of the scheme has been thoroughly studied. Although the McEliece scheme is relatively simple and offers a fast key encapsulation process, its public key size is significant, making it less feasible for certain applications [7]. Specifically, its public key size is a multiple of thousands compared to other candidates, as shown in Table II.

3) BIKE [13]: is a family of post-quantum public key encryption schemes that are based on code-based cryptography. BIKE has the most competitive performance among the non-lattice-based KEM schemes. The quasi-cyclic structure of BIKE enables public key and ciphertext sizes comparable to the structured lattice KEM schemes [7].

4) HQC [14]: is a family of post-quantum public-key encryption schemes that are based on codes and follows an LWE-like encryption protocol. HQC offers strong security assurances and a mature decryption failure rate analysis. However, despite its quasi-cyclic structure allowing for reasonable sizes of public keys and ciphertexts, HQC's public keys and ciphertexts are still larger compared to other structured code- and lattice-based KEM schemes [7].

C. Open Quantum Safe

The Open Quantum Safe project is an open-source initiative founded in 2016 that focuses on the development of algorithms resistant to quantum computers. It also involves incorporating

TABLE II
MEMORY REQUIREMENTS COMPARISON OF NIST PQC KEM SCHEMES.

Scheme	NIST Level	Private Key $[B]$	Public Key $[B]$	Ciphertext $[B]$
Kyber	1	1632	800	768
Kyber	3	2400	1184	1088
Kyber	5	3168	1568	1568
McEliece	1	6452	261120	128
McEliece	3	13568	524160	188
McEliece	5	13892	1044992	240
BIKE	1	5223	1541	1573
BIKE	3	10105	3083	3115
BIKE	5	16494	5122	5144
HQC	1	2289	2249	4481
HQC	3	4562	4522	9026
HQC	5	7285	7245	14469

NIST level 1 is equivalent to AES128, level 3 is equivalent to AES192, level 5 is equivalent to AES256 [10].

its library into various cryptographic libraries and applications, such as OpenSSL. The direction of development of the project follows the resolution of NIST on post-quantum cryptography to maintain maximum security and relevance [2].

1) The liboqs library: The liboqs library is the main work of the Open Quantum Safe project, written in the C programming language. It contains both KEM schemes, such as Kyber, BIKE, HQC, etc., and digital signature algorithms, such as Dilithium, Falcon, and SPHINCS+. For Dilithium, it offers two versions, one using the SHAKE256 hash function, and the other using the AES block cipher in counter mode. The most developed implementation is SPHINCS+, which offers not only robust version options but also a choice of hash functions, including SHAKE256, SHA-256, and Haraka, a hash function specifically developed for post-quantum cryptography. The library maintains a common API for easy implementation across all algorithms [2].

The current version of the liboqs library is available on Linux, Windows, and macOS operating systems, and supports Intel, AMD, x86_64, and ARM processor architectures. Furthermore, it provides official wrappers to integrate the library into C++, .NET, C#, Java, Go, Python, and Rust programs [2].

III. EXPERIMENTAL SETUP

This section outlines the experimental setup for both iOS and Android platforms. First, the iOS implementation and integration of the liboqs library are described, followed by the integration of the liboqs library into our Android implementation. Furthermore, issues during Android integration are described.

A. iOS Setup

For iOS implementation we used the Swift programming language, development environment Xcode version 14.2 and for benchmarking of the schemes the XCTest framework. The liboqs was incorporated using CMake with toolchains for iOS and was linked to the application in the Xcode project. As Swift supports integration with C, it was pretty straightforward, and only the bridging header needed to be added, to make C functions visible for Swift code.

B. Android Setup

The implementation of the Android version took place in the .NET and C/C++ Microsoft Visual Studio 2022 Community environment, to which the .NET MAUI framework and the Android SDK software development kit package version 33.0.2 were added. The liboqs library was built using the build script provided from the Github repository. To use the liboqs library in C#, it is advisable to create a wrapper that implements its functions. For the purposes of this project, the official wrapper of the Open Quantum Safe project for the .NET framework was used with slight modifications.

IV. RESULTS AND DISCUSSION

We evaluated the performance of the liboqs library on various post-quantum schemes that are either being standardized or are in the fourth round of the standardization process. Evaluation was carried out on iOS and Android platforms, and computation time was measured for different security levels across each scheme's key generation, signing and verification phases for signature schemes, and key generation, key encapsulation and key decapsulation phases for KEM schemes. For the signature schemes, 20 runs were evaluated, and for the KEM schemes, 100 runs were evaluated. The average times are presented in Figures 1 and 2.

Performance measurements were conducted on two mobile devices: iPhone 12 mini with Apple A14 Bionic 6-core CPU, 3.1 GHz and Samsung A53 with Exynos 1280 8-core CPU, 2 - 2,4 GHz.

Figure 1 shows the results of our performance evaluation of digital signature schemes. We can see from the results that Dilithium outperforms other schemes in key generation and signing phases, whereas Falcon is significantly faster during the verification phase. Also, as shown in Table I, Falcon has a smaller signature size and public key than Dilithium, making it a better choice for constrained devices where only verification would be required.

Specifically, Dillithium 2 AES demonstrated the fastest computation times with 96 μs for key generation, 302 μs for signing, and 65 μs for verification on iOS. In contrast, SPHINCS+ Haraka 192s exhibited the longest computation times with 173,576 μs for key generation, 1,764,918 μs for signing, and 1,727 μs for verification on iOS.

Figure 2 shows the results of the performance measurements for the KEM schemes. The performance of each post-quantum scheme and mobile device was compared, with Kyber demonstrating superior performance compared to all candidate schemes from the fourth round. Furthermore, all operations were faster on both iOS and Android platforms for the same level of security.

Specifically, Kyber 512 demonstrated the fastest computation times with 24 μs for key generation, 30 μs for key encapsulation, and 34 μs for key decapsulation on iOS. In contrast, McEliece 8192128f exhibited the longest computation times with 298,277 μs for key generation, 476 μs for key encapsulation, and 330 μs for key decapsulation on iOS.

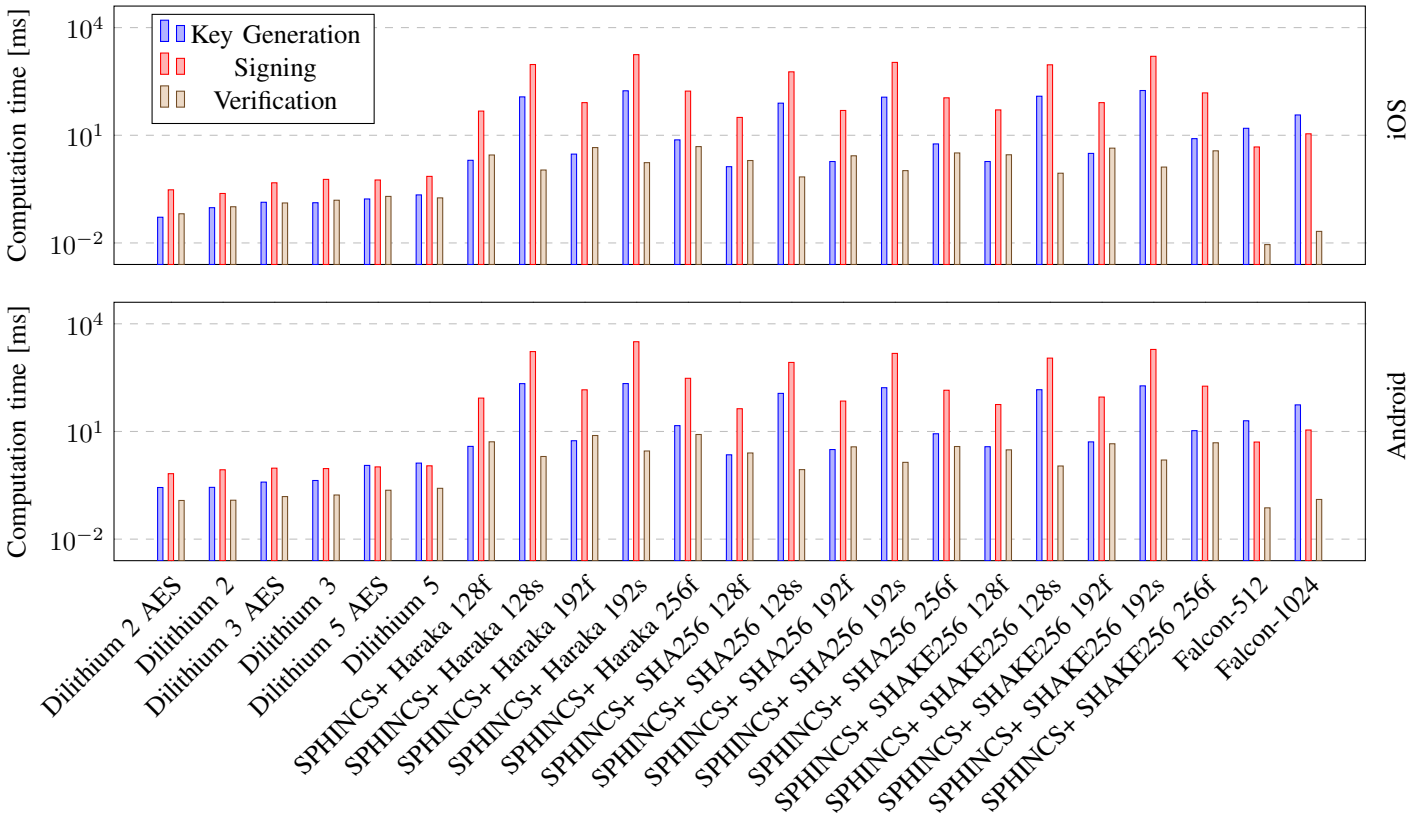


Fig. 1. Computation times of signature schemes on iOS and Android

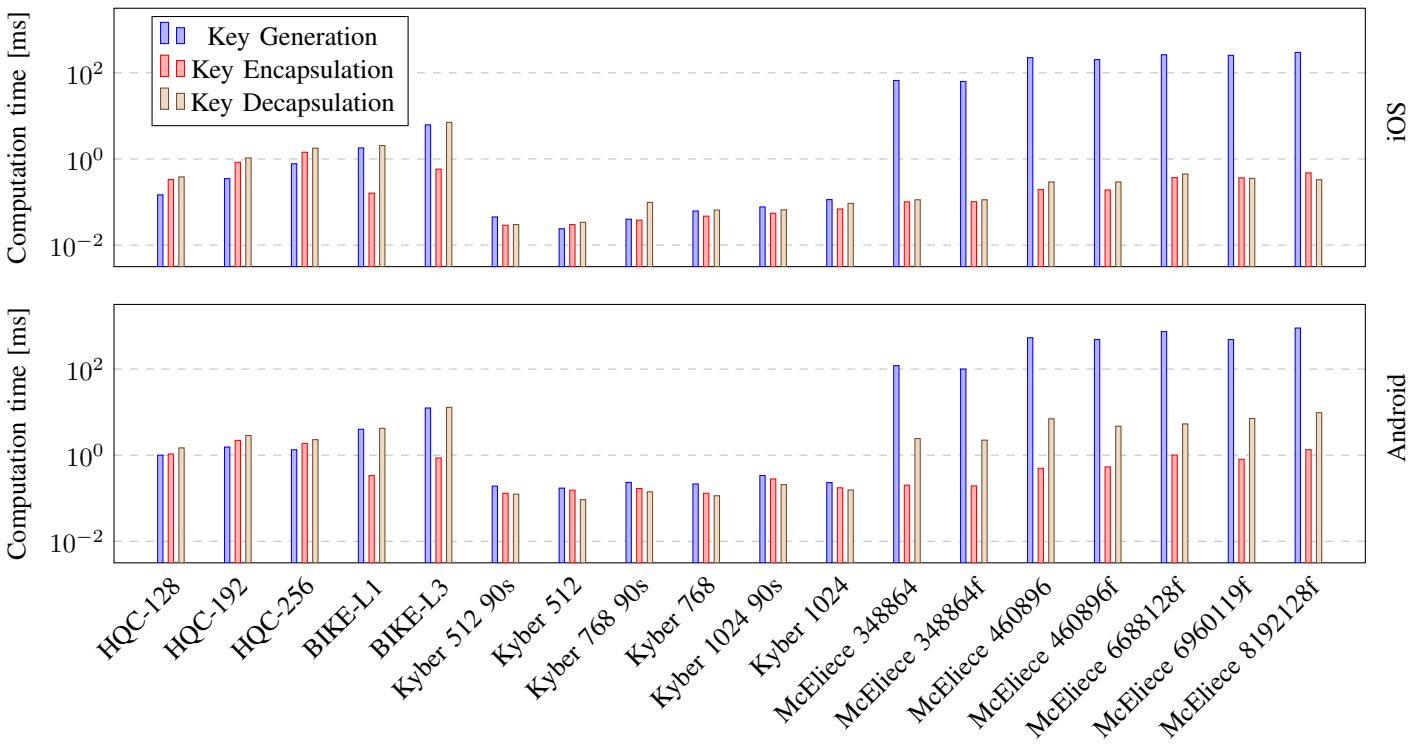


Fig. 2. Computation times of KEM schemes on iOS and Android

V. Conclusion

This article presented the performance evaluation of PQC schemes on mobile platforms. The schemes were measured on the iOS and Android platforms. The results indicated that lattice-based schemes demonstrated superior performance compared to other schemes. For the signature schemes, Dilithium showed signing times in the range of hundreds of μs, while SPHINCS+ took about a thousand times more. The verification times for Falcon were tens of μs, while SPHINCS+ took a hundred times longer. In terms of key encapsulation and decapsulation for the KEM schemes, Kyber showed the lowest times, taking only tens of μs, while BIKE took up to units of ms. The key generation phase demonstrated a significant difference from Classic McEliece showing times that were about a thousand times longer than Kyber.

References

[1] L. Chen, L. Chen, S. Jordan, Y.-K. Liu, D. Moody, R. Peralta, R. A. Perlner, and D. Smith-Tone, *Report on post-quantum cryptography*. US Department of Commerce, National Institute of Standards and Technology, 2016, vol. 12.

[2] D. Stebila and M. Mosca, "Post-quantum key exchange for the internet and the open quantum safe pr oject," in *Selected Areas in Cryptography – SAC 2016*, R. Avanzi and H. Heys, Eds. Cham: Springer International Publishing, 2017, pp. 14–37.

[3] L. Malina, L. Popelova, P. Dzurenda, J. Hajny, and Z. Martinasek, "On feasibility of post-quantum cryptography on small devices," *IFAC-PapersOnLine*, vol. 51, no. 6, pp. 462–467, 2018.

[4] C. Paquin, D. Stebila, and G. Tamvada, "Benchmarking post-quantum cryptography in tls," in *Post-Quantum Cryptography: 11th International Conference, PQCrypto 2020, Paris, France, April 15–17, 2020, Proceedings 11*. Springer, 2020, pp. 72–91.

[5] M.-J. O. Saarinen, "Mobile energy requirements of the upcoming nist post-quantum cryptography standards," in *2020 8th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud)*. IEEE, 2020, pp. 23–30.

[6] L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, P. Schwabe, G. Seiler, and D. Stehlé, "Crystals-dilithium: A lattice-based digital signature scheme," *IACR Transactions on Cryptographic Hardware and Embedded Systems*, vol. 2018, no. 1, p. 238–268, Feb. 2018. [Online]. Available: https://tches.iacr.org/index.php/TCHES/article/view/839

[7] G. Alagic, D. Apon, D. Cooper, Q. Dang, T. Dang, J. Kelsey, J. Lichtinger, C. Miller, D. Moody, R. Peralta *et al.*, "Status report on the third round of the nist post-quantum cryptography standardization process," *US Department of Commerce, NIST*, 2022.

[8] D. J. Bernstein, D. Hopwood, A. Hülsing, T. Lange, R. Niederhagen, L. Papachristodoulou, M. Schneider, P. Schwabe, and Z. Wilcox-O'Hearn, "Sphincs: practical stateless hash-based signatures," in *Advances in Cryptology–EUROCRYPT 2015: 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part I 34*. Springer, 2015, pp. 368–397.

[9] P.-A. Fouque, J. Hoffstein, P. Kirchner, V. Lyubashevsky, T. Pornin, T. Prest, T. Ricosset, G. Seiler, W. Whyte, Z. Zhang *et al.*, "Falcon: Fast-fourier lattice-based compact signatures over ntru," *Submission to the NIST's post-quantum cryptography standardization process*, vol. 36, no. 5, 2018.

[10] National Institute of Standards and Technology, "Post-quantum cryptography," 2017, accessed: 2023-03-30. [Online]. Available: https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization/evaluation-criteria/security-(evaluation-criteria)

[11] J. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. M. Schanck, P. Schwabe, G. Seiler, and D. Stehle, "Crystals - kyber: A cca-secure module-lattice-based kem," in *2018 IEEE European Symposium on Security and Privacy (EuroS&P)*, 2018, pp. 353–367.

[12] T. Chou, C. Cid, S. UiB, J. Gilcher, T. Lange, V. Maram, R. Misoczki, R. Niederhagen, K. Paterson, and E. Persichetti, "Classic mceliece: conservative code-based cryptography, 10 october 2020," 2020.

[13] N. Aragon, P. S. Barreto, S. Bettaieb, L. Bidoux, O. Blazy, J.-C. Deneuville, P. Gaborit, S. Gueron, T. Guneysu, C. A. Melchor *et al.*, "Bike: bit flipping key encapsulation," 2017.

[14] C. A. Melchor, N. Aragon, S. Bettaieb, L. Bidoux, O. Blazy, J.-C. Deneuville, P. Gaborit, E. Persichetti, G. Zémor, and I. Bourges, "Hamming quasi-cyclic (hqc)," *NIST PQC Round*, vol. 2, no. 4, p. 13, 2018.

Towards 5G Networks Optimization through Machine Learning-aided Processing of Large Datasets

1st Elham Younesian
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
elham.younesian@vut.cz

2nd Jiri Hosek
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
hosek@vut.cz

***Abstract*—This paper presents a novel approach for optimizing 5G networks through the use of machine learning techniques for processing large datasets. With the increasing demand for high-speed and low-latency communications, 5G networks have emerged as a promising solution. However, the optimization of these networks remains a challenge due to the complexity of the systems and the vast amount of data generated by them. To address this issue, we propose a framework that utilizes data preprocessing in time series data to analyze large datasets and extract useful information. It can be used to optimize various aspects of the 5G network. Specifically, we present results of implementing seasonal decomposition of time series which is one of the techniques in time series forecasting. This preprocessed data will be implemented to time series machine learning models to predict the future delay between routers and finally propose the best available route. Overall, our work presents a valuable contribution to the field of 5G network optimization and highlights the potential of machine learning techniques in this domain.**

***Index Terms*—Machine learning, Time series forecasting, Data preprocessing, Beyond 5G networks, Network slicing**

I. Introduction

The rapid growth in cellular technologies is expected to provide unprecedented levels of network's capability and efficiency. These improvements lead to a more complex network and also require more resources to provide better services for different devices and various applications. However, lack of application awareness in the networks results in dissatisfaction of the end-users. Application-aware routing which implements different kinds of routing policies corresponding to differentiated applications, can overcome this drawback [1]. Network slicing is a method of creating multiple unique logical and virtualized networks over a common multi-domain infrastructure. It is one the main technologies in 5G networks that enables multiple networks to be created on top of common physical infrastructures.

Next-generation networks are expected to learn the diverse characteristics of the user's ambience as well as human behavior, in order to autonomously determine the optimal system configurations. These smart networks have to rely on sophisticated learning and decision-making. Machine Learning (ML), as one of the most powerful artificial intelligence tools, constitutes a promising solution [2]. ML has been applied to various mundane and complex problems arising in network operation and management.

Data preprocessing is a crucial step in ML, and it plays an essential role in building accurate and reliable models. It helps to improve the quality of the data by removing irrelevant, incomplete, or noisy data. Data with inconsistencies or errors can affect the accuracy of machine learning models, and preprocessing can help to eliminate these issues. In many cases, data may be collected from multiple sources and may need to be integrated into a single dataset. Data preprocessing can help to integrate the data from various sources, which can make it easier to build accurate models.

This work models the data traffic of a proposed topology which is based on the data collected from one macro cell in a residential area that covers 53,000 users each day. These data is being implemented for an intelligent network slicing in 5G beyond networks. We visualized the average delay between different routers in a network that contains 6 routers to analyse the behaviour of the data for further implementation of time series forecasting. It will help to predict the delay between different routers and predict the best available route for different applications in next-generation networks.

The remainder of this paper is structured as follows. Section 2 provides background of this work. Section 3 describes the process of calculating delay using time series forecasting including a description about our dataset, how it was preprocessed and visualized. Section 4 discusses about the obtained results. Finally, Section 5 summarizes and concludes this paper and mentions the future works.

II. Background

Within the new-generation mobile networks, the amount of traffic rises compared to the previous generation (4G) wireless systems according to mobile end-user data speed and growth [3]. Network capability needs to satisfy the diverse requirements of various services for the fully connected smart society, such as enhanced mobile broadband (eMBB), massive

machine type communication (mMTC), and ultra-reliable and low latency communication (URLLC). Since each of these services requires different types of network capabilities, they cannot be provided effectively over a single homogeneous network. Therefore, 5G networks are being enabled with the capability of the on-demand constructions of network slices with sufficient resources by using network function virtualization (NFV) and software defined networking (SDN) technologies. A network slice is configured and controlled dynamically by software, which is also called network softwarization. The network slice is a logically separated, self-contained, independent and secured part of the network, targeting different services with different requirements on speed, latency and reliability [4].

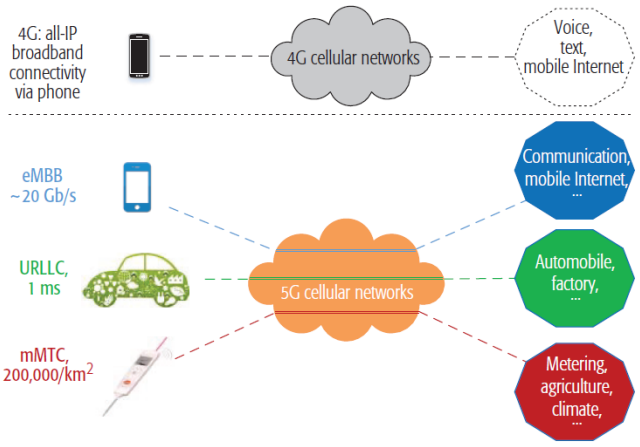


Fig. 1: 5G cellular networks: a key enabler to all mobile devices across all industries.

As depicted in Fig. 1 [4], due to the development of smarter 5G networks, it will be feasible to provide customized end-to-end network slices (NS) [5] to simultaneously satisfy distinct service requirements, such as ultra-low latency in URLLC and ultra-high throughput in eMBB [4].

In this contex, Cianfrani et al. [6] discussed about the Traffic Engineering (TE) which has the goal to facilitate efficient and reliable network operations while simultaneously optimizing network resource utilization and traffic performance. In addition, the authors in [6] discussed about the drawbacks of End-to-End Explicit (E2E) routing solutions which are the complexity of the control plane and their scalability. Thus, Segment Routing (SR) paradigm has been defined which is fully compatible with both IP and multiprotocol label switching (MPLS) data planes. Its main feature is the possibility to make use of available shortest paths, referred to as Segments, to create unconstrained network paths, which are represented by the so called Segment Lists.

Moreover, Van Tong et al. [1], discussed how the lack of application-awareness results in dissatisfaction of the end-users. They introduced Application-aware routing which implements different kinds of routing policies corresponding to differentiated applications, can overcome this drawback. Therefore, many network operators (e.g., NTT, Vodafone, etc.) implement segment routing (SR) in their network infrastructure as a solution [7], [8]. The core idea of SR architecture is based on the notion of source routing [9] and tunneling to guarantee the scalability property in decreasing the amount of state information to be processed in the core network. Besides, SR's main benefit is to fix the scalability issues and limitations of the multiprotocol label switching (MPLS) approach.

To improve Quality of Service (QoS) of the network and based on the difficulties for a human to create and operate network slices manually, the tasks automation has been proposed by applying Machine Learning (ML) techniques [9]. Applying ML in routing optimization and traffic engineering is a new trend. To self-organize parameters that become significantly larger and also gain sufficient flexibility for network maintenance, it is essential for cellular networks to observe environment variations, learn uncertainties, plan response actions, and configure the networks properly. Coincidentally, AI mainly solves how to learn the variations, classify the issues, forecast future challenges, and find potential solutions, by interacting with the environment [4]. Thus, cellular networks could interact with the environment using AI to fully accelerate the evolution.

III. Delay Calculation in Beyond 5G Networks using Time Series Forecasting

Time series forecasting is a technique used to make predictions about future values of a time series based on past observations. Time series data is a sequence of observations made over time, and forecasting involves analyzing this sequence to identify patterns and trends that can be used to make predictions about future values [10].

Machine learning (ML) forecasting proved to be the most effective in capturing the patterns in the sequence of both structured and unstructured data and its further time series analysis forecasting.

There are several techniques that can be used for time series forecasting, including [11]:

- **Autoregressive integrated moving average (ARIMA):** This is a popular model for time series forecasting that uses three parameters (p, d, q) to model the autoregression, differencing, and moving average components of the time series.
- **Exponential smoothing:** This is a family of models that use a weighted average of past observations to make predictions about future values. There are several types of exponential smoothing models, including simple, Holt's linear, and Holt-Winters.
- **Seasonal decomposition of time series (STL):** This is a method that decomposes a time series into three components: trend, seasonal, and residual. Each component is analyzed separately, and predictions are made by combining the results.
- **Neural networks:** This is a type of machine learning algorithm that can be used for time series forecasting. There are several types of neural networks that can be

used, including recurrent neural networks (RNNs) and long short-term memory (LSTM) networks.

- **Prophet:** This is a forecasting library developed by Facebook that uses an additive model to decompose a time series into trend, seasonality, and holidays. It can handle missing data, outliers, and changes in trend over time.

When choosing a forecasting method, it is important to consider the specific characteristics of the time series being analyzed, such as seasonality, trend, and the presence of outliers or missing data. It is also important to evaluate the accuracy of the forecasts using metrics such as mean absolute error (MAE), mean squared error (MSE), or root mean squared error (RMSE).
In this section, we present the seasonal decomposition of time series forecasting based on the real dataset. The following subsections provide a description about the dataset which is used, data preprocessing and visualizing the data using STL.

A. Dataset Description

In this work, we worked on network records of three eNodeBs belongs to a cell site in a residential area that covers 53,000 users on each day with 375 gigabytes of traffic volume. This data was collected by monitoring each eNodeB during 200 days. These records were modified to be implemented as network slices in our proposed topology. The design of our network architecture is based on the standard design of the transport part of 5G networks. A simplified topology of the considered network can be seen in Fig. 2.

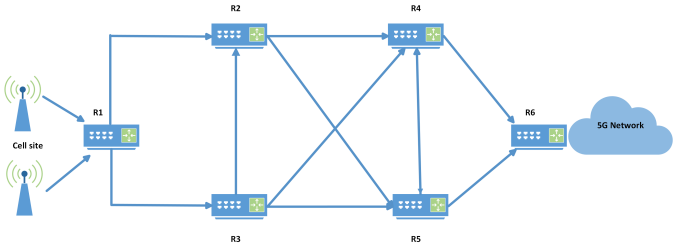


Fig. 2: Proposed topology.

The structure of our mentioned dataset is as follows: Source IP, destination (with this format: fdXX::100 which XX is the router's number), packet size, segment left, date/time and epoch time. Based on the proposed topology, this data was collected from 6 different routers for 15 seconds measurements in a 15 minutes time slot. A sample of the dataset is shown in Table I . As it is shown in this example, the IPs format is IPv6. The destinations are router 2 and router 3 with different packet sizes. The segment left shows how many hops are remained for the packet to leave the topology.

B. Preprocessing Data

Machine learning success depends on many factors. The first and foremost of them is the quality of the data and how it is represented. In ML process, data preparation and filtering steps could take the most amount of the processing time. Data preprocessing includes data cleaning, normalization, transformation, feature extraction and selection, etc. The product of data preprocessing is the final training set [12], [13].
Data preprocessing is an essential step in the analysis of time series data. Time series data is characterized by observations that are made over time, and these observations are often dependent on each other. The goal of data preprocessing in time series data is to transform the raw data into a form that is suitable for analysis. The following list introduces some common techniques used in data preprocessing for time series data:

- **Data cleaning:** This involves identifying and removing any errors or anomalies in the data, such as missing values, outliers, and duplicates.
- **Data normalization:** This involves transforming the data so that it is on a common scale. This can be done using techniques such as min-max scaling or z-score normalization.
- **Data smoothing:** This involves removing noise from the data to make it easier to analyze. This can be done using techniques such as moving averages or exponential smoothing.
- **Feature engineering:** This involves creating new features from the existing ones. For example, in financial time series data, one might create new features such as moving averages or technical indicators.
- **Data aggregation:** This involves summarizing the data over different time periods. For example, daily data might be aggregated to weekly or monthly data.
- **Handling missing values:** Time series data often contains missing values, which can be handled by imputing the missing values using techniques such as interpolation or regression.
- **Feature scaling:** This involves scaling the features to a similar range to avoid one feature dominating the analysis.

In this work, for preprocessing the data, first datasets for different time slots of each routers were merged together to have one dataset for each router using a Python script. Second, the average delay was calculated between different routers to be implemented later for the ML algorithms for prediction. The average delay between different routers was calculated based on these three assumptions:

- We consider the IPs that came out of the Router 1 (R1) where their next destinations would be R2 or R3. Therefore, the delay between R1-R2 and R1-R3 needs to be calculated.
- For the internal routers, the delay between these routes were calculated: R2-R4, R2-R5, R3-R2, R3-R4, R3-R5, R4-R5 and R5-R4.
- At the end point of the topology (R6), we considered just the IPs which were coming from R4 and R5.

Then all the calculated average delays for different time

TABLE I: A sample of utilized dataset.

Source IP	Destination	Packet Size (Byte)	Segment left	Date/Time	Epoch Time
e000::ee6:587d	fd33::100	248	4	Jan 31, 2023, 21:00:36.336862500 UTC	1675198836
c000::ee6:587d	fd22::100	248	4	Jan 31, 2023, 21:00:36.336868600 UTC	1675198836
d000::ee6:587d	fd33::100	248	4	Jan 31, 2023, 21:00:36.336872900 UTC	1675198836
a000::ee6:587d	fd22::100	232	3	Jan 31, 2023, 21:00:36.336872500 UTC	1675198836
b000::ee6:587d	fd33::100	232	3	Jan 31, 2023, 21:00:36.336877400 UTC	1675198836

slots was gathered in a table for 24 hours. An example of the calculated average delay for 4 time slots is shown in Table II.

C. Data Visualization

Data visualization is an essential tool in time series forecasting as it helps to understand the patterns and trends in the data, which can help to identify potential forecasting models and improve the accuracy of the forecast. There are different techniques for data visualization in time series forecasting such as: line charts, seasonal decomposition, autocorrelation plots, heatmaps and box plots [14]. Overall, data visualization is a crucial step in the time series forecasting process as it helps to identify patterns and trends in the data, which can inform the selection of appropriate forecasting models and improve the accuracy of the forecast.

In this work, we implemented the seasonal decomposition technique for visualizing our dataset. It is a technique to extract multiple types of variation from the dataset. There are three important components in the temporal data of a time series: seasonality, trend, and noise [15].

- **Seasonality:** It is a recurring movement that is present in time series variable.
- **Trend:** It can be a long-term upward or downward pattern.
- **Noise:** It is the part of the variability in a time series that can neither be explained by seasonality nor by a trend. When building models, you end up combining different components into a mathematical formula. Two parts of such a formula can be seasonality and trend. A model that combines both will never represent the values of the dataset perfectly: an error will always remain. This is represented by the noise factor.

IV. Experimental Results

To decompose our dataset, we used the statsmodels library. Statsmodels is a Python module that provides classes and functions for the estimation of many different statistical models, as well as for conducting statistical tests, and statistical data exploration. This library provides a method to represent a time series as a sum (or, sometimes, a product) of three components - the linear trend, the periodic (seasonal) component, and random residuals called seasonal_decompose(). It requires to specify whether the model is additive or multiplicative. Additive decomposition argues that time series data is a function of the sum of its components. Rather than a sum, the multiplicative decomposition argues that time series data is a function of the product of its components.

The seasonal_decompose() function returns a result object that contains arrays to access four pieces of data from the decomposition. The result object provides access to the trend and seasonal series as arrays. It also provides access to the residuals, which are the time series after the trend, and seasonal components are removed. Finally, the original data is also stored [16].

As it is shown in Table II, there are different values per route in a specific time slot. Therefore, we have to choose each route separately and implement the decomposition for it to be used for predicting the future delay. Thus, here we present the data visualization for R1-R2.

Listing 1: Python script for visualizing delay for R1-R2 route

```
import numpy as np
import pandas as pd
from datetime import datetime
import matplotlib.pyplot as plt
from sklearn import datasets
from sklearn import metrics
from sklearn.model_selection import train_test_split
from statsmodels.tsa.seasonal import seasonal_decompose
import matplotlib.pyplot as plt

# Load the time series data
df = pd.read_csv(r"/content/delay.csv")
# Convert data to time series
ts = pd.Series(df['12'].values, index=df['datetime'])
ts.index = pd.DatetimeIndex(ts.index)

# Seasonal decomposition
result = seasonal_decompose(ts, model='additive',
period=15)

result.plot()
```

Average delay between R1-R2 route for 24 hours is illustrated in Fig. 3. As it is shown, we can observe a dynamic behaviour of the average delay. To extract multiple types of its variation, we implemented the seasonal decomposition using the seasonal_decompose library in a Python script that is presented in Listing 1.

The seasonal decomposition for the average delay between R1-R2 route is presented in Fig. 4. As we can see, there are four graphs which are observed, trend, seasonal and residual. The trend line is missing some observations from the beginning and from the end of our data set and comparing it with the observed graph, it shows the same behaviour. The seasonality component shows the seasonal variations on the time series remains unchanging. Ideally, trend and seasonality should capture most of the patterns in the time series. Hence, the residuals represent what's left from the time series and it

TABLE II: Average delay between different routes for four time slots in seconds.

Time	N1-N2	N1-N3	N2-N4	N2-N5	N3-N2	N3-N4	N3-N5	N4-N5	N4-N6	N5-N4	N5-N6
15:00	0.000063	0.000058	0.000050	0.000065	0.000078	0.000067	0.000069	0.000049	0.000002	0.000048	0.000002
15:15	0.000068	0.000060	0.000052	0.000069	0.000085	0.000072	0.000075	0.000053	0.000002	0.000053	0.000002
15:30	0.000046	0.000045	0.000038	0.000050	0.000066	0.000050	0.000051	0.000038	0.000002	0.000037	0.000002
15:45	0.000056	0.000048	0.000046	0.000051	0.000065	0.000064	0.000057	0.000045	0.000002	0.000042	0.000002

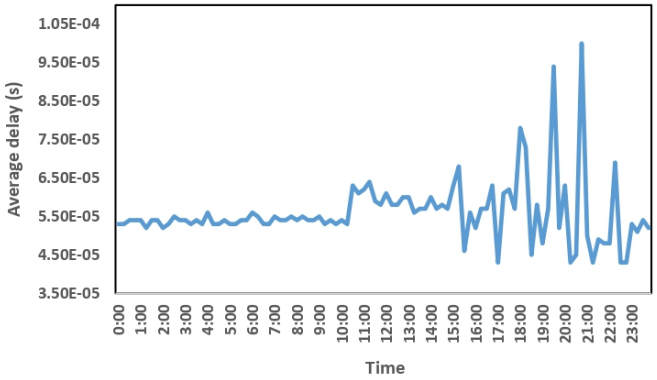


Fig. 3: Average delay between R1-R2 route in 24 hours in seconds.

needs to be as small as possible.

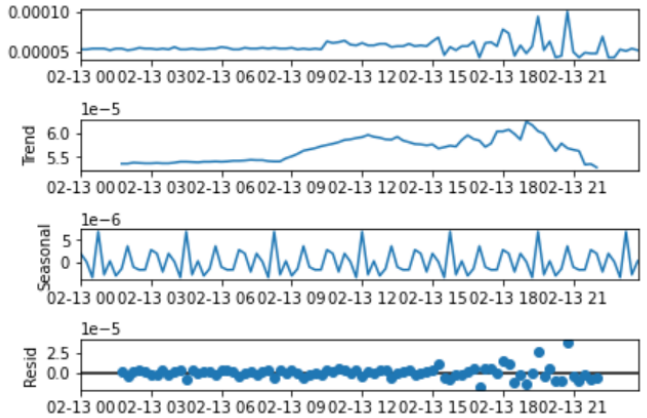


Fig. 4: Delay seasonal decomposition between R1-R2 route in 24 hours.

V. Conclusion and Future Works

In conclusion, data decomposition is a powerful technique for breaking down large, complex datasets into smaller, more manageable parts. By decomposing data, researchers can analyze and understand complex relationships that would otherwise be impossible to discern. There are several methods of data decomposition. Overall, data decomposition is a valuable tool for researchers and practitioners working with large datasets. Its versatility makes it an essential component of modern data analysis.

In this work, we implemented the seasonal decomposition technique. Based on the presented graphs, it seems that our time series dataset has some outliers in some points that needs to be investigated for implementing anomaly detection in time series forecasting. As it is illustrated in the figures of the previous section, in some specific time slots, delay between some routes are much more that the other time slots. Next step will be splitting the dataset to training and testing sets to train the machine learning model based on it. Then the ML models could predict the future behaviour of the delay and propose the best available routes.

References

[1] V. Tong, S. Souihi, H. A. Tran, and A. Mellouk, "Sdn-based application-aware segment routing for large-scale network," *IEEE Systems Journal*, 2021.

[2] M. Van Der Schaar and F. Fu, "Spectrum access games and strategic learning in cognitive radio networks for delay-critical applications," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 720–740, 2009.

[3] D. Banupriya, S. P. Pandi, and R. R. Prathisha, "A study on use of artificial intelligence in wireless communications," *Asian Journal of Applied Science and Technology*, vol. 2, no. 1, pp. 354–360, 2018.

[4] R. Li, Z. Zhao, X. Zhou, G. Ding, Y. Chen, Z. Wang, and H. Zhang, "Intelligent 5g: When cellular networks meet artificial intelligence," *IEEE Wireless communications*, vol. 24, no. 5, pp. 175–183, 2017.

[5] X. Zhou, R. Li, T. Chen, and H. Zhang, "Network slicing as a service: enabling enterprises' own software-defined cellular networks," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 146–153, 2016.

[6] A. Cianfrani, M. Listanti, and M. Polverini, "Incremental deployment of segment routing into an isp network: A traffic engineering perspective," *IEEE/ACM Transactions on Networking*, vol. 25, no. 5, pp. 3146–3160, 2017.

[7] C. Filsfils, N. K. Nainar, C. Pignataro, J. C. Cardona, and P. Francois, "The segment routing architecture," in *2015 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2015, pp. 1–6.

[8] P. L. Ventre, S. Salsano, M. Polverini, A. Cianfrani, A. Abdelsalam, C. Filsfils, P. Camarillo, and F. Clad, "Segment routing: a comprehensive survey of research activities, standardization efforts, and implementation results," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 182–221, 2020.

[9] Y. Tian, Z. Wang, X. Yin, X. Shi, Y. Guo, H. Geng, and J. Yang, "Traffic engineering in partially deployed segment routing over ipv6 network with deep reinforcement learning," *IEEE/ACM Transactions on Networking*, vol. 28, no. 4, pp. 1573–1586, 2020.

[10] "Tableau," https://www.tableau.com/learn/articles/time-series-forecasting/, 2023.

[11] "Vitalflux," https://vitalflux.com/different-types-of-time-series-forecasting-models/#google_vignette/, 2023.

[12] S. B. Kotsiantis, D. Kanellopoulos, and P. E. Pintelas, "Data preprocessing for supervised leaning," *International journal of computer science*, vol. 1, no. 2, pp. 111–117, 2006.

[13] E. Younesian, M. Siklosi, N. Khatib, and J. Hosek, "Syslog anomaly detection using supervised machine learning models," in *2021 13th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. IEEE, 2021, pp. 78–84.

[14] "Codecademy," https://www.codecademy.com/article/visualizing-time-series-data-with-python/, 2023.

[15] "Neptune," https://neptune.ai/blog/select-model-for-time-series-prediction-task/, 2023.

[16] "Machinelearningmastery," https://machinelearningmastery.com/decompose-time-series-data-trend-seasonality/, 2023.

Efficient Multi-UAV Network Placement for Wireless Coverage in Challenging Environments

Balaji Kirubakaran[1,2,*], and Jiri Hosek[1,2], Radim Burget[1]
[1]Department of Telecommunications, Brno University of Technology, Brno, Czech Republic
[2]Unit of Electrical Engineering, Tampere University, Tampere, Finland
[*]Email: balaji.kirubakaran@vut.cz

***Abstract*—Providing wireless coverage for IoT devices in remote/rural and urban environments can be challenging for different reasons. In remote areas, lack of infrastructure limits coverage, while urban areas face congestion, building materials, network capacity, and cost issues. Due to these limitations, traditional cellular networks may need help to provide reliable coverage in both areas. Recently, Unmanned Aerial Vehicle (UAV) networks have been gaining attention as a potential solution to provide wireless connectivity to ground users, particularly in areas that are hard to reach. This approach shows promise in overcoming the challenges associated with ground-based wireless networks in remote and challenging areas. However, determining the most efficient placement of UAV-based stations in a network comprising multiple UAVs is a complex task with various constraints. In this study, we propose a solution that formulates fewer UAVs in a Multi-UAV Network as an Integer Linear Programming (ILP) problem. The proposed solution includes a highly efficient Genetic Algorithm (GA) that utilizes a novel chromosome representation to optimize the placement of UAVs, ensuring wireless coverage for all ground users while minimizing the number of deployed UAVs. To ensure reliable wireless connectivity for all deployed UAVs, we introduce a backhaul-aware K-means clustering algorithm that identifies which UAV will be the backhaul for the other UAVs in the network. Moreover, the solution incorporates a centroid, tethered UAV connected to a robotic Cell on Wheels (COW). Our simulation results show that the proposed solution outperforms existing approaches regarding solution quality and computational efficiency while ensuring reliable wireless connectivity for all deployed UAVs and users. The proposed solution has significant potential to provide wireless coverage to remote/rural and dense urban areas where traditional cellular networks may require more coverage. This study demonstrates the applicability and effectiveness of the proposed solution, making it a promising alternative for providing wireless coverage in challenging areas.**

***Index Terms*—UAV networks, Wireless coverage, Genetic algorithms, Multi-UAV network, Backhaul connectivity**

I. Introduction

As the population grows, the telecommunications sector faces significant challenges, including limited coverage and data traffic capacity. These issues can negatively impact the quality of services provided to consumers, making it more difficult to ensure reliable and high-speed connectivity. Innovative solutions are urgently needed to address the challenges the telecommunications sector faces. One such solution that has emerged is intelligent systems, including smart cities, buildings, vehicles, and grids, which can help address the challenges faced by the telecommunications sector by improving efficiency, reducing costs, and enhancing connectivity for consumers. These systems leverage the inter-connectivity of the Internet of Things (IoT) [1] to offer better services and resource management to the population. However, the vast amount of data generated and collected by these intelligent systems poses a challenge, requiring a powerful communication paradigm to ensure high-quality services in all areas. The need for such a paradigm is crucial in today's world, where reliable and efficient communication is essential for the smooth functioning of various sectors, from healthcare to finance to transportation. Developing such communication systems requires a multidisciplinary approach involving experts in telecommunications, data management, and system design. Only through collaboration can we ensure that the potential of intelligent systems is fully realized and that they can continue to improve people's lives worldwide.

Traditional cellular networks, such as 3G and 4G, have limitations in efficiently supporting various types of communication due to their primary focus on broadband communication optimization [2]. Despite offering extensive coverage, low deployment cost, and high security, these networks can only sometimes meet the requirement of supporting communication with high efficiency. However, the emergence of 5G+ has presented a promising solution for addressing the limitations of traditional cellular networks. 5G+ has been specifically designed to offer users high Quality of Service (QoS), boasting maximum bandwidth, reduced latency and error rates, and improved uptime [3]. The increased data rate, reduced delay, and enhanced cellular coverage of 5G+ make it a promising solution for a wide range of communication needs, such as video conferencing, live streaming, and online gaming. Additionally, 5G+ ability to cater to heterogeneous access technology opens up numerous possibilities for efficient communication services.

The challenge of providing wireless coverage to remote/rural, and dense urban areas where traditional cellular networks are sparse or unavailable is addressed in this paper. The proposed solution involves deploying Unmanned Aerial Vehicles (UAVs) as flying base stations (BSs) [4], which can be flexibly placed where they are most needed, providing efficient and flexible 5G+ coverage for users at all times. Despite increasing interest in using UAV-BS to improve 5G+ coverage and QoS, determining the optimal number and placement of UAVs in a given 3D space remains a challenging problem. The proposed solution formulates the situation as an Integer

Linear Programming (ILP) problem, enabling a novel Genetic Algorithm (GA) [5] with a unique chromosome representation to optimize the placement of UAVs and ensure maximum 5G+ coverage and QoS. This approach provides service providers with fast and effective solutions for deploying UAV-BS to ensure maximum 5G+ coverage and QoS.

Additionally, the proposed solution incorporates a backhaul-aware K-means clustering algorithm, which determines the optimal UAV to serve as the backhaul for the other UAVs in the network. Furthermore, the proposed solution includes a centroid UAV, a tethered UAV connected to a robotic Cell on Wheels (COW). The centroid UAV serves as a charging station for the other UAVs and provides backhaul connectivity to the UAV base station. This component enhances the solution's overall efficiency by enabling the other UAVs to recharge their batteries and receive data without frequent trips back to a ground-based station.

The paper is organized as follows. Section II provides a literature review of related work. Section III identifies and discusses some of the main challenges UAV networks face. Section IV presents the proposed model of the system. Section V describes the K-means clustered algorithm for backhaul connectivity. Section VI presents the simulation results and discussions. Finally, Section VII provides the conclusions and highlights future work.

II. Related Work

The efficient positioning of UAVs has attracted significant attention from researchers and academicians. For instance, Zorbas et al. [6] introduced a minimum-cost UAV location problem in a two-dimensional terrain to find the optimal location and the number of UAVs needed to observe given targets in a specified area. Tuba et al. [7] employed brainstorm optimization algorithms to determine the optimal position of static UAVs in a monitored site to maximize coverage. Shakhatreh et al. [8] utilized UAVs as aerial wireless base stations in natural disasters, finding the optimal position for the UAVs to maximize the sum of time durations of uplink transmissions. Kalantari et al. [9] proposed a heuristic algorithm to optimally place UAV base stations in a region with different target densities, ensuring the quality of service constraints was met. These studies demonstrate the growing interest in developing efficient UAV positioning algorithms to enhance wireless coverage and communication services. Wu et al. [10]. This paper proposes a solution to address the optimal positioning of UAV base stations in a multi-UAV network, considering various constraints such as backhaul connectivity and minimizing the number of deployed UAVs.

III. Challenges of UAV Based Network

UAVs have gained significant attention in recent years as a means of providing wireless coverage in areas with limited or no cellular infrastructure. However, UAV-based networks come with their own set of challenges. The following are some of the challenges of UAV-based networks:

- UAVs have limited resources such as power, memory, and communication bandwidth, making resource allocation a significant challenge.
- Efficient power consumption is crucial to extend the UAVs operation time. It can be achieved by carefully planning the UAV's flight path to optimize power consumption and leveraging sensors to send back information about wind direction and speed during flight.
- Communication in a UAV-based network is unique, and UAVs must exchange data such as current position, speed, and direction with each other while in flight. It requires a communication link with low latency and a wide range of communication to ensure effective data exchange.
- Determining the most efficient placement of UAV-based stations in a network comprising multiple UAVs is a complex task with various constraints. Optimal placement is necessary to ensure wireless coverage for all ground users while minimizing the number of deployed UAVs. Out of all the challenges faced in UAV-based networks, optimal placement is the most critical factor for ensuring efficient and effective operation.

IV. System Model

Our proposed UAV-based network system aims to establish efficient communication between UAVs and a standard sink (UAV-BS COW or BS), which serves as the final data relay. To ensure reliable communication with the UAV-BS COW or BS, we employ a relay system whereby the UAV sends data to the next available UAV until the communication range with the UAV-BS COW or BS is attained. This approach minimizes the system's energy requirements and eliminates the need for high-power, long-range transmitters.

Optimal positioning UAVs is crucial to maximizing 5G+ coverage while minimizing the number of UAVs used. However, UAV placement is a challenging research area due to UAVs limited sensing and communication range and restricted resources, such as coverage. In this paper, we address this challenge by formulating the coverage problem mathematically and proposing a solution that utilizes a GA to calculate the number of UAVs and their respective positions in a given 3D deployment area while maintaining coverage and lifetime constraints. Lifetime constraints refer to the maximum operating time of a UAV-based network system before its batteries need to be recharged or replaced. In other words, it is the period for which the UAV-based network system can operate without interruption or significant maintenance.

In our proposed system, UAVs can fly at heights ranging from v_{min} to v_{max} over a rectangular area with dimensions a_{max} and b_{max}. Targets, defined as areas where terrestrial BS is sparse or unavailable, can be located at any arbitrary position within the area. We assume each UAV can be located instantaneously at a position (a, b, v). Let U denote the set of available UAVs, and T be the set of targets, where each target $t_i \in T$ has a position (A_{t_i}, B_{t_i}), and each UAV $u \in U$ has

a position (A_u, B_u, v_u). The distance between target t_i and UAV u can be calculated using the equation:

$$L_{t_i}^{u_a,u_b} = \sqrt{(A_{t_i} - A_u)^2 + (B_{t_i} - B_u)^2}. \quad (1)$$

Each UAV u has a communication range β in the form of a disc over the area of $a_{max} \times b_{max}$, where the disc radius is determined by its height v_u. The radius r^{v_u} increases with the value of v_u, resulting in a more extensive communication range. The two critical decisions that need to be made are determining the position (A_u, B_u, v_u) of each UAV and finding the target $t_i \in T$ within the specified area of interest. To solve the first problem, we introduce a decision variable γ_{abv}^u that takes a value of 1 if the UAV u is located at (a, b, v) and 0 otherwise. For the second problem, we define a decision variable $\theta_{t_i}^u$ that takes a value of 1 if the target t_i is within the communication range of UAV u and 0 otherwise.

We aim to cover all the target areas using a minimum number of UAVs while minimizing the total energy consumed. The energy consumed by each UAV is given by:

$$M = (\alpha + \delta v)t + P_{max}\left(\frac{v}{s}\right), \quad (2)$$

where α is the minimum power required to hover at almost zero altitudes, δ is the motor speed multiplier, P_{max} is the maximum motor power, s and t are the speed and operating time, respectively. $P_{max}(v/s)$ denotes the power required to raise the drone to a height v at speed s. Note that α and δ depend on the weight of the UAV and the motor characteristics used.

We formulate our placement problem as follows:

$$f(\theta) = O - \sum_{u \in U} \theta_{abv}^u * O_n', \quad (3)$$

where O is the total area to be covered, and O_n' is the area covered by the nth UAV.

Our proposed system is similar to the study in [11], which focuses on minimizing the number of UAVs and energy consumption by formulating static drone location problems (SDLP) and mobile drone location problems (MDLP). However, we use a GA to optimize the placement of UAVs in the 3D deployment area, ensuring reliable coverage and efficient operation. Compared to other approaches, GA offers several advantages, such as the ability to handle complex and nonlinear optimization problems, faster convergence to optimal solutions, and the ability to avoid local optima. These advantages make GA a suitable and effective optimization tool for the placement of UAV-based networks.

V. GA Algorithm for UAV Placement Optimization

We propose using a GA to determine the minimum number of UAVs and their optimal positions for a given set of conditions. The GA algorithm consists of the following steps:

- Initialization: Define the population size (PS), maximum generation number (Gmax), probability of crossover (PC), and probability of mutation (PM).
- Generate an initial set of random solutions.
- Evaluate the fitness of each initial solution using the fitness function.
- Select the best fitness solution.
- Iterate over the generation number, Gmax.
- For each iteration, iterate over half of the PS.
- Select two parents for crossover.
- Produce two children using the crossover operator.
- Mutate the produced children using PM.
- Replace parents with the produced children.
- Update the best-found solution (BFS) by applying the produced solutions to the fitness function and finding the best solution.
- Compare the newly found BFS with the previous BFS and update it if it is better than the previous.
- End the iteration over the PS.
- Return the best solution found.

This algorithm evolves from one generation to another by generating new solutions using crossover and mutation operators. The fitness function is used to evaluate the quality of each solution, and the best solution is retained for further evolution. The GA algorithm is well-known for optimizing and producing high-quality solutions for complex problems. It has been widely used in various fields, including wireless communication and network optimization.

VI. K-means clustering of UAVs and users

The K-means clustering algorithm is a popular unsupervised machine-learning approach that can be applied to various scenarios, including our proposed UAV-based network system. Our system utilizes K-means clustering to determine the most suitable location for the UAV base station for charging and backhaul. The UAV base station operates as a swarm of UAVs, and K-means clustering is employed to locate the centroid of the swarm.

To initiate the learning process, we define the number of centroids required in the dataset, which is one centroid UAV for charging and backhaul. The K-means algorithm assigns every UAV in the swarm to the nearest cluster while keeping the centroid as small as possible to identify this centroid. The K-means algorithm selects a random group of UAVs as the starting points for the clusters. It performs iterative calculations to optimize the positions of the UAVs until the defined number of iterations is achieved or the centroids have stabilized due to successful clustering. Upon identifying the centroid UAV through the K-means algorithm, it is linked to the robotic COW for charging and backhaul. The centroid UAV can also serve as a TUAV (Tethered UAV) to provide wireless coverage in the area. The robotic COW is equipped with a charging station and a backhaul module, enabling it to charge the UAVs and provide backhaul connectivity. When a UAV requires charging, it autonomously flies to the robotic COW and lands on the charging station, which charges the UAV's batteries. Once charged, the UAV resumes its operation, providing wireless coverage. This process guarantees that the UAVs in the swarm are charged and prepared for their next mission. By utilizing K-means clustering to determine

the optimal candidate UAV for charging and backhaul, we can ensure the efficient utilization of resources and optimal performance of the UAV swarm.

Our proposed solution utilizes K-means clustering to find the most suitable UAV among the others within the swarm and GA to determine the optimal placement of the UAVs in the 3D deployment area while maintaining coverage and lifetime constraints.

VII. Simulation Results and Discussions

In this section, we will analyze the GA algorithm's optimal placement results for the UAVs and K-means clustered-based backhaul. We consider predetermined simulation parameters [12] (see Table I) for UAV-based network solutions in urban scenarios. We carefully chose these simulation parameters to ensure our proposed system is tested under realistic conditions. We chose the Manhattan model, which is widely used to simulate urban environments, and a large area of interest (2000 m $\times$ 2000 m) to provide a realistic scenario. We selected the dimensions of the buildings and roads to reflect real-world environments. The number of users (20000) is a reasonable estimate for an urban environment, and we chose the height and coverage radius of the UAV station to ensure effective communication with ground users. We estimated the number of available UAVs (50) to reflect a small-scale UAV-based network system.

TABLE I
Simulation Parameters

Parameter	Value
Model	Manhattan
Area	2000 m $\times$ 2000 m
Building height	50 m
Building width	100 m
Road width	200 m
Number of users	1000
UAV-BS station height	25 m
UAV-BS station coverage radius	400 m
Base station height	80 m
Available UAVs	50
Max generations	100
Population size	50
Max stall generations	100

We selected the GA parameters to ensure the algorithm can efficiently and effectively optimize UAV placement. We chose a maximum of 100 generations, and a population size starts with 50 to balance computation time and optimization quality. We also set a maximum of 10 stall generations to prevent the algorithm from getting stuck in local optima. These simulation parameters provide a realistic and robust test bed for our proposed UAV-based network system. The results of our simulation demonstrate the effectiveness of the proposed GA algorithm in optimizing the placement of UAVs in an urban area to provide maximum coverage to all users with minimal resources. Fig. 1 shows the optimal placement of UAVs, and it is evident that 15 to 20 UAVs can provide complete coverage to all 20000 users in the area, depending on the placement strategy used. This suggests that UAV-based networks can provide efficient network coverage in urban environments where traditional network infrastructure may be difficult or expensive to deploy. The approach proposed in this study can support the design of future UAV-based network solutions in urban areas, enabling more efficient and reliable connectivity for users.

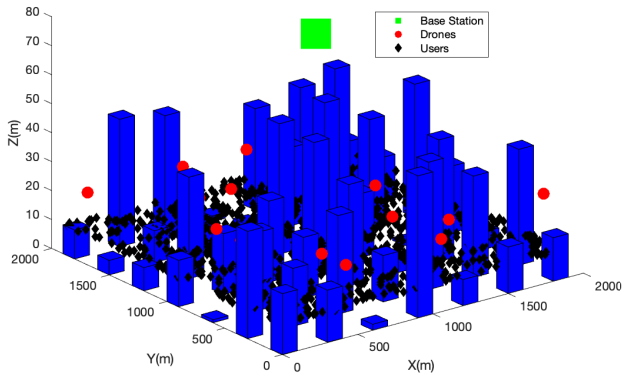


Fig. 1. Optimal placement of UAVs using GA

Furthermore, the use of K-means clustering algorithm in identifying centroid points for each cluster of users and UAVs is an effective solution for establishing backhaul connectivity and charging infrastructure in urban areas. The centroid points were identified for each cluster, providing backhaul connectivity and charging for the UAVs. If a base station is damaged, the tethered COW can be used for backhaul connectivity, ensuring uninterrupted network coverage. The tethered COWs can be connected to nearby ground access points with the help of other TUAVs. The TUAVs can act as a relay node, facilitating communication between the tethered COWs and the ground network access points. The use of centroid UAVs for charging reduces the need for multiple charging stations, allowing for a more streamlined and efficient network. Fig. 2 shows the K-means clustered view of users and UAVs, highlighting the importance of grouping users and UAVs based on their proximity using the K-means clustering algorithm. We found two clusters of UAVs with two centroid points and 18 clusters of users with 18 other centroid points for users. By clustering users, we can determine the areas with the highest demand for wireless coverage. Fig. 3 shows the bar graph for the clusters versus the number of users and UAVs in each cluster. The graph visually represents the distribution of users and UAVs across the different clusters. It is clear from the graph that some clusters have a higher number of users and UAVs than others, which highlights the importance of the K-means clustering algorithm in grouping the users and UAVs based on their proximity.

The graph also shows that the number of UAVs in each cluster

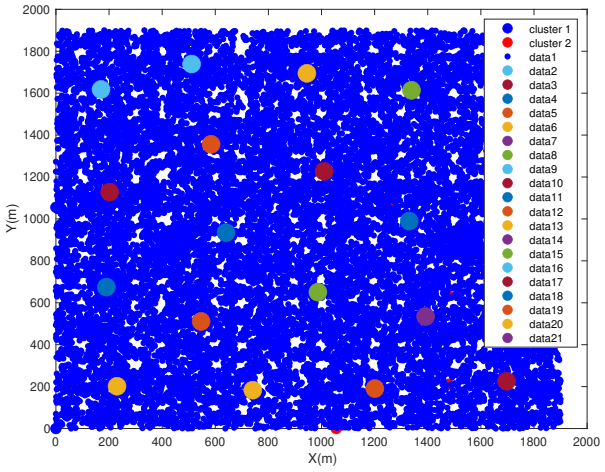


Fig. 2. K-means Clustering of users and UAVs

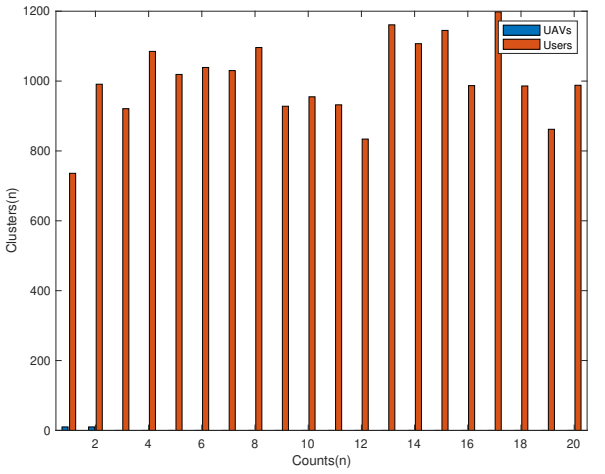


Fig. 3. Number of UAVs and users in each cluster

is significantly lower than the number of users. This is because the UAVs are strategically placed to provide maximum coverage with minimal resources while the users are distributed more evenly across the area. Using UAVs in the network provides an efficient and cost-effective solution for coverage in hard-to-reach areas or areas with limited infrastructure.
Overall, the bar graph provides valuable insights into the distribution of users and UAVs across the different clusters. It highlights the effectiveness of the K-means clustering algorithm in optimizing the placement of UAVs and providing efficient network coverage in urban environments.

VIII. Conclusion

In conclusion, this study proposes a novel and efficient solution for wireless coverage to ground users in remote and hard-to-reach areas using a Multi-UAV network. The proposed solution formulates the placement of UAVs as an ILP problem and employs a GA algorithm to optimize the placement of UAVs while minimizing the number of deployed UAVs. The backhaul-aware K-means clustering algorithm is utilized to identify which UAV will serve as the backhaul for the other UAVs in the network. The solution also incorporates a centroid UAV connected to a robotic COW for charging purposes. Simulation results demonstrate that the proposed solution outperforms existing approaches regarding solution quality and computational efficiency, ensuring reliable wireless connectivity for all deployed UAVs. The proposed solution suits remote/rural, and dense urban areas where traditional cellular networks may need more coverage. Future work can focus on further improving the proposed solution, including real-world testing and integration with existing wireless networks to enhance connectivity in challenging areas.

References

[1] Y. Koucheryavy, S. Balandin, and S. Andreev, "Correction to: Internet of things, smart spaces, and next generation networks and systems," in *Internet of Things, Smart Spaces, and Next Generation Networks and Systems: 21st International Conference, NEW2AN 2021, and 14th Conference, ruSMART 2021, St. Petersburg, Russia, August 26–27, 2021, Proceedings*. Springer, 2022, pp. C1–C1.

[2] R. Mozny, M. Stusek, P. Masek, K. Mikhaylov, and J. Hosek, "Unifying multi-radio communication technologies to enable mmtc applications in b5g networks," in *2020 2nd 6G Wireless Summit (6G SUMMIT)*, 2020, pp. 1–5.

[3] V. Petrov, M. A. Lema, M. Gapeyenko, K. Antonakoglou, D. Moltchanov, F. Sardis, A. Samuylov, S. Andreev, Y. Koucheryavy, and M. Dohler, "Achieving end-to-end reliability of mission-critical traffic in softwarized 5g networks," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 3, pp. 485–501, 2018.

[4] B. Kirubakaran and J. Hosek, "Extending uav's operational time through laser beam charging: System model analysis," in *2022 45th International Conference on Telecommunications and Signal Processing (TSP)*, 2022, pp. 322–328.

[5] M. H. Rahman, M. Al-Naeem, A. Banerjee, and A. Sufian, "Eeto-ga: Energy efficient trajectory optimization of uav-iot collaborative system using genetic algorithm," *Applied Sciences*, vol. 13, no. 4, p. 2535, 2023.

[6] D. Zorbas, L. D. P. Pugliese, T. Razafindralambo, and F. Guerriero, "Optimal drone placement and cost-efficient target coverage," *Journal of Network and Computer Applications*, vol. 75, pp. 16–31, 2016.

[7] E. Tuba, R. Capor-Hrosik, A. Alihodzic, and M. Tuba, "Drone placement for optimal coverage by brain storm optimization algorithm," in *Hybrid Intelligent Systems: 17th International Conference on Hybrid Intelligent Systems (HIS 2017) held in Delhi, India, December 14-16, 2017*. Springer, 2018, pp. 167–176.

[8] H. Shakhatreh and A. Khreishah, "Optimal placement of a uav to maximize the lifetime of wireless devices," in *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)*. IEEE, 2018, pp. 1225–1230.

[9] E. Kalantari, H. Yanikomeroglu, and A. Yongacoglu, "On the number and 3d placement of drone base stations in wireless cellular networks," in *2016 IEEE 84th vehicular technology conference (VTC-Fall)*. IEEE, 2016, pp. 1–6.

[10] V. Buggiani, J. C. Ú. Ortega, G. Silva, J. Rodríguez-Molina, and D. Vilca, "An inexpensive unmanned aerial vehicle-based tool for mobile network output analysis and visualization," *Sensors*, vol. 23, no. 3, p. 1285, 2023.

[11] K. Bezas, G. Tsoumanis, C. T. Angelis, and K. Oikonomou, "Coverage path planning and point-of-interest detection using autonomous drone swarms," *Sensors*, vol. 22, no. 19, p. 7551, 2022.

[12] Y. Zeng, I. Guvenc, R. Zhang, G. Geraci, and D. W. Matolak, *UAV Communications for 5G and Beyond*. John Wiley & Sons, 2020.

NTT for Resources-Optimized Hardware Implementation of CRYSTALS-Kyber

1st Petr Jedlička
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
xjedli23@vut.cz

2nd Jan Hajný
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
hajny@vut.cz

***Abstract*—In this work, we deal with a hardware-accelerated implementation of the Number Theoretic Transform for CRYSTALS-Kyber, a lattice-based quantum-safe key encapsulation mechanism. Described implementation is reduced to one butterfly which makes it suitable also for area-optimized designs. Waiting between iterations is preserved using our method of polynomial interleaving. We present a novel method of reusing hardware resources in the programmable logic between the butterflies and the state machines of NTT (*Number Theoretic Transform*) and its inverse variant. Our implementation is also, as the first one, equiped with an interface following the AXI4-Stream standard which makes it suitable for both FPGA-based (*Field Programmable Gate Array*) designs and Zynq-based hardware/software co-designs.**

***Index Terms*—NTT, CRYSTALS-Kyber, Post-Quantum Cryptography, FPGA, VHDL**

I. Introduction

In recent years, there has been a significant progress in the research of quantum computers. The major technology companies involved in this area include Google and IBM. Each of these two companies developed their first quantum computers with 53 qubits in 2019 [1] [2]. The first quantum computer comprising 127 qubits was developed in 2021 [3]. IBM has a plan to finish their first quantum computer that comprises more than 1000 qubits at the end of 2023 [4]. If the research in that field keeps the same trend as it is outlined by this timeline many conventional and widely used cryptoschemes will be at a risk of being broken because quantum computers comprising a sufficient number of qubits will be able to solve hard problems such as the integer factorization and the discrete logarithm much faster than conventional computers. As a consequence of these facts, research teams started to deal with so-called post-quantum cryptography. The PQC (*Post-quantum Cryptography*) represents cryptosystems that are believed to be not vulnerable to quantum computer attacks. CRYSTALS-Kyber is one the lattice based PQC key encapsulation mechanism that has been chosen by the NIST (*National Institute of Standards and Technology*) for standardization. In general, the disadvantage of the post-quantum algorithms is their higher computational complexity. For example, lattice-based algorithms perform many arithmetic operations over long polynomials and their multiplication is one of the most time consuming operations in this schemes. To accelerate this operation, an algorithm of the Number Theoretic Transform is used in this schemes.

This work presents improved VHDL-based (*Very High Speed Integrated Circuit Hardware Description Language*) hardware implementation of NTT and inverse NTT for CRYSTALS-Kyber. Implemented algorithms has been designed according to the reference implementation of Kyber in the C programming language [5].

A. Related Work

During the last years, authors have already implemented NTT and inverse NTT on various hardware platforms but many of the works are not related to post-quantum cryptography. NTT is also useful in homomorphic cryptography for instance. Implementations for different applications work with different prime numbers and different lengths of the polynomials which results in different computational times and different hardware utilizations. Thus, we mention only recent works related to PQC in the following paragraphs.

Ricci *et al.* [6] presented an implementation of Kyber components including NTT, inverse NTT and other CRYSTALS-Kyber components using Xilinx Virtex UltraScale+ architecture. They reached a frequency of 632 MHz on this architecture and relatively low latency thanks to the usage of 4 butterflies in 2x2 layout.

On the other hand, Chen *et al.* [7] presented the implementation of NTT using lightweight platforms such as Xilinx Artix 7 and Xilinx Spartan 6. Even with regard to the used low performance platform, they reach very low frequency of 136 MHz and high latency. Their implementation has relatively low resource utilization though.

Another relevant work was done by Dang *et al.*, [8]. They present results of the implementation of NTT using high performance Zynq UltraScale+. Their implementation has relatively high frequency at the expense of high resource utilization. They also have quite high latency.

The last recent work was done by Jedlicka *et al.*, [9]. They described two implementations of both NTT and inverse NTT. The implementations work with 4 parallel butterflies in 2x2 arrangment processing two input polynomials in parallel by interleaving them in the iterations. The arrangment of the butterflies and the polynomial interleaving result in a relatively

TABLE I
THE IMPLEMENTATION RESULTS AND THE COMPARISON WITH THE OTHER IMPLEMENTATIONS

	LUT	FF	DSP	BRAM	LUTRAM	Frequency [MHz]	Latency [cycles]	Platform
NTT								
Ricci *et al.* [6]	1107	1407	28	3.5	-	637	405	Virtex UltraScale+
Chen *et al.* [7]	442	237	1	1.5	-	136	2055	Artix 7
Dang *et al.* [8]	2040	3223	24	5	-	500	1271	Zynq UltraScale+
Jedlicka *et al.* [9]	1189	1469	28	1.5	302	637	322	Virtex UltraScale+
Inverse NTT								
Jedlicka *et al.* [9]	1568	2161	50	1.5	368	637	334	Virtex UltraScale+

low latency. They used Xilinx Virtex UltraScale+ architecture and reached a frequency of 632 MHz.

Only one of the works mentioned above ([9]) shows the implementation results of inverse NTT and uses interleaving of the input polynomials. None of the works mentions neither reusing of the hardware resources between NTT and NTT nor any standard interface like AXI4-Stream. The implementation results of the works are summarized in the Table I.

B. Contribution

The main contribution of this work is the optimization of NTT for lightweight platforms for which hardware utilization is usually the main optimization goal. The implemented component can work as both NTT and inverse NTT reusing the same hardware resources. Its interfaces follow the AXI4-Stream standard which makes the integration of the component very easy also for Zynq-based hardware/software co-designs without any need of the implementation of additional interface layers.

Despite the fact that our work is targeted mainly on area-optimized designs, the implementation is still also suitable for performance-optimized applications thanks to its good ratio between the hardware utilization and the computational time.

II. PRELIMINARIES

The Number theoretic transform (NTT) is a generalization of the discrete Fourier transform, which is carried out in a finite field instead of complex numbers. Assuming a polynomial $f(x)$ with coefficients $(f_0, f_1, ..., f_{n-1})$ and their representation in the NTT domain $(\hat{f}_0, \hat{f}_1, ..., \hat{f}_{n-1})$, the transformations can be expressed by the following formulas [10]:

$$\hat{f}_i = \sum_{j=0}^{n-1} f_i r^{ij} \mathsf{mod}(q), i = 0, 1, ..., n-1,$$

$$f_i = \frac{1}{n} \sum_{j=0}^{n-1} \hat{f}_i r^{-ij} \mathsf{mod}(q), i = 0, 1, ..., n-1,$$

where $r = 17$ is the $2n$-th primitive root of unity in modulo q arithmetic [5].

Kyber performs operations in the polynomial ring $\mathcal{R}_q = \mathbb{Z}[x]_q/(x^n + 1)$ where n and q are two integers, i.e. $n = 256$ and $q = 3329$. Using the textbook definition for the polynomial multiplication, the achieved time complexity is $O(n^2)$. However, the polynomial multiplication accelerated by the properties of the NTT reduces the time complexity to $O(n \log n)$. The multiplication of polynomials $f(x)$ and $g(x)$) using the NTT can be expressed by the following formula [5]:

$$f(x) \times g(x) = NTT^{-1}(NTT(f(x)) \odot NTT(g(x))),$$

where $\odot$ is the point-wise multiplication. The NTT can be used only for the $q = 1 \mathsf{mod}(2n)$ [5].

III. IMPLEMENTATION

The design was implemented in Vivado 2022.1 using the VHDL language. Regarding the target platform, two different platforms were chosen – Xilinx Zynq-based board with programmable logic equivalent to the Artix7 series and Xilinx Zynq UltraScale+ board. FPGAs from the Artix7 series can be considered as cheaper, lightweight platforms with less hardware resources and lower maximum clock frequency. On the other hand, UltraScale+ represents the top of the line series. The main optimization goal of the design is the area but also a good ratio between hardware resources and the time of computation. The second goal makes the implementation suitable also for performance-oriented designs. For example, parallel instantiation of several NTT components could be a good solution for such a design.

A. Interfaces

As it has been already mentioned, input and output interfaces follow the specification of the AXI4-Stream standard. The advantage of standard interface is that it is well-known and hence it is easier for other developers to integrate the component to their designs. The design is also ready to use for hardware/software co-designs on SoC (System on Chip) platforms like Zynq from Xilinx. Many IP blocks from Xilinx also use AXI4-Stream interfaces which makes our design easier for its integration among this IP blocks. For example, the processing system of the Zynq platform uses memory mapped AXI4 or AXI4-Lite interfaces and the Xilinx IP block for the transition of a stream of data to the memory mapped domain uses AXI4 and AXI4-Stream interfaces. An example of such design is in the Figure 1. The NTT component is directly connected to a Xilinx IP block that ensures a direct memory access to the Zynq processing system. All Xilinx IP blocks use AXI4 interfaces. The simulation of the AXI4-Stream interface can be seen in the Figure 3. One transaction is performed

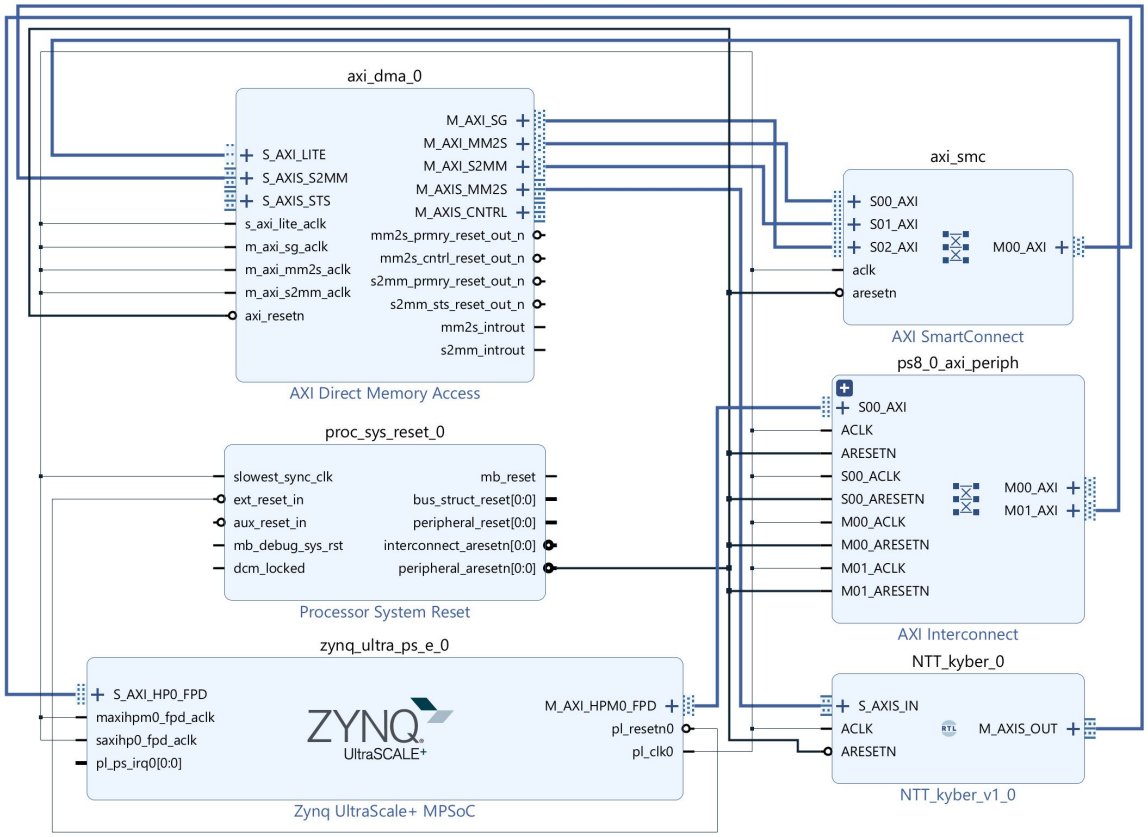


Fig. 1. Intregration of NTT with Zynq Processing System

per two clock cycles because one transaction comprises 4 coefficients but the internal block memory has only 2 ports so it takes two clock cycles to save all the coefficients.

B. Design Description

The presented implementation is based on our previous work presented in Jedlicka *et al.* [9]. Specifically, we use the method of interleaving the input polynomials, the block memories for saving the intermediate results, coefficients and twiddle factors and we modified the state machine and the butterflies for reusing the hardware resources. We use only one butterfly.

Polynomials interleaving is a method of switching between two polynomials after iterations to preserve waiting after each iteration to have saved all the coefficients in the memory before starting another iteration. Instead, we process a given iteration of both polynomials and immediately continue to the next iteration. The number of saved clock cycles per each iteration corresponds to the latency of the butterfly structure.

Reusing the hardware resources is based on using the same components for both NTT and inverse NTT by changing the internal connections or state transitions. The order of processed coefficients in each iteration is inverse between NTT and inverse NTT. It means that the order in the 1st iteration of NTT corresponds to the order of 7th iteration of the inverse NTT etc. Each state in the state machine corresponds to a concrete coefficients order. The same state machine can be used, only the state transitions are different and each state represents different iterations for NTT and inverse NTT. The state machine diagram is in the Figure 4. The butterflies for both of the NTT versions use the same reduction algorithms and arithmetic blocks, only their connections differs. Adding a set of multiplexers changing these connections enables reusing the building blocks. This set of multiplexers is represented by the interconnect component in the block scheme in the Figure 5. The simulation of NTT can be seen in the Figure 2. At first, data is received at the input AXI4-Stream interface. Then, the `tready` signal is going low and the transformation is performed. In the end, the output data is send from the output AXI4-Stream interface.

IV. Results and Discussion

The described design was synthesized for two different platforms (Artix7 and Zynq UltraScale+). The results of the synthesises are summarized in the Table II. The NTT and the inverse ware sythesized as one component so the amount of the hardware resources is the same for both of them, only the computational time slightly differs.

The presented design was compared with the previous ones assuming the ratio between the hardware utilization and the computational time and considering the used platform. The design presented by Chen *et al.* [7] is approximately 50

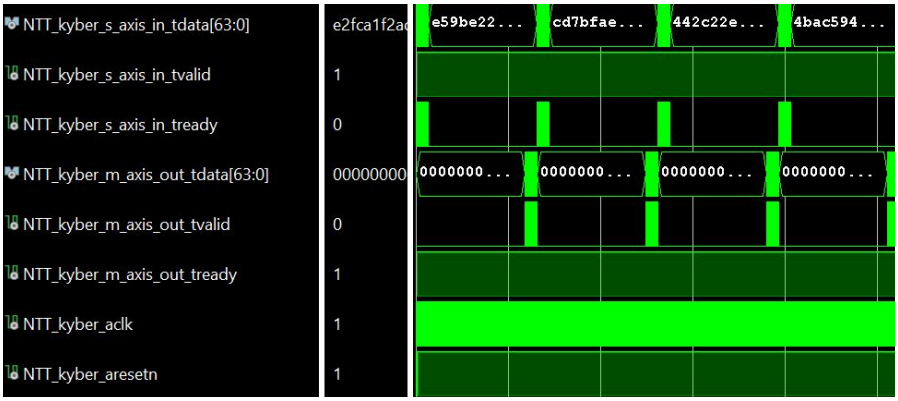


Fig. 2. Simulation of NTT transformations

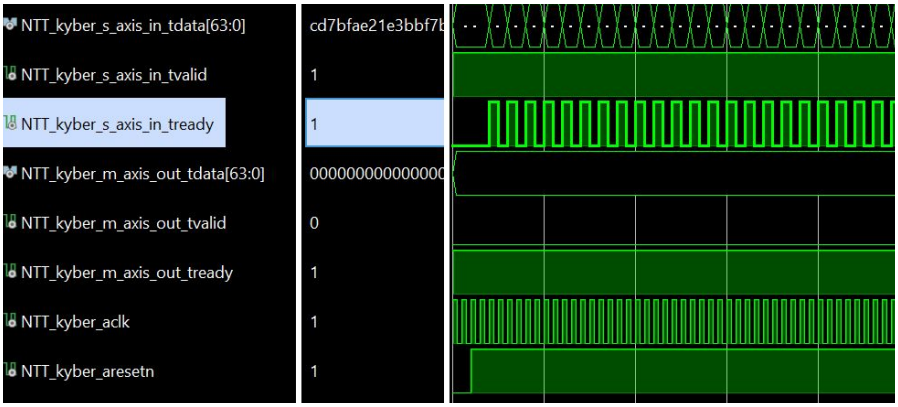


Fig. 3. Simulation of AXI4-Stream transactions

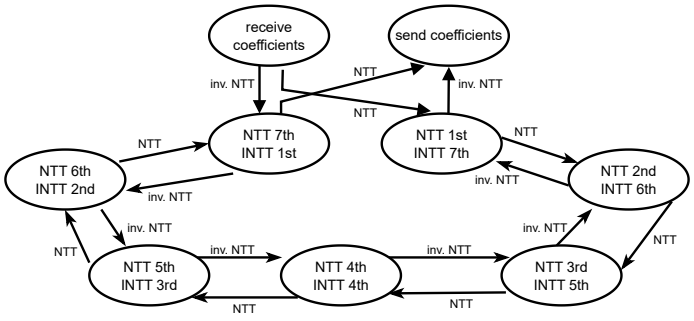


Fig. 4. State Machine Diagram

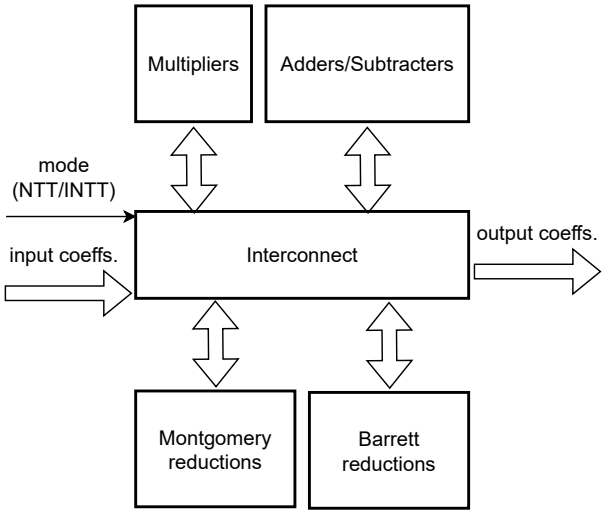


Fig. 5. Butterfly Interconnect

times less effective. Much better results were reached in the work done by Dang *et al.* [8]. Their implementation is approximately 4 times less effective than our work. On the other hand, works presented by Ricci *et al.* [6] and Jedlicka *et al.* [9] are approximately 1.5 times more effective but they need to instantiate two components for both NTT and inverse NTT. If both of the NTT variants are needed in the design our design will the most effective.

V. Conclusion

To the best of the author's knowledge, we proposed the first implementation of NTT and inverse NTT with reusing

TABLE II
HARDWARE UTILIZATION.

	Artix7	**Zynq UltraScale+**
LUT	978	969
FF	952	956
DSP	17	17
BRAM	1.5	1.5
LUTRAM	130	134
Frequncy [MHz]	455	637
Latency [cycles] (NTT)	920	920
Latency [cycles] (inv. NTT)	986	986

the hardware resources between these components in this work. Our design also includes AXI4-Stream interface which makes it suitable for its integration with other AXI4-based IP blocks. Regarding the ratio between hardware utilization and performance, our design is the most effective compared to other recent implementations if both of the NTT versions are used in the design. The implementation is targeted to area-optimized designs but it also suitable for performance-oriented applications. The design was synthesized for two platforms. The maximum reached frequency was 455 MHz for Artix7 and 637 MHz for Zynq UltraScale+.

VI. ACKNOWLEDGEMENT

This work is supported by Ministry of the Interior of the Czech Republic under grant VJ01010008.

REFERENCES

[1] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019.

[2] "On "quantum supremacy"," Aug 2020. [Online]. Available: https://www.ibm.com/blogs/research/2019/10/on-quantum-supremacy/

[3] "On "ibm představila nový kvantový čip: díky němu budou výkonné počítače působit zastarale"," Nov 2021. [Online]. Available: https://www.chip.cz/novinky/ibm-predstavila-novy-kvantovy-cip-diky-nemu-budou-vykonne-pocitace-pusobit-zastarale//

[4] J. Gambetta, "Ibm's roadmap for scaling quantum technology," Sep 2022. [Online]. Available: https://research.ibm.com/blog/ibm-quantum-roadmap

[5] R. M. Avanzi, J. W. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. M. Schanck, P. Schwabe, G. Seiler, and D. Stehlé, "Crystals-kyber algorithm specifications and supporting documentation," 2017.

[6] S. Ricci., P. Jedlicka., P. Cibik., P. Dzurenda., L. Malina., and J. Hajny., "Towards crystals-kyber vhdl implementation," in *Proceedings of the 18th International Conference on Security and Cryptography - SECRYPT,*, INSTICC. SciTePress, 2021, pp. 760–765.

[7] Z. Chen, Y. Ma, T. Chen, J. Lin, and J. Jing, "Towards efficient kyber on fpgas: A processor for vector of polynomials," in *2020 25th Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2020, pp. 247–252.

[8] V. B. Dang, F. Farahmand, M. Andrzejczak, K. Mohajerani, D. T. Nguyen, and K. Gaj, "Implementation and benchmarking of round 2 candidates in the nist post-quantum cryptography standardization process using hardware and software/hardware co-design approaches," Cryptology ePrint Archive, Paper 2020/795, 2020, https://eprint.iacr.org/2020/795. [Online]. Available: https://eprint.iacr.org/2020/795

[9] P. Jedlicka and J. Hajny, "Vhdl-based implementation of crystals-kyber components on fpga," in *Proceedings I of the 28th Conference Student EEICT*. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2022, pp. 297–301.

[10] H. Nejatollahi, N. D. Dutt, S. Ray, F. Regazzoni, I. Banerjee, and R. Cammarota, "Post-quantum lattice-based cryptography implementations," *ACM Computing Surveys (CSUR)*, vol. 51, pp. 1 – 41, 2019.

Application of YOLOv7 neural network model for control of laboratory processes

1st Martin Kiac
Dept. of Telecommunications
Brno University of Technology
Brno, Czech Republic
xkiacm00@vutbr.cz

2nd Kamil Říha
Dept. of Telecommunications
Brno University of Technology
Brno, Czech Republic
rihak@vut.cz

3rd Ondřej Krajsa
Dept. of Telecommunications
Brno University of Technology
Brno, Czech Republic
krajsao@vutbr.cz

***Abstract*—Currently, the world is experiencing an ever-increasing boom in the use of artificial intelligence, especially deep learning. Deep learning and its applications are increasingly popular and used in many fields, such as industry, security systems, and even medicine. This work deals with the problem of detection and evaluation of the pipetting process. The entire system is based on the use of a camera, respectively the camera of a mobile device or tablet, which is conveniently positioned to capture the scene being captured. All image processing, object detection using the YOLOv7 (You Only Look Once 7. version) model and subsequent other operations are performed on a mobile device or tablet. The YOLOv7 neural network model was trained on our own prepared dataset. This training set was created specifically for the analysis of the pipetting process. The result of this work is a prepared annotated dataset and a trained YOLOv7 neural network model, which is aimed at detecting the entire pipette and the tip of this pipette in the image scene. The output of the work is also an implemented algorithm that can perform a complex analysis of the pipetting process using the trained YOLOv7 model. All materials used, dataset and scripts used in this work are available at https://github.com/KicoSVK/Application-of-YOLOv7-for-control-of-laboratory-processes.**

***Index Terms*—pipette, microplate, wells, laboratory processes, image processing, convolutional neural network, object detection**

I. Introduction

The field of artificial intelligence, or deep learning, is already one of the most used techniques today, yet deep learning is constantly evolving and is still expanding into many other fields. Techniques such as computer vision and deep learning are used today in many fields, such as industry, security systems, and even medicine. With the advent of modern technologies, deep learning thus becomes more and more accessible in many applications and devices. These techniques are already implemented and users are using them without knowing about this fact. [1], [2].

Deep learning is a very complex technique that can replace classical techniques. Among the basic techniques of computer vision using deep learning is the ability of a computer to recognize an object [3]. In order for a computer to be able to recognize individual objects, it is necessary to implement this ability, or to teach this computer to recognize objects in some way. There are several ways to implement this ability on a computer. One of those options is to use the aforementioned deep learning technique, or neural network. The technique of neural networks today is among the most widespread techniques in the field of artificial intelligence and achieves very good results [2].

This thesis deals with the issue of detection and evaluation of the pipetting process. The practical application of deep learning techniques, or the use of neural networks for complex analysis of the pipetting process, can be widely used in small experimental laboratories in the civil sector as well as in large professional laboratories. The use of the neural network model learned in this way and the implemented algorithms can greatly simplify and make more efficient the resulting quality of the pipetting process, see image 1.



Fig. 1: Practical demonstration of the pipette process.

II. Related works

The task of solving this work was to design and then implement an application that would be able to process and then analyze individual images coming from the camera. The solution of the work follows the previous work *Control of laboratory processes using modern methods of image processing* [4]. The practical use of the application is focused on the analysis of the pipetting process in laboratories. This means that the resulting application should serve primarily

in laboratories as a laboratory aid, or a simple tool for the complete analysis of the experimental pipetting process.

Currently, there are very few publications and works dealing with this issue in the field of image processing or deep learning. There are several studies and research works that focus on the analyze pipetting process by using intelligent vision technologies. For example, Sadak *et al.* [5] present a vision-based algorithm that can detect the injection and holding pipette as a part of the injection system and then automatically align them in three orthogonal axes. The system works with a pipette tip at the level of microscopic magnification. The authors present a rather complex system performing the pipetting process. However, due to the architecture of the presented system, it is intended for professional laboratories and thus is not intended for small experiments. Gonzalez *et al.* [6] proposed a detection and tracking algoritm, specifically convolutional neural network based on ResNet101, to identify and correct pipette positioning errors. In this work, the authors also work with the pipette tip at the microscopic level. The solution and also the benefit of the work itself is really described in detail. But even in this case, the solution is directly focused on the pipetting process at the level of the microscopic scale.

Directly related to the work investigated in this work, there are few related works, e.g., [7], [8], [9], that focus on the detecting and classifying pipette tip in microplate. Instead of other related works, our system is proposed to analyze and monitor an entire pipetting process within a microplate in order to increase the efficiency and simplicity of the pipetting process. To the best of our knowledge and deep research literature, we present a novelty solution that uses the integration of image processing methods and deep learning techniques to greatly simplify and increase the efficiency resulting quality of the pipetting process.

III. Proposed system solution

Our proposed solution deals with the detection and evaluation of the pipetting process. An illustration of the architecture of the proposed solution can be seen in the following 2.

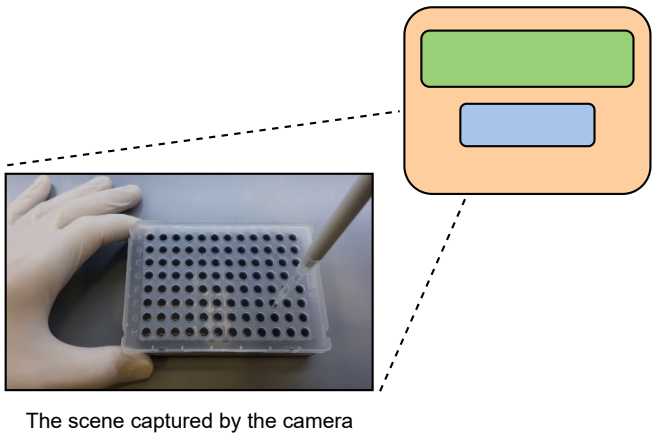


Fig. 2: The architecture of the proposed solution.

A. Analysis of the proposed solution

The practical use of the application is focused on the analysis of the pipetting process in laboratories. This means that the resulting application should serve primarily in laboratories as a laboratory aid, or a simple tool for the complex analysis of the pipetting process. The idea for solving the problem of this work was created in cooperation with the company TATAA Biocenter s.r.o.. Before solving the work, it was necessary to become more familiar with the given problem. This part was also important, because the next direction and way of solving the work depended on it. Pipettes themselves are among the basic laboratory tools used in laboratories. These are measuring containers specially designed so that it is possible to easily handle liquids [10]. By default, pipettes have an oval tubular shape at the end of one side and are gradually narrowed to the very tip of the pipette. Subsequently, using a pipette, it is possible to carry out the pipetting process. Under the term pipetting, it is possible to simply imagine a controlled process of liquid manipulation. For example, in laboratories, a so-called microplate is standardly used as part of the pipetting process, into which individual liquid samples are pipetted [11].

B. Methods of pipetting process analysis

In the course of development, all important algorithms were gradually implemented in the application, such as the recognition of the microplate in the image, the detection of wells in this microplate and the main algorithm for detecting the pipette itself in the image. The implementation of the solution itself was based on previous work [4]. The algorithm for detecting the microplate in the image works on the basis of the analysis of the contours in the image. The subsequent detection of holes in this microplate takes place with the help of geometric shape detection using the so-called *Hough transformation* [12]. The algorithm that serves to detect the pipette and subsequently the tip of this pipette was originally implemented using classic segmentation methods. Background subtraction technique was used to detect the moving pipette in the image, specifically the *mixture of Gaussians* (MoG) [13] algorithm and subsequent processing of the obtained segmentation. This solution was relatively accurate, but showed a high error rate and was also computationally demanding. Optimizing performance, or using physical resources effectively, is very important because individual image processing methods must run not only on computer, but also on a mobile device. Considering this fact, it was necessary to design an algorithm that would use a suitable neural network model to detect the pipette and then the tip of this pipette.

Based on the obtained position of the pipette tip and the detected microplate in the image scene, it was subsequently possible to analyze the pipetting process itself. The algorithm for the analysis of the pipetting process is based on the evaluation of the position of the pipette tip. It uses simple mathematical operations and a timer with which it is possible to evaluate each well and how long the pipette tip was there. After thorough testing of the created functional

unit, the entire source code could be migrated to the Android platform. The following example 3 shows the final form of the implemented solution. The mobile device is conveniently placed on a practical stand, and by using the camera and the implemented software on the Android device, it is possible to effectively evaluate the pipetting process in real-time.



Fig. 3: Practical demonstration of the use of the resulting application in the pipetting process.

C. Selection of a suitable object detector

The main and important output of this work is the software part, i.e. actually the prepared dataset and the learned neural network model, respectively its conclusions and results. In the very beginning of the solution of the work, it was necessary to choose a suitable detector, or a model of a neural network. The selection of a suitable neural network model for the detection of objects in the image represents a basic and essential part of the entire process of selecting a suitable detector. This is due to the fact that each existing neural network achieves different detection accuracy due to its architecture, and also differs significantly in its computational complexity during image analysis.

There are therefore several criteria by which decisions can be made when choosing a suitable neural network model. In the case of solving the problems of our work, one of the main criteria was the sufficient accuracy of the detector and that the chosen detector was also able to work in real time. This means that the chosen model of the neural network should be able to achieve high accuracy and the lowest possible detection error rate at a low or optimal computational complexity. Due to the fact that the resulting neural network model must work on a mobile device or tablet, which has limited performance, this fact had to be taken into account when choosing.

D. Model of neural network

Nowadays, there are already many high-quality models of convolutional neural networks that achieve high accuracy, low error rate, and at the same time can be used in real time. Older models of neural networks such as SSD (*Single Shot MultiBox Detector*) [15], RetinaNet [16] gradually overcome and replace more modern, more powerful models such as EfficientDet [17], YOLO citeyolov7 their other variations. Good results are achieved by the mentioned EfficientDet network model [17]. This network consists of a relatively complex architecture compared to older models of convolutional neural networks. The EfficientDet model does not only use classic subsampling of the input image, this model can also work with the width and depth of the network model architecture, depending on its configuration. The architecture of this network is thus optimized for efficiency and scalability. For this reason, different versions of the EfficientDet model can have substantially different architectures.

The YOLO convolutional neural network has been among the most used models for object detection in real time for several years. It is precisely because of the offered speed and accuracy of this neural network model [14]. This model is mainly used in applications where the emphasis is on low computing power while maintaining optimal accuracy of the resulting detection of the used model. The YOLO convolutional neural network belongs to the algorithms that are based on the so-called regression. Such types of detectors are often referred to as single-stage detectors. This means that classification and subsequent detection are performed simultaneously within one step. The first version of the YOLO convolutional neural network was first published in 2016 by four authors [18], [19] including its main author Joseph Redmon. Starting this year, the YOLO model itself is undergoing gradual development, the result of which are increasingly powerful and accurate versions. It is currently 2023 and there are up to 6 versions or official modifications of this neural network available [20], [21]. Version 6 of the model was skipped for some reason, so the latest version is YOLOv7 [22], [14].

In this work, due to our requirements, we chose the YOLOv7 version of the [14] model. When choosing this model, we were based on professional articles by the authors themselves and materials aimed at comparing the YOLOv7 model with its competitors [14]. It was a matter of course that high accuracy and the lowest possible error rate of the resulting detection, considering the performance of the network. Due to the mentioned facts, the YOLOv7 neural network model is one of the most widely used methods for detecting objects in an image in real time, mainly due to the speed of detection it achieves while maintaining high accuracy. Its accuracy reaches up to 56.8 % mean average precision (mAP) at a detection speed in the range from 5 FPS to 160 FPS [14]. The model achieves these values when detecting objects within the *Microsoft COCO dataset* [23] database on the Tesla V100 graphics card.

YOLOv7 improves speed and accuracy by introducing several architectural reforms. Basic elements in the architecture of YOLOv7 are similar to the older version of YOLOv4. This fact is because YOLOv7 is written by the same authors as YOLOv4. The architecture of model YOLOv7 consists of three important parts: the backbone, neck, and head. This appearance of the basic elements of the architecture is also similar

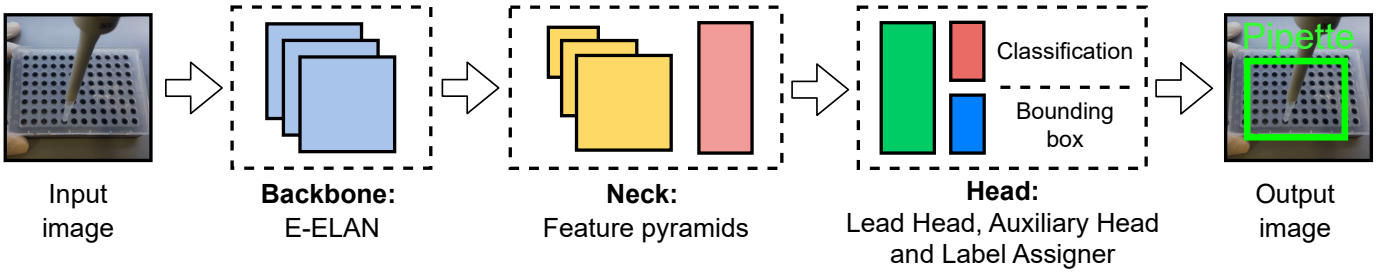


Fig. 4: Architecture of the YOLOv7 convolutional neural network model [14].

with the older version of YOLOv4 [21], [24]. The backbone is responsible for the extraction of features from the given input images, the neck mainly generates the feature pyramids, and the head performs the final detection as an output. In the case of YOLOv7, the authors introduced many architectural changes, such as compound scaling, the extended efficient layer aggregation network (E-ELAN), a bag of freebies with planned and reparameterized convolution, coarse for auxiliary and fine for lead loss [14]. The mentioned architecture of the YOLOv4 model can be seen in the image 4.

E. Preparation of training and validation data

Since there was no dataset available that focused on the pipettes themselves, or on the pipetting process, it was necessary to create such a dataset. For this purpose, the supplied test videos from TATAA Biocenter and our own simulated test videos capturing the pipetting process were used. From these test videos, individual images were created that capture the process of pipetting in different phases and positions of the pipette under different lighting conditions. This is due to the diversity and quality of the training set. A model trained on such a dataset can thus have a better discrimination ability. A total of 750 images were taken, which were subsequently divided into training and test sets. The training set consisted of 450 images and the test set consisted of 300 images. Individual images then had to be manually annotated. The freely available *Yolo_Label* [25] annotation tool was used for annotation. The dataset prepared in this way consists of two classes: the *pipette* and the *pipette tip* class.

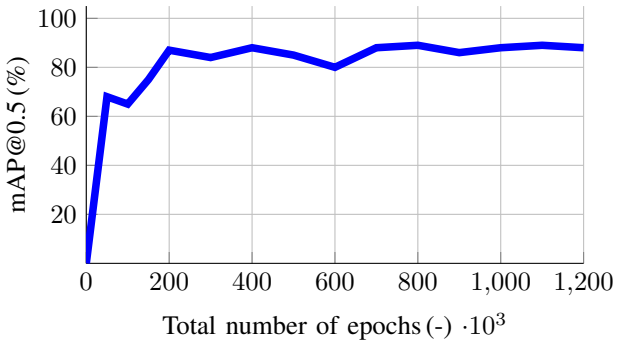


Fig. 5: Progression of mAP metrics while training the YOLOv7 model.

The training of the YOLOv7 convolutional neural network itself took place using the PyTorch framework [26] and publicly available training scripts [27]. The total number of training iterations, or epochs, was set to 1125. The computing device on which the learning itself took place was a computer set with an Nvidia GTX 1060 3 GB graphic card, CPU i7-3770 @ 3.4 GHz and 16 GB RAM memory. Windows 10 operating system, Python 3.9, PyTorch 1.11.0 and CUDA version 11.3 were installed on the device.

TABLE I: Table of the number of annotations of individual objects in the dataset.

Class name	Annotations
Pipette	750
Pipette tip	750

F. Model training process and learning results

There are several parameters that can be used to effectively evaluate the quality of a trained neural network model. These parameters include AP, True Positive (TP) and False Positive (FP). In the sample 5, it is possible to see the learning process of the YOLOv7 neural network on the prepared dataset. In the chart, the blue color is the expressed mAP value, which shows the achieved average prediction (AP) accuracy. As can be seen, after reaching approximately 200 epochs, the mAP parameter does not change significantly and only oscillates around the value of 85% mAP. This value is calculated as the average of the achieved AP accuracies for each defined class. More detailed accuracy results of the trained YOLOv7 model can be seen in II. The parameter TP expresses the dependence of correctly positive detections and the parameter FP expresses false positive detections. The metric AP indicates the average precision with which the given object or class is detected. As can be seen from the table, the average

TABLE II: The resulting accuracy of the learned YOLOv7 network model.

Class name	AP [%]	TP	FP
Pipette	99.3	444	79
Pipette tip	79.4	514	10

AP prediction accuracy for both classes achieves quite good

accuracy. The *pipette* class achieved 99.3% AP and the *pipette tip* class achieved slightly less, namely 79.4% AP. As a result, it can be said that the neural network model learned in this way achieves sufficiently good results for our use, as the average accuracy of the mAP model reached a value of 89.4 %. Model achieves these accuracy values at 19.8 FPS on the Nvidia GTX 1060 3 GB graphic card and on the Samsung Galaxy S22 Android device it reaches 4.2 FPS.

IV. CONCLUSION AND FUTURE WORK

This thesis deals with the issue of detection and evaluation of the pipetting process. The practical application of deep learning techniques, or the use of neural networks for complex analysis of the pipetting process has a strong potential of use. The proposed solution using deep learning is among the original solutions, considering its implementation and the technologies used. There are various other solutions, using various techniques, but our designed solution, which uses deep learning and camera images to analyze the pipetting process, is among the unique solutions. The main and important novelty of this work is the software part, i.e. actually the prepared dataset and the learned neural network model, respectively its conclusions and results. This model was trained using our own dataset with the number of 750 images, which contains 2 classes: *pipette* and *pipette tip*. The learned model achieved sufficiently high-quality results, when processing images with a speed of around 20 FPS on the Nvidia GTX 1060 3 GB graphic card, it was able to detect objects with an average accuracy of 89.4 % mAP. This result was achieved by accelerating calculations on the graphics card using NVIDIA CUDA technology. The resulting solution was also migrated to the Android mobile platform, where it reached 4,2 FPS with the defined accuracy. The algorithm for the analysis of the pipetting process is based on the evaluation of the position of the pipette tip. Based on the obtained position of the pipette tip and the detected microplate in the image scene, it was subsequently possible to analyze the pipetting process itself. In the follow-up work leading to the expansion of this model of the YOLOv7 neural network, it would be possible to continue on the prepared dataset, or to expand the dataset especially with specific images containing other types of pipettes and microplates.

REFERENCES

[1] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, B. C. Van Esesn, A. A. S. Awwal, and V. K. Asari, "The history began from alexnet: A comprehensive survey on deep learning approaches," *arXiv preprint arXiv:1803.01164*, 2018.

[2] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[3] S. Weidman, *Deep learning from scratch*, first edition ed. Beijing: O'Reilly, [2019].

[4] M. Kiac, "Control of laboratory processes using modern methods of image processing," online, Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních, pp. 335–339, june 2020.

[5] F. Sadak, M. Saadat, and A. M. Hajıyavand, "A vision-guided methodology for the automation of biological cell injection," in *2020 2nd International Conference on Electrical, Control and Instrumentation Engineering (ICECIE)*, 2020, pp. 1–9.

[6] M. M. Gonzalez, C. F. Lewallen, M. C. Yip, and C. R. Forest, "Machine learning-based pipette positional correction for automatic patch clamp in vitro," *eNeuro*, vol. 8, no. 4, 2021. [Online]. Available: https://www.eneuro.org/content/8/4/ENEURO.0051-21.2021

[7] K. Koos, J. Molnar, and P. Horvath, "Pipette hunter: Patch-clamp pipette detection," 05 2017, pp. 172–183.

[8] Y. Zhang, K. Tan, and S. Huang, "Vision-servo system for automated cell injection," *Industrial Electronics, IEEE Transactions on*, vol. 56, pp. 231 – 238, 02 2009.

[9] H. Hile, J. Kim, and G. Borriello, "Microbiology tray and pipette tracking as a proactive tangible user interface," vol. 3001, 04 2004, pp. 323–339.

[10] P. fakulta MU, "Práce s různými druhy pipet," http://www.ped.muni.cz/wchem/sm/hc/labtech/pages/prace_pipetami.html, 2008.

[11] ——, "Obecné laboratorní postupy," https://is.muni.cz/el/1431/jaro2013/Bi5220c/um/1_Obecne_laboratorni_postupy.pdf, 2013.

[12] Sklansky, "On the hough technique for curve detection," *IEEE Transactions on Computers*, vol. C-27, no. 10, pp. 923–926, 1978.

[13] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking," in *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, vol. 2, 1999, pp. 246–252 Vol. 2.

[14] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," *arXiv preprint arXiv:2207.02696*, 2022.

[15] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," *Lecture Notes in Computer Science*, p. 21–37, 2016. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-46448-0_2

[16] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2999–3007.

[17] M. Tan, R. Pang, and Q. V. Le, "Efficientdet: Scalable and efficient object detection," 2020.

[18] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[19] J. Redmon and A. Farhadi, "Yolo9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.

[20] J. R. A. Farhadi and J. Redmon, "Yolov3: An incremental improvement," *Retrieved September*, vol. 17, p. 2018, 2018.

[21] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," 2020.

[22] Ultralytics, "Yolov5," https://github.com/ultralytics/yolov5, 2020.

[23] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision*. Springer, 2014, pp. 740–755.

[24] S. R. Kukil, "Yolov7 object detection paper explanation & inference," https://learnopencv.com/yolov7-object-detection-paper-explanation-and-inference/, 2022.

[25] Y. Kwon, "Yolo_label," https://github.com/developer0hye/Yolo_Label, 2018.

[26] 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*. Curran Associates, Inc., 2019, pp. 8024–8035. [Online]. Available: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

[27] WongKinYiu, "Yolov7," https://github.com/WongKinYiu/yolov7, 2022.

vitesco
TECHNOLOGIES

PŘEMÝŠLÍŠ, CO DÁL PO ŠKOLE?

TAK UŽ SE ROZHOUPEJ!

Nastartuj kariéru u nás a proměň si dětské sny v realitu. Místo s autíčky si hrajeme s technologiemi pro auta budoucnosti.

Ve Vitesco Technologies Trutnov, Frenštát pod Radhoštěm a Ostrava nabízíme příležitosti pro absolventy i studenty.

VITESCO.JOBS.CZ



BĚHEM STUDIA I PO NĚM TI RÁDI NABÍDNEME SPOLUPRÁCI:

Odborné placené praxe v oblastech

- softwarového a elektronického vývoje senzorů
- elektrotechniky
- testováni automobilového softwaru a elektroniky
- systémového inženýrství
- industry 4.0

Spolupráci na bakalářské či diplomové práci

Účast na Student Day

- den otevřených dveří pro studenty vč. exkurze do vývojových laboratoří

Globální absolventský program

- aktuálně nabízíme 18-měsíční program v oblasti softwarového vývoje, v rámci kterého se podíváš do různých týmů i do zahraničí

Pracovní příležitosti pro absolventy a absolventky

UPLATNĚNÍ U NÁS NAJDEŠ ZEJMÉNA S TĚMITO OBORY FEKT:

- Elektronika a komunikační technologie
- Mikroelektronika
- Telekomunikační a informační technika
- Kybernetika, automatizace a měření
- Biomedicínské inženýrství a bioinformatika
- Elektroenergetika
- Elektrotechnická výroba a management
- Silnoproudá elektrotechnika a výkonová elektronika
- a další!

Visual defect detection in real-world industrial applications using convolutional neural networks

1st Stepan Jezek
Dept. of telecommunications, FEEC
Brno University of Technology
Brno, Czech Republic
xjezek16@vutbr.cz

2nd Radim Burget
Dept. of telecommunications, FEEC
Brno University of Technology
Brno, Czech Republic
burgetrm@vutbr.cz

***Abstract*—This paper presents a benchmark of current deep learning-based visual anomaly detection methods. Anomaly detection is a critical problem in various field such as fraud detection, cybersecurity and also industrial quality control. Modern advancements in the area of deep learning have shown a significant promise in the ability to automatically detect defects in the image data and this can be exploited in many industrial applications. With the recent development in the field of deep learning, many new visual anomaly detection methods have been introduced recently. These methods are usually evaluated using datasets that do not cover several industrial use cases very well. Performance of many methods may therefore differ, when used in new applications.**

In this paper we perform the image-level and pixel-level AUROC evaluation of recent state of the art visual anomaly detection methods using the MPDD dataset that aims to simulate complex conditions that can be encountered during data acquisition in metal parts manufacturing. These conditions include changes in rotation and position of objects to the camera, changing number of parts in the image or limited number of training samples. We show that recent state of the art mathods may have a significantly lower precision (by up to 30 % AUROC) when deployed in several industrial use cased evaluated in our benchmark.

***Index Terms*—visual anomaly detection, industrial visual inspection, convolutional neural networks, image processing**

I. Introduction

Anomaly detection presents a common and important problem in many areas, e.g. healthcare, cybersecurity, finance and industry, where anomaly detection methods can be used to automatically detect defects that occur during fabrication of products [1]. Quality control is still in many cases performed manually. Recent advancements in the field of machine learning offer many ways for automation of the anomaly detection task. As industrial quality control is very often performed using image data, it can mainly exploit recent developments in the area of deep learning and convolutional neural networks.

In general, anomaly detection task is aimed to detect some form of deviation from a defined state of normality [2]. With this general definition, the task of anomaly detection may have, in some cases, a simple solution. In practice however, the input data can contain a significant ammount of noise, the normal state may not be easily definable or the deviations considered anomalies may be very insignificant, i.e. hard to detect. The problem of anomaly detection can therefore become very challenging and complex in many real-world scenarios [9].

Anomaly detection has traditionally been performed using shallow learning methods such as support vector machines (SVM). For the purpose of processing a high dimensional data, recent deep learning-based methods have shown significantly better results [8]. For this reason, deep learning is also deployed in most recent state of the art anomaly detection methods. The main purpose of the deep learning-based model is to extract features from the visual input data, these features can then be used to discriminate anomalies. Depending on the anomaly detection methods used, the detection itself may be performed on the raw features [23] or further feature processing can be used before the detection [3]–[5].

The task of anomaly detection can be solved via supervised learning, where the detection model is trained as a classifier on a labeled dataset. This approach can reach satisfactory results, in practice however, it is not always feasible to collect a sufficient ammount of data with annotations. In industrial defect detection it is mainly due to optimized manufacturing processes, where production of defective samples is usually minimized. In order to learn a precise supervised model, balanced ratio of normal and anomalous samples is usually required [6]. Detecting anomalies in cases where defective samples are not available is usually solved via unsupervised or semi-supervised learning. With these approaches, models are trained using normal data only and therefore learn the state of normality that is later used to discriminate anomalies [7].

To train and evaluate the anomaly detection models, the classical datasets such as MNIST or CIFAR-10 were traditionally deployed. Using these datasets, one class was typically selected as a normal state and the other classes were anomalies. Although this approach provided a way to perform benchmarks of different methods, this case of anomaly detection did not represent real-world use cases very well. Recently, new domain-specific datasets were introduced to better represent the real anomaly detection use in practice [10], [11]. One of the most widely used dataset in the field of visual anomaly detection is the MVTec-AD [9], which contains 15 classes that represent different industrial applications (see fig. 1). Many new methods focused specifically on solving the problems presented in the MVTec-AD dataset have recently been released and already reached a very good results (e.g. [15], [23]).

Although the MVTec-AD dataset is very high quality, it

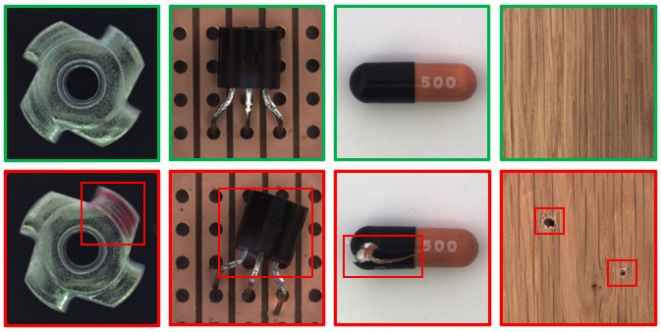


Fig. 1. Sample images from the MVTec-AD dataset (normal training samples are displayed in the first row, testing samples with highlighted anomalies are displayed in the second row).

does not represent some real-world applications very well. All classes in the MVTec-AD dataset are usually very well positioned in the image centre and rotation of objects is not very significant. In many real-world industrial applications, such conditions cannot be easily achieved. For this reason, in our previous work, we introduced the MPDD dataset [12], which is aimed at representing more complex real-world use cases.

In this paper, we present an evaluation of recent state of the art methods using the MPDD dataset and provide a benchmark of these methods in real-world applications. The main contribution of the paper can be sumarized as follows:

- We perform evaluation of recent state of the art visual anomaly detection methods in real-world industrial applications with complex conditions for data acquisition.
- We show that recent state of the art methods may not provide enough precision in some real-world industrial use cases.

The rest of the paper is structured as follows: in the next section, we provide an overview of recent state of the art visual anomaly detection methods. The next section describes the approach to the problem and methodology used for our experiments. In the last section, we present the results of our evaluation.

II. Related Work

Current visual anomaly detection methods can be divided into two following categories: reconstruction-based and feature extraction-based methods [26]. Reconstruction-based methods create reconstructions of the input data and anomalies are detected via L2-distance of the reconstructed and the input data. Examples of reconstruction-based methods include *DRAEM* [13] or *Skip-GANomaly* [14]. Even though the reconstruction-based methods provide satisfactory results in some use cases, in many scenarios the learned representation is also able to reconstruct anomalies, which degrades the anomaly detection precision [20].

Feature extraction-based methods recently proved to achieve higher precision in most visual anomaly detection tasks [9], [21], [23]. The main principle of the feature extraction-based methods is the detection of anomalies at the level of features extracted from the input data by a convolutional neural network (CNN) or a transformer model [15]. The feature extractor can be pretrained on the ImageNet dataset or can be trained from scratch [19], [25]. The anomaly detection itself is performed on the extracted features using a traditional machine learning algorithm (e.g. *k*-NN [23]) or a statistical model (e.g. multivariate Gaussian distribution [20]).

According to the MVTec-AD dataset benchmark, the best anomaly detection performance can currently be achieved by the PatchCore [23] method. In this method, feature extraction is done by a *ResNet* model pretrained on the ImageNet dataset. Features are extracted for all training samples and then placed into a memory bank, which is further reduced via the *CoreSet* algorithm [24]. Anomaly detection during testing phase is performed by the *k*-NN algorithm. A sample is considered anomalous when a defined threshold distance value from the neighboring training features is exceeded.

The next successful branch of anomaly detection methods deploy *normalizing flows* to model the true distribution of extracted features as it is expected that extracted features cannot always be represented well by a simple probability distributions (e.g. Gaussian distribution). Examples of *normalizing flows*-based methods include the *FastFlow* [15], *CS-Flow* [17], *AST* [18] or the *CFLOW-AD* [21].

Feature extraction models used in recent anomaly detection methods are usually pretrained using the ImageNet dataset. This approach is however not suitable for all anomaly detection methods as certain datasets may differ significantly from what can be encountered during training with the ImageNet dataset. Some methods therefore exploit transfer learning (i.e. start with a ImageNet-pretrained model and then train the model further on the custom dataset). We can find the above-mentioned transfer learning approach in the *CFA* method [16].

III. Evaluation Methodology

Our benchmark task is aimed at discriminating samples with defects in the MPDD dataset [12] that presents images from the field of metal parts manufacturing. Methods are trained using the non defective images only, we therefore follow the semi-supervised learning approach.

We assume a trainig set χ_M with M training images that can be denoted as x_m where $m = 1, 2, \ldots, M$. With the semi-supervised approach used during the evaluation, every x_m is an image without defect. Evaluation of detect detection for each method is performed using the testing dataset χ_T that contains N testing images denoted as x_n and labels y_n where $n = 1, 2, \ldots, N$ and $y_n \in \{0, 1\}$, (0 denoting normal sample and 1 anomalous sample). The exact number of training and testing samples can be seen in the table I.

As a part of the evaluation, we also perform defect localization, in which each method marks every pixel in the image as anomalous or non-anomalous. The precision of defect localization is measured using the standard AUROC metric.



Fig. 2. Sample images from the used MPDD dataset (normal samples without defects are displayed in the first row, anomalous samples with highlighted defects are in the second row).

TABLE I
NUMBER OF TRAINING, TESTING NORMAL AND TESTING ANOMALOUS IMAGES FOR ALL CLASSES IN THE MPDD DATASET USED FOR EVALUATION.

Class	Train	Test norm.	Test defect.
Bracket Black	289	32	47
Bracket Brown	185	26	51
Bracket White	110	30	30
Connector	128	30	14
Metal Plate	54	26	71
Tubes	122	32	69
Total	**888**	**176**	**282**

IV. EXPERIMENT

In this section, we describe the detailed setup used for our benchmark of selected anomaly detection methods. Our main goal was to select recent state of the art methods that currently provide the best precision on the MVTec-AD dataset and then test the selected methods on the MPDD dataset to observe the precision methods can provide in more complex real-world applications. Although the document [12] does not provide the name of the dataset – MPDD – explicitly, it can be viewed and also downloaded at the following URL[1].

We resized all input images to 256 × 256 pixels before processing the sample. We used the *ResNet-18* as a feature extraction backbone for the *PatchCore*, *FastFlow* and *CFA* methods. The *AST* and the *CS-Flow* methods deploy the *Efficient-b5* feature extractor. To measure the anomaly detection performance, we used the standard Area Under Receiver Operating Curve (AUROC) [27] (both for image-level and pixel-level detection).

[1]The dataset is available at: https://github.com/stepanje/MPDD

V. RESULTS AND DISCUSSION

We present evaluation results in the table II, where corresponding image-level and pixel-level AUROC scores are reported. The highest image-level AUROC scores – 96.1 % (averaged for all classes) – were achieved by the *AST* and the *CS-Flow* methods. Following these results, the second highest AUROC score in average is reported for the *CFA* method. According to the results, we can see that the best performance was achieved by the *normalizing flows*-based methods. In the pixel-level defect detection AUROC scores however, the *normalizing flows*-based methods did not provide a good precision compared to other methods. The highest pixel-level detection AUROC score was achieved by the *CFA* (97.6 %) and then the *PatchCore* (97.4 %) method. We also report the average MVTec-AD dataset scores for each evaluated method. According to our results, we observed that performance of all anomaly detection methods can be significantly different when used on the dataset that captures more complex conditions that can be encountered during manufacturing (e.g. MPDD).

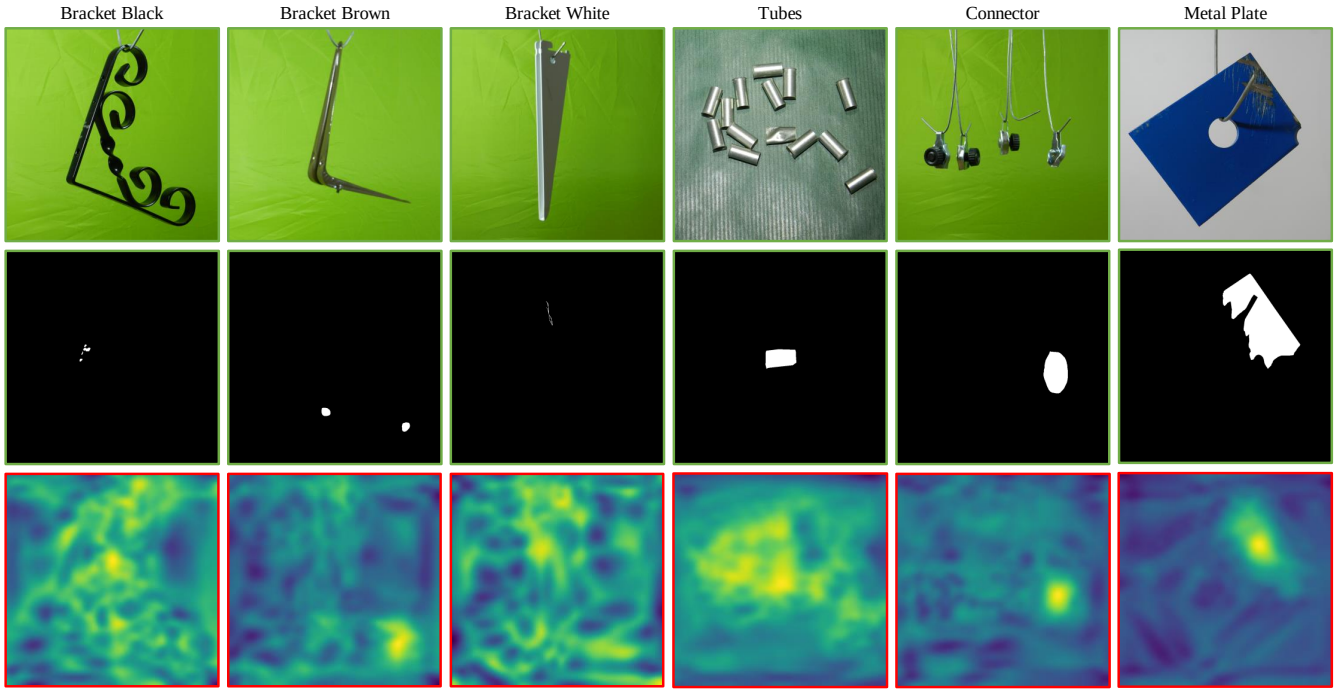


Fig. 3. Examples of resulting defect segmentation images of the PatchCore method (output images of the method are displayed in the bottom row, input samples with the corresponding ground truth images are displayed in the top and the middle row).

TABLE II
IMAGE-LEVEL AND THE CORRESPONDING PIXEL-LEVEL DETECTION AUROC FOR SELECTED STATE OF THE ART SEMI-SUPERVISED ANOMALY DETECTION METHODS AND THE MPDD DATASET.

Class	Avg.	PatchCore	FastFlow	CFA	AST	CS-Flow
Bracket Black	(88.3, 87.8)	(78.3, 93.0)	(70.3, 93.6)	(74.3, **96.2**)	(**98.1**, 68.4)	(95.6, –)
Bracket Brown	(86.9, 87.5)	(92.9, **99.0**)	(46.3, 86.4)	(95.4, 98.0)	(**100.0**, 66.6)	(**100.0**, –)
Bracket White	(86.2, 83.3)	(81.6, **99.1**)	(67.6, 65.6)	(82.7, 95.7)	(**100.0**, 72.7)	(99.3, –)
Connector	(88.6, 90.6)	(**100.0**, **99.6**)	(51.4, 76.9)	(96.7, 98.7)	(98.8, 87.2)	(96.2, –)
Metal Plate	(98.6, 91.1)	(**100.0**, 98.0)	(93.7, 80.2)	(99.9, **98.6**)	(**100.0**, 87.5)	(99.2, –)
Tubes	(79.5, 93.1)	(57.6, 95.9)	(85.1, 91.9)	(**88.5**, **98.6**)	(79.9, 85.9)	(86.2, –)
Avg.		(85.1, 97.4)	(69.1, 82.4)	(89.6, **97.6**)	(**96.1**, 78.1)	(**96.1**, –)
Avg. MVTec AD		(**99.6**, 98.2) [23]	(99.4, **98.5**) [15]	(99.3, 98.2) [16]	(99.2, 95.0) [18]	(98.7, –) [17]

When comparing results for different classes in the dataset, we can see the best precision – 98.6 % AUROC in average – was achieved on the *metal plate* class. This is probably due to the bigger size and good visibility of defects in the input images. The lowest precision was achieved on the *tubes* class in which all methods scored 79.5 % AUROC in average. This class captures the change in the number of parts, which have proven challenging for current anomaly detection methods.

VI. CONCLUSION

In this paper, we focused on the evaluation of recent state of the art visual anomaly detection methods in complex industrial defect detection tasks. Compared to most current datasets, we perform the evaluation in scenarios where conditions for data acquisition are complex and do not allow collection of images suitable for current visual anomaly detection methods (i.e. objects are always aligned, do not have significant position changes and the relative size of defects is high). The highest average image-level AUROC scores – 96.1 % – are reported for the *AST* and the *CS-Flow*. In terms of defect localization, the highest average pixel-level AUROC scores – 97.6 % – were achieved by the *CFA* method. High average pixel-level AUROC score – 97.4 % – is also reported for the *PatchCore* method.

According to our results, current visual anomaly detection do not provide sufficient precision in some evaluated defect detection use cases. Also, compared to the results that methods report on the MVTec-AD dataset, we observe singificantly lower averace AUROC scores in our use case. We believe more benchmarks of recent methods with more varied real-world datasets would help to further contribute to the development of new anomaly detection methods.

REFERENCES

[1] V. Chandola, A. Banerjee and V. Kumar. "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, 2009.

[2] G. Pang et al. "Deep learning for anomaly detection: A review," ACM Computing Surveys (CSUR), vol. 54, n. 2, 2021, pp. 1-38

[3] T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth and G. Langs, "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," in *Information Processing in Medical Imaging - 25th International Conference, IPMI 2017*, Boone, NC, USA, June 25-30, 2017, Proceedings, Vol. 10265 of Lecture Notes in Computer Science, Springer, 2017, pp. 146–157.

[4] S. Akçay, A. Atapour-Abarghouei and T. P. Breckon, "Ganomaly: Semi-supervised anomaly detection via adversarial training," in *Asian Conference on Computer Vision*, Springer, 2018, pp. 622–637.

[5] C. Huang, F. Ye, J. Cao, M. Li, Y. Zhang, C. Lu, "Attribute Restoration Framework for Anomaly Detection," in *arXiv:1911.10676*, 2020.

[6] D. Zhang, K. Song, J. Xu, Y. He and Y. Yan, "Unified detection method of aluminium profile surface defects: Common and rare defect categories," *Optics and Lasers in Engineering*, 126, 105936, 2020.

[7] L. Ruff et al. "Deep one-class classification," *International conference on machine learning*, PMLR, 2018.

[8] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.

[9] P. Bergmann, M. Fauser, D. Sattleger and C. Steger, "Mvtec ad - a comprehensive real-world dataset for unsupervised anomaly detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9592–9600.

[10] D. Tabernik, S. Šela, J. Skvarč, and D. Skočaj, "Segmentation-Based Deep-Learning Approach for Surface-Defect Detection," in *Journal of Intelligent Manufacturing*, May 2019.

[11] Y. Huang, C. Qiu, Y. Guo, X. Wang and K. Yuan, "Surface Defect Saliency of Magnetic Tile," in *2018 IEEE 14th International Conference on Automation Science and Engineering (CASE)*, 2018, pp. 612-617, doi: 10.1109/COASE.2018.8560423.

[12] S. Jezek, M. Jonak, R. Burget, P. Dvorak and M. Skotak, "Deep learning-based defect detection of metal parts: evaluating current methods in complex conditions," 2021 13th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), Brno, Czech Republic, 2021, pp. 66-71, doi: 10.1109/ICUMT54235.2021.9631567.

[13] V. Zavrtanik, M. Kristan and D. Skocaj, "DRÆM – A discriminatively trained reconstruction embedding for surface anomaly detection," in 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, 2021 pp. 8310-8319.

[14] S. Akçay, A. Atapour-Abarghouei and T. P. Breckon, "Skip-ganomaly: Skip connected and adversarially trained encoder-decoder anomaly detection," in *2019 International Joint Conference on Neural Networks (IJCNN)*, 2019, pp. 1–8.

[15] J. Yu et al., "FastFlow: Unsupervised Anomaly Detection and Localization via 2D Normalizing Flows", arXiv.org, 2022. [Online]. Available: *https://arxiv.org/abs/2111.07677v2*.

[16] S. Lee, S. Lee and B. Song, "CFA: Coupled-hypersphere-based Feature Adaptation for Target-Oriented Anomaly Localization", arXiv.org, 2022. [Online]. Available: *https://arxiv.org/abs/2206.04325v1*.

[17] M. Rudolph, T. Wehrbein, B. Rosenhahn, and B. Wandt, "Fully Convolutional Cross-Scale-Flows for Image-based Defect Detection," Jan. 2022. [Online]. Available: https://arxiv.org/pdf/2110.02855.pdf

[18] M. Rudolph, T. Wehrbein, B. Rosenhahn, and B. Wandt, "Asymmetric Student-Teacher Networks for Industrial Anomaly Detection," Jan. 2023.

[19] C. L. Li, K. Sohn, J. Yoon and T. Pfister, "CutPaste: Self-Supervised Learning for Anomaly Detection and Localization," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9664-9674, 2021.

[20] T. Defard, A. Setkov, A. Loesch, R. Audigier, "Padim: A patch distribution modeling framework for anomaly detection and localization," in *Pattern Recognition. ICPR International Workshops and Challenges - Virtual Event*, January 10-15, 2021, Proceedings, Part IV, Vol. 12664 of Lecture Notes in Computer Science, Springer, 2020, pp. 475–489.

[21] D. Gudovskiy, S. Ishizaka and K. Kozuka, "CFLOW-AD: Real-Time Unsupervised Anomaly Detection with Localization via Conditional Normalizing Flows," in *arXiv preprint arXiv:2107.12571*, 2021.

[22] J. Yi and S. Yoon, "Patch svdd: Patch-level svdd for anomaly detection and segmentation," in *Proceedings of the Asian Conference on Computer Vision*, 2020.

[23] K. Roth and L. Pemula, J. Zepeda, B. Schölkopf, T. Brox and P. Gehler, "Towards Total Recall in Industrial Anomaly Detection," in *arXiv:2106.08265*, 2021.

[24] O. Sener and S. Savarese, "Active learning for convolutional neural networks: A core-set approach," in *arXiv preprint arXiv:1708.00489*, 2017.

[25] S. Gidaris, P. Singh and N. Komodakis, "Unsupervised Representation Learning by Predicting Image Rotations", OpenReview, 2018. [Online]. Available: *https://openreview.net/forum?id=S1v4N2l0-*.

[26] Rippel, O., Mertens, P., & Merhof, D. (2021, January). Modeling the distribution of normal data in pre-trained deep features for anomaly detection. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 6726-6733). IEEE.

[27] T. Fawcett, "ROC graphs: Notes and practical considerations for researchers," in *Machine learning*, vol. 31(1), pp. 1-38, 2004.

Loudspeaker Enclosure with Directional Radiation of Frequencies below 100 Hz

Š. Skvaril
Faculty of electrical engineering and communication
Brno University of Technology
Brno, Czech Republic
xskvar08@vut.cz

***Abstract*— This work deals with development of loudspeaker unit whose radiation is directional at low frequencies. A concept that uses transmission line enclosure is proposed and verified with numerical simulations using K-wave toolbox for MATLAB. The simulations results for far field are verified by measurement at enclosure prototype.**

Keywords—sound, FDTD, loudspeaker, subwoofer, speaker enclosure

I. Introduction

Most of the state-of-the-art low-frequency loudspeakers (subwoofers) have omnidirectional radiation pattern due to a large wavelengths of low frequency sound waves in comparison with the membrane radius. This phenomenon leads to numerous problems like energy build-up on stage or disturbance of nearby residential areas. [1] These problems are solved by using multiple loudspeaker units in connection with processor unit which controls delay and phase shift of input signal of the individual loudspeakers. Those configurations are called subwoofer arrays, of which two are most common - End Fire array and Two-Element In-Line array (also called Gradient array). [1]

A. End Fire array

This configuration uses multiple (minimum of 3) loudspeakers placed in line behind each other with time delay added to the input signal of the individual loudspeakers. The delay is set in such manner that the loudspeaker generates the soundwave in phase with the sound coming from the loudspeaker behind it. This results in in-phase addition of sound waves in the front of the loudspeaker, but it cancels the waves in the back. However, this cancellation is not ideal and it is frequency-dependent as can be seen from Fig. 1 and Fig. 2.

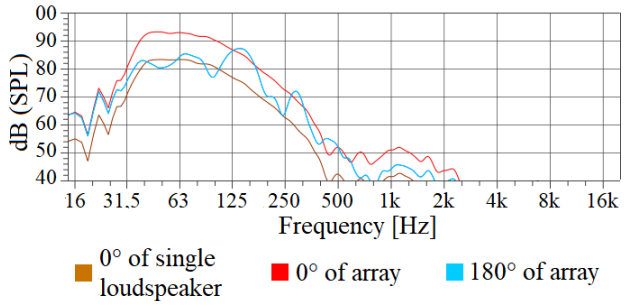


Fig. 1. Calculated frequency response of the End Fire subwoofer array

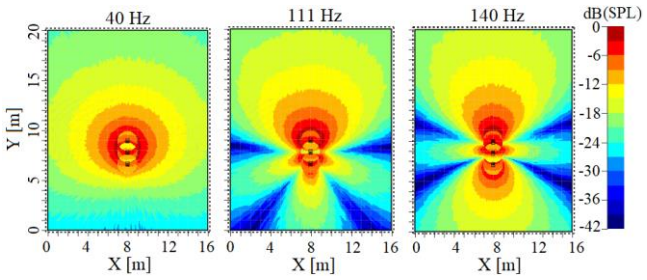


Fig. 2. Calculated polar response of the End Fire subwoofer array for three different frequencies

B. Gradient array

The gradient subwoofer array incorporates two loudspeakers in series behind each other. The input signal of the loudspeaker at the back has inverted phase and it is delayed by the amount of time in which the sound from the front loudspeaker propagates to the back loudspeaker. Thanks to the inverted phase, zero sound pressure is achieved in the back and thanks to the distance between loudspeakers, the phase is shifted and addition of waves is achieved in the front. However, this addition is frequency dependent as can be seen from Fig. 3 and Fig. 4.

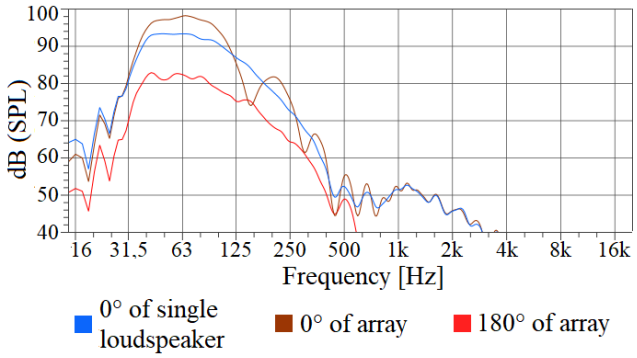


Fig. 3. Calculated frequency response of the Gradient subwoofer array

Few companies managed to create single loudspeaker units with directional radiation at lower frequencies, but, in the end, their functionality is based on Gradient loudspeaker array as in example published in [2]. The loudspeaker enclosure contains

more transducers with built-it amplifiers and processor unit which all increases the cost.

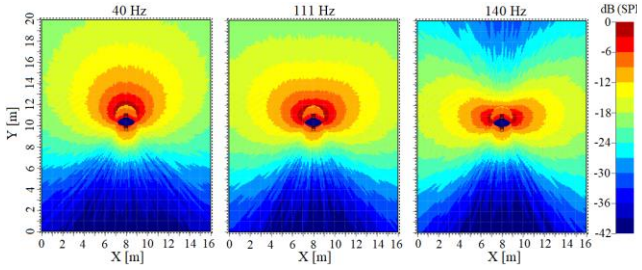


Fig. 4. Calculated polar response of Gradient array for three different frequencies

The aim of this work is to design a single loudspeaker unit, which achieves directional radiation patter at low frequencies without usage of a signal processing unit or multiple loudspeakers and is based on diploma thesis of the author [3].

II. Design of Enclosure

A. Principle of Operation

Chosen principle of operation was based on the gradient loudspeaker array since it creates perfect cancelation behind the loudspeakers. The frequency-dependent character of the addition in the front is not an issue, as low frequency loudspeakers often operate in narrow frequency band.

Proposed design of loudspeaker enclosure uses wave from the back side of the membrane, which has inverted phase to the wave generated by front side of the membrane. By incorporating the waveguide, this wave is delayed thus it acts as the second loudspeaker from the gradient subwoofer array.

B. Enclosure Design

When adding two harmonic waveforms that do not have the same phase the resulting sound pressure is calculated with (1) where p is resulting sound pressure, p_1 , p_2 are the amplitudes of added signals and φ_1, φ_2 their initial phase. [4]

$$p = \sqrt{p_1^2 + p_2^2 + 2p_1p_2cos(\varphi_1 - \varphi_2)} \quad (1)$$

The phase shift of two delayed signals is determined by (2) where f is the frequency of signals in Hz and τ is the time delay in seconds.

$$\varphi = f \cdot \tau \cdot 360 \quad (2)$$

By using (1) and (2) it can be found that the signals are added with a gain of 6 dB to 3 dB, if there is a phase shift from 0° to 90° and from 270° to 360° between them. At the same time, the back side of the membrane generates a signal that has phase shift of 180°. If this signal is delayed by another 180° compared to the original, a perfect addition and gain of 6 dB is reached. When shifted by another 90° it reaches the chosen limit of +3 dB. It follows that the total phase shift of the given signal must not be more than 270°. Therefore, if a frequency of 110 Hz whose wavelength is 3.127 m is taken as the upper limit a phase shift of 270° corresponds to a distance of 2.345 m. Since the sound wave must travel twice the length of the enclosure to the front part in front of the enclosure, it follows that the enclosure must not be longer than 1.173 m.

At the same time, if 35 Hz is taken as the lower limit, with the difference that the goal is not a 3 dB gain, but in order for no subtraction to occur, the phase shift of the two waves at this frequency needs to be at least 240°. The wavelength of this frequency is 9.829 m. The back side of the membrane generates a wave with phase shift of 180°. In order to avoid a subtraction in the front part, this signal must be shifted by another 80° at least, which results in a total path length of 2.184 m. Therefore the length of the enclosure must be at least 1.09 m.

With a maximum enclosure length of approximately 117 cm and a minimum length of approximately 109 cm, the final length of 115 cm was chosen.

III. Simulations

The FDTD method (Finite Difference Time Domain) in 3D was chosen for simulation of the loudspeaker in the designed enclosure. [5] K-wave toolbox for Matlab was used for the simulations. [6] The toolbox has the possibility to assign different values of density and speed of sound to each individual point in its computational grid.

The simulations showed, that the enclosure without damping behaves as half-closed tube, with strong resonance at 70 Hz. This resonance is unwanted as the amplification disbalances the ratio between the both signals which leads to the breakup of radiation pattern as seen in Fig.5.

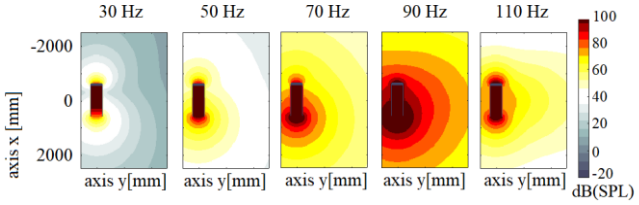


Fig. 5. Results of simulation of sound pressure distribution of empty enclosure

A damping material in the opening of the waveguide was used to suppress this resonance. However, the K-wave toolbox is not able to simulate sound absorption defined using the absorption coefficient as it is defined for the acoustic applications. [6] This problem has been overcome by inserting artificial material with speed of sound set to such low value that any sound entering this material will not be radiated from the material until the end of the simulation. Resulting calculations show that the resonance is dampened, and polar pattern is of the desired directionality as shown in Fig.6.

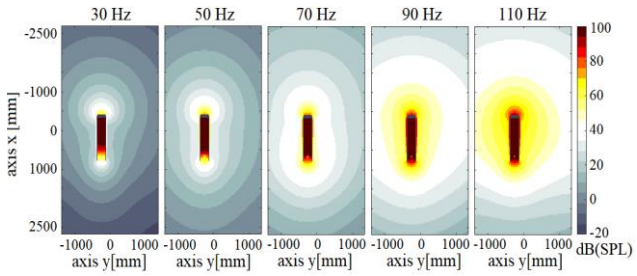


Fig. 6. Results of simulation of sound pressure distribution in damped enclosure

IV. Measurements

Based on the simulations, full scale prototype of the loudspeaker enclosure was built. The medium density fibreboard (MDF) with thickness of 18 mm and 15" loudspeaker Dayton Ultimax UM 15-22 was used to build the prototype. It was chosen because it has sufficient maximal membrane excursion limit, which makes it more resistant to damage during reproduction of the lowest frequencies. Picture of the prototype is shown in Fig. 7.



Fig. 7. Prototype of constructed experimental loudspeaker enclosure

A. *Near Field Measurements*

The measurements in close fields of the loudspeaker transducer and opening were conducted in first to see the effect of different amount of damping to the frequency response. This was crucial since as it is important that waves from both sides are of similar intensity so that their cancellation is effective. Four different amounts of damping were measured:

- None
- Light damping (5 cm of polyurethane foam)
- Medium damping (10 cm of mineral wool)
- Heavy damping (10 cm of PU foam and 5 cm of Re-bonded PU foam)

Result of these measurements are shown in Fig. 8 and Fig. 9. Strong resonance at 70 Hz can be seen according to the simulations but also on higher frequencies around 200 Hz. The effect of damping is evident as higher amount of damping reduces the module of the resonance but also changes its frequency to the lower values. It can be also observed that the module of resonance has effect on frequency dependence of the sound pressure generated by the membrane. The resonance is not fully eliminated even with heavy damping, but its middle frequency is lowered bellow 25 Hz which is out of the effective frequency range of the loudspeaker. The effect of damping on attenuation of the higher resonance is not very noticeable except their frequency shift. However, this resonance is not a problem, because it is also out of the effective frequency range of the loudspeaker.

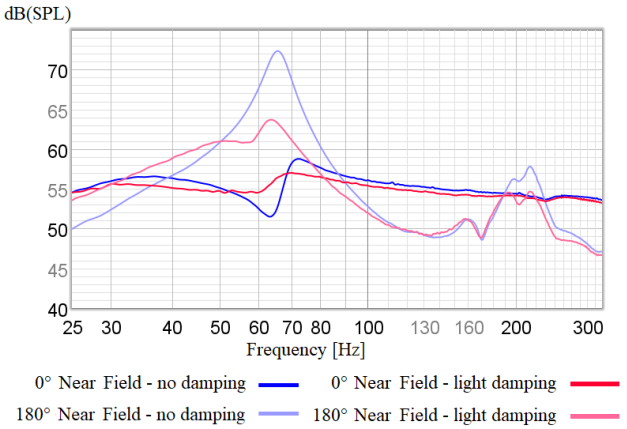


Fig. 8. Measurements in near field of empty and lightly damped enclosure

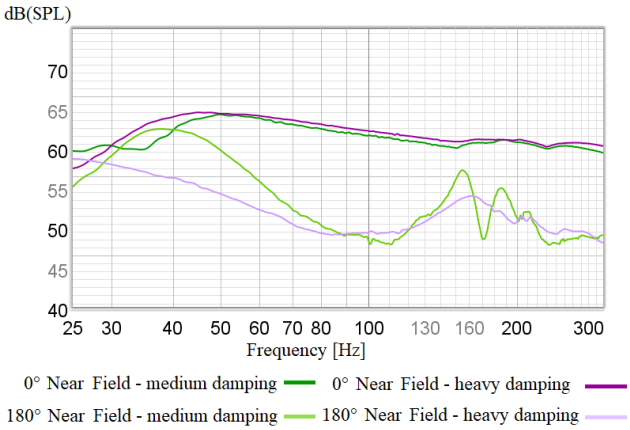


Fig. 9. Measurements in near field of medium and lightly damped enclosure

B. *Far Field Measurements*

After verifying that sufficient resonance attenuation is achieved with strong damping, The polar response of the prototype was measured in far field. The measurement of enclosure without damping showed resonance at 70 Hz according to the simulations alongside with the breakup of the radiation pattern as can be seen from Fig. 6 and Fig. 7.

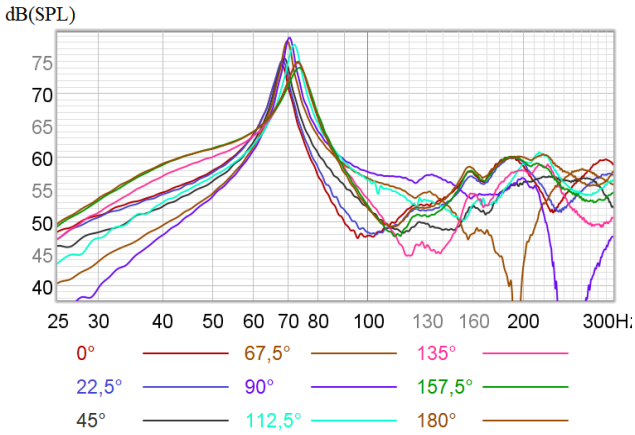


Fig. 10. Results of simulation of sound pressure distribution empty in enclosure

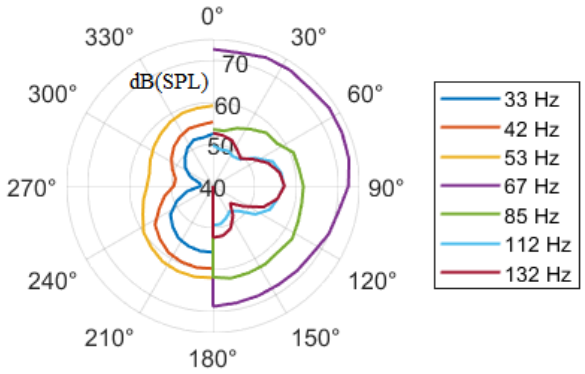


Fig. 11. Results of simulation of sound pressure distribution empty in enclosure

Another measurement was done with damping material placed in the enclosure. The results in Fig. 12 and Fig. 13 show directional characteristic with cardioid pattern up to 70 Hz with shift to super-cardioid pattern above 70 Hz. It can be seen that the unwanted resonance is successfully removed.

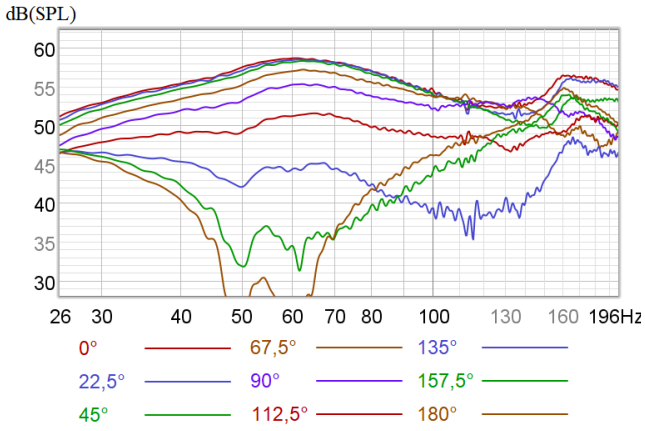


Fig. 12. Results of simulation of sound pressure distribution empty in enclosure

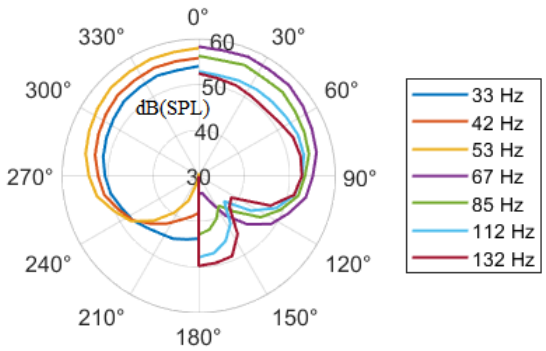


Fig. 13. Results of simulation of sound pressure distribution empty in enclosure

V. Conclusion

The proposed experimental loudspeaker enclosure was successfully simulated using the FDTD method. The prototype loudspeaker unit was constructed and measured in the far field afterwards. The measurement showed the better results of addition and cancellation than the simulation. The difference between the front and back side of the loudspeaker was around 4 dB. Measurements showed that the difference in acoustic pressure levels in the front and back of the loudspeaker is more than 12 dB in the band from 35 Hz to 85 Hz and more than 6 dB in the band from 28 Hz to 110 Hz. The differences between the simulations and measurements are probably caused by the inability of the K-wave toolbox to incorporate the absorption and by solution of this by using an artificial material as the absorber medium. This modification in calculations probably caused computational error.

As this concept has been proven as functional, it could be possibly extended to broader range of frequency spectrum. It could be possibly used in multi-way loudspeaker enclosure where each transducer would have its own waveguide designed for the required frequency range.

The drawback is that it is necessary to have the loudspeaker enclosure of sufficient length for lower frequencies which becomes undesirable for other applications than public address systems where loudspeaker placement is not limited by space.

REFERENCES

[1] MCCARTHY, Bob. Sound systems: design and optimization: modern techniques and tools for sound system design and alignment. Third edition. New York: Focal press, Taylor & Francis group, 2016. ISBN 978-0-415-73099-0.

[2] Kii Audio three. Kii audio [online]. Dingen: Kii audio, 2023 [cit. 2023-03-30]. Dostupné z: https://www.kiiaudio.com/electronics.php

[3] SKVARIL, Šimon. Reproduktorová soustava pro reprodukci nízkých kmitočtů se směrovým vyzařováním [online]. Brno, 2022 [cit. 2023-03-30]. Dostupné z: http://hdl.handle.net/11012/204989. Diplomová práce. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav telekomunikací. Vedoucí práce Jiří Schimmel.

[4] SERWAY, Raymond A. a John W. JEWETT. Physics for scientists and engineers: with modern physics. Tenth edition. Boston: Cengage, [2019]. ISBN 978-1-337-55329-2.

[5] Cox TJ, D'Antonio P. Acoustic Absorbers and Diffusers Theory, Design and Application, Third Edition. 2016. https://doi.org/10.1201/9781315369211.

[6] K-Wave - A MATLAB toolbox for the time-domain simulation of acoustic wave fields. K-Wave - A MATLAB toolbox for the time-domain simulation of acoustic wave fields [online]. Treeby, Cox, Jaros, c2010-2022 [cit. 2023-03-11]. Dostupné z: http://www.k-wave.org/index.php

FPGA Design with TDM Module for Real-Time Sound Acquisition from MEMS Microphone Array

Matej Liska
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
xliska19@vutbr.cz

***Abstract*—The paper summarizes the major contributions to the theory for the acquisition and processing of multiple channels of audio data for further real-time applications. It describes a particular type of Micro Electro Mechanical Systems (MEMS) microphone selected with Time Division Multiplexed (TDM) data format output. It also shows calculations and clock signals generation capabilities on Field Programmable Gate Array (FPGA) boards. The paper contains options for real-time data capture into the Matlab for further signal operations.**

***Index Terms*—MEMS, FPGA, TDM, MATLAB**

I. Introduction

The ever-expanding challenge of capturing real-time audio and speech for teleconferencing or broadcasting becomes even more difficult when hands-free lavaliere microphones are not an option. This is particularly true in situations such as live sports broadcasts or company meetings where large capture zones, such as rooms or halls, make microphone placement challenging. If it is not practical or financially feasible to have a microphone for every person, one solution is to cover specific areas near the edges or, if possible, areas in the middle from the top with simple shotgun microphones. However, these microphones only provide sporadic coverage, and their major limitation is fixed spatial selectivity in post-processing. This means that audio missed in real-time due to misaiming is permanently lost.

The main advantage of a microphone array is its electronic steerability, which allows multiple beams from different directions. This feature provides the same functionality in real-time signal processing as it does in post-processing. Additionally, arrays can be designed to have better directivity than with shotgun microphones. Although these features are attractive, they are not yet fully exploited in the field of audio acquisition, as in the case of using a spherical microphone array. There is a significant amount of well-documented research on acoustic sensors and spherical microphone arrays in [1] and [2].

With MEMS technology, it is possible to create controllable mechanical-movable structures using IC processing. Over the last four decades, MEMS technology has developed gradually [3] and has had a significant impact on the acoustic industry, particularly in the production of acoustic sensors. The introduction of MEMS microphones has opened up new possibilities for the development of microphone arrays, which, due to their size, price, and interfaces, offer significant competition to other technologies for the production of sound sensors.

A more thorough description of the selection of the acoustic sensor using MEMS technology is provided in section II. The significant advantage of the selected microphone is that it includes analog to digital conversion on one chip, thus creating a microphone with a digital output.

In section III, a detailed description of the FPGA board chosen for processing the output digital data from several microphone sensors is provided, along with the design created for this purpose.

The selection of the FPGA board model was based on its support for real-time connectivity with the MATLAB, as described in section IV.

This allowed for the processing of the received sound signals of individual microphones on the computer.

II. MEMS Microphone Selection

During the initial selection of the MEMS sensors, we relied on the implementation of the spherical array described in [4], which outlined three possibilities for the digital output of the MEMS sensors using interfaces *I2S, PDM and TDM.*

Although the interfaces are essentially similar, they primarily work with a one or two-channel audio signal. Only the TDM interface is capable of transmitting more than two channels of audio data. Therefore, the selection process was simplified to specific types of MEMS sensors with TDM interface output.

A. MEMS Sensor ICS-52000

The specific type of MEMS sensor was chosen based on availability in the market and its price.

The ICS-52000 is bottom port microphone with a digital TDM output. The complete ICS-52000 solution consists of a MEMS sensor, signal conditioning, an analog-to-digital converter, decimation and anti-aliasing filters, power management, and an industry standard 24-bit TDM interface.

The TDM interface which is detailed described in section II-B allows to connect an array of multiply microphones directly to digital processors, such as DSPs and microcontrollers, without the need for an audio codec in the system [5].

TABLE I
OPERATING CHARACTERISTICS.

Parameter	*Conditions*	*Value*	*Unit*
Supply Voltage		1.62 to 3.63	V
Directionality		Omni	
Sensitivity	1kHz, 94 dBSPL	-26	dBFS
Sensitivity Tolerance		1.0	dB
Signal-to-Noise Rat. (SNR)		65	dBA
Tot. Harm. Distor. (THD)	105 dBSPL	0.3	%
Frequency Response		50 to 20 000	Hz
Low Current Consumption		1.0	mA
Acoustic Overload Point		117	dB
Temperature Range		-40 to +85	°C

The ICS-52000 is available in a small 4 × 3 × 1 mm surface-mount package, and its electrical and acoustical operating characteristics are specified in Tab. I. All microphones in an array sample their acoustic signals synchronously, enabling precise array processing [5].

Its functional block diagram is shown in Fig. 1 and typical frequency response of ICS-52000 is shown in Fig. 2.

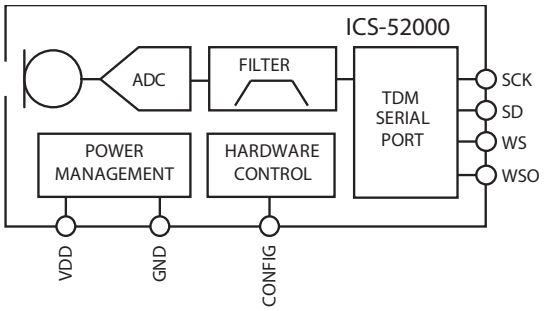


Fig. 1. Functional block diagram of ICS-52000 [5].

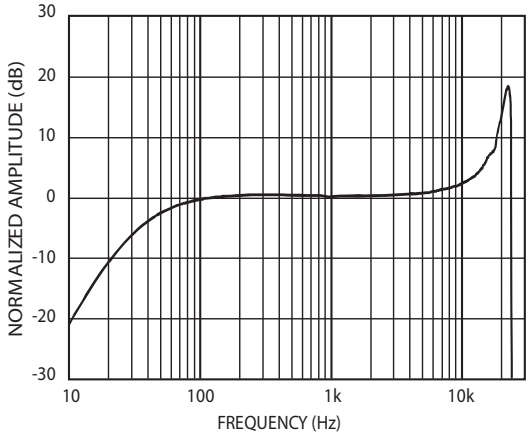


Fig. 2. Typical frequency response of ICS-52000 [5].

B. TDM Interface

The interface allows for the transmission of multiple channels of data on a single data line. The interface is similar to two-channel serial audio interfaces (I2S or PDM), with the exception that more channels, typically 4, 6, 8, or 16, are transmitted within a sample frame or sample period. As with the two-channel serial audio interface, the TDM interface is comprised of two control clocks, a frame synchronization pulse and serial clock, and the serial audio data line [6].

The signals of the TDM interface for the ICS-52000 microphone are labeled with the following abbreviations: *SCK* for input serial clock, *WS* for frame synchronization input pulse, *WSO* for output pulse, and *SD* for data output.

Serial data port's format is TDM, 24-bit, twos complement and up to 16 ICS-52000 microphones can be daisy-chained together on a single data bus. There must be 64, 128, 256 or 512 SCK cycles in each WS frame. Each microphone will output 24-bit data in subsequent 32-bit slots. The ICS-52000 will always be a follower on the TDM bus. The ICS-52000's WS clock input is sampled on the rising edge of SCK and the falling edge of WS can come anywhere before the start of the next frame [5].

The frequency of SCK will depend on the number of microphones in the system and should be calculated as:

$$n * 32 * WS, \tag{1}$$

where *n* is a power of two (2, 4, 8, or 16) equal to or greater than the number of ICS-52000s on the bus [5] and *WS* is frame synchronization pulse based on audio sample frequency (in our case 48 kHz).

Calculated SCK clock frequencies based on the number of microphones for our purpose is shown in Tab. II.

The ICS-52000 connected directly to the system's clock leader will output its data in the first TDM slot, the next microphone in the chain will output its data in the second TDM slot, and so on. The word select/word clock signals of the microphones in the system will be daisy-chained so that the clock leader drives WS of the first ICS-52000, whose WSO will drive WS of the second ICS-52000, and so on; the last ICS-52000 in the chain can leave WSO disconnected [5].

An illustration of these connections is shown in Fig. 3.

TABLE II
CALCULATED FREQUENCIES FOR SCK SIGNALS.

Number of ICS-52000	*SCK clock frequency [MHz]*
1-2	3.072
3-4	6.144
5-8	12.288
9-16	24.576

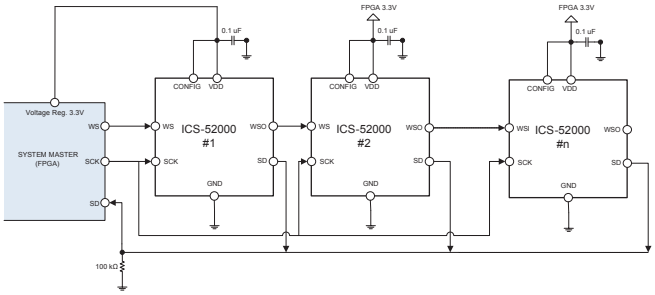


Fig. 3. Block diagram of system connection [5].

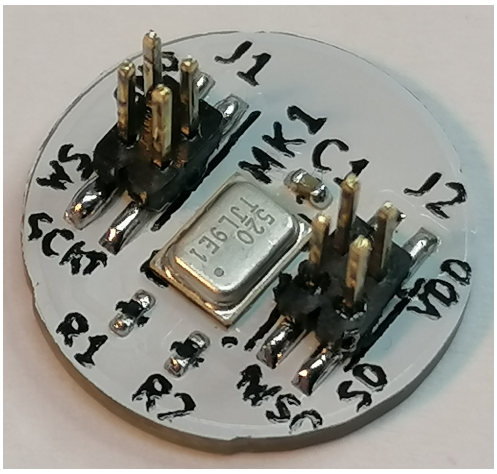


Fig. 4. ICS-52000 microphone PCB prototype [4].

The microphones on multiple PCBs will be attached to a spherical mount to form the final spherical microphone array.

The created prototype of PCB with MEMS ICS-52000 from [4] is shown in Fig. 4.

III. Main control device

The choice of a device for processing the digital output of the TDM interface from MEMS microphones was primarily focused on the possibility of processing this signal in real-time. The computing power of today's audio processors or microcomputers, such as Raspberry Pi or ADAU DSP, etc., was also important, but their availability on the market and price were decisive.

Due to the possibility of flexible design and connection with the MATLAB , we finally decided to implement processing using an FPGA board, which is described in detail in section III-C.

To process multi-channel audio data in combination with FPGA-based evaluation boards, an IP module was needed to transmit audio and configuration data. Section III-A contains more details about the design of the module for calculating clock signals, and section III-B describes the module for transmitting audio via the TDM interface.

A. Clock Module

As mentioned in section II-B, the TDM interface requires two clocking signals to sample and process the output audio data format. The first clock signal, named *WS*, is determined by the sampling frequency of the audio data - in our case, 48 kHz, and the *SCK* signal can be calculated using (1).

The main clock signal of the FPGA board, called *ddr_clock*, has a frequency of 100 MHz. To obtain a frequency of 48 kHz, it is necessary to use a frequency divider.

There are two options for creating a clock signal on the FPGA boards by division:

- using the *Clocking Wizard* IP core integrated on the FPGA,
- creating own entity and code in VHDL for dividing the input clock signal by an integer.

For both variants, it is necessary to divide the clock signal at least twice in order to obtain a common multiple of the numbers 100 MHz and 48 kHz.

1) Clocking Wizard: is an IP core that is available on FPGA boards and offers the ability to divide and create output frequencies using the Mixed Mode Clock Manager (MMCM). For our design, we used two IP blocks of the Clocking Wizard core. In the first block, the input frequency was ddr_clock (100 MHz), and the output frequency was set to 12.288 MHz. The next step was to connect this generated frequency as the input of the second block, from which we created frequencies for the individual numbers of microphones according to Tab. II. Using the MMCM Counter Cascading function described in [7], we created the resulting WS signal with a frequency of 48 kHz.

2) Clock Divider: is the second possibility of creating a frequency divider. However, to generate individual frequencies, the VHDL code needs to call the created block several times with the calculated divider values. The frequency divider works similarly to the Clocking Wizard. A description of how the code process works is shown in Listing 1.

```
1  --Dividing number is called DIV_FACT.
2  --Output clock signal is called ce_o.
3  signal cnt_div : unsigned :=(OTHERS=> '0' );
4  BEGIN
5  PROCESS (clk) BEGIN
6      IF rising_edge (clk) THEN
7          IF cnt_div = DIV_FACT THEN
8              cnt_div <= (OTHERS => '0');
9              ce_o <=  '1';
10         ELSE
11             cnt_div <= cnt_div + 1;
12             ce_o <=  '0';
13         END IF;
14     END IF;
15 END PROCESS;
```

Listing 1. Clock divider process.

B. TDM Module

To obtain data from connected MEMS microphones with TDM output, it was necessary to create a module that would implement the entire connection process.

At startup of the ICS-52000, the start of the frame sync (WS) signal should be delayed from the start of the serial clock (SCK) by at least 10 ms. This enables the microphone's internal circuits to completely initialize before starting the synchronization sequence with other microphones in the TDM array. The data is valid to use after the initial 262,144 SCK cycles (85 ms) [5].

The TDM module generates the clock signal WS, which is delayed by 15 ms based on the SCK clock signal. The frequency of the WS signal was set to 48 kHz, which equates to a time interval of approximately 0.02 ms. To ensure correct timing, the module waits for 4080 pulses of the WS signal, which takes approximately 85 ms, before starting to insert data. The rising edge of the 4081st pulse triggers the insertion of the captured data. The captured 32-bit slot of data is inserted into a 24-bit slot in MSB-first, LSB-last order, with the last 8 bits being discarded. The TDM module's operation is synchronized with the clock signals to ensure accurate data transfer, and its use of MSB-first, LSB-last order ensures that the most significant data is transmitted first, leading to better overall data quality.

The created entity of the TDM module including its signals is shown in Listing 2 and the simulation of one microphone with the clock signals is shown in Fig 5.

```
entity TDM_controller is
  generic(
    SYSTEM_CLK       :  integer   :=  100_000_000;
    SAMPLING_FREQ    :  integer   :=  48_000;
    SLOT_WIDTH       :  integer   :=  32;
    SAMPLE_BITS      :  integer   :=  24;
    NUMBER_SLOTS     :  integer   :=  16
   );
   port(
    clock            :  in std_logic;
    reset            :  in std_logic;
    enable_tdm       :  in std_logic;
    data_o           :  out std_logic_vector (...);
    --- TDM
    SCK_clock        :  out std_logic;
    WS_clock         :  out std_logic;
    data_i           :  in std_logic_vector (...)
   );
end TDM_controller;
```

Listing 2. TDM controller entity.

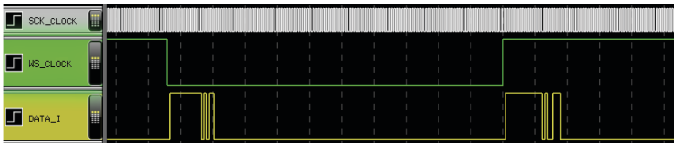


Fig. 5. Simulation of TDM interface signals.

C. FPGA Board Selection

The created entities for generating clock signals and retrieving data from the TDM interface can be utilized for designing on various FPGA boards. The selection of our specific board was primarily based on the possibility of connecting to the MATLAB for real-time connection and processing of captured data. Secondly, we required the chosen FPGA board to have a 100 MHz ddr_clock signal, at least three or more digital general-purpose input/output (GPIO) pins, and controllable buttons or switches for expanding the options.

The list of FPGA boards that support connection with MATLAB can be found in [8], from which the *Digilent Arty S7-25* board was chosen.

Digilent Arty S7-25 board features the Xilinx Spartan-7 FPGA and is the latest member of the Arty FPGA development board family from Digilent. Arty S7 was designed to be MicroBlaze ready and comes out of the box ready to use with the free Xilinx WebPack licensing with the Vivado Design Suite [9].

The main features of the Digilent Arty S7-25 board, which is shown in Fig. 6, are the following:

- 100 MHz external and internal clock,
- USB-JTAG programming circuitry,
- interaction and sensory devices (buttons, switches),
- powered from USB or any 7 V - 15 V source,
- 3.3 V output from voltage regulator,
- 32 total FPGA I/O,
- 256 MB DDR3L with a 16-bit bus @ 650 MHz,
- 16 MB Quad-SPI flash,
- three clock management tiles,
- 80 DSP slices and 23.360 logic cells [9].

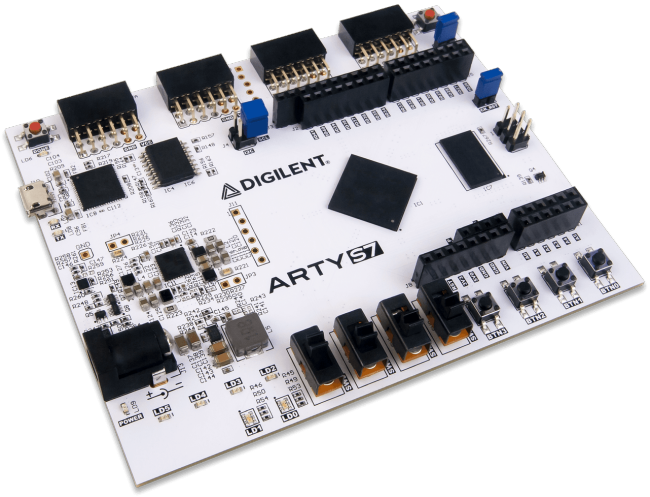


Fig. 6. Digilent Arty S7 board [9].

IV. MATLAB DATA CAPTURE

Based on the selection of the FPGA board that supports connection with the MATLAB , it was possible to obtain and capture data from MEMS microphones using the *FPGA Data Capture Component Generator* tool. A JTAG-USB cable was used for the connection.

To set up the FPGA Data Capture Component Generator tool, it is necessary to open it in MATLAB and set the input and output parameters. To integrate the FPGA Data Capture IP, we must open the Vivado project, edit and run the line_tdmmodule.tcl script. Then, we need to add all the FPGA Data Capture IP files to the project, define the path to the VHDL files, and add all the VHDL files to the project in the script shown in Listing 3. Finally, running the Capture Data tool provides the ability to display the data using a logic analyzer. After confirming the parameters, the FPGA generates an IP core for interconnection.

```
set projectDir ../../projectset projectName myname
set srdDir ../../src
set xdcDir ../../xdc
set memDir ../../memory_content
set bdDir ../../script/2022.1/bd
set dtDir ../../../hdlsrc_datacapture
add_file [glob $srdDir/ce_gen.vhd]
add_file [glob $srdDir/TDM_controller.vhd]
add_file [glob $dtDir/*.vhd]
```

Listing 3. Paths defining and adding VHDL files.

Detailed instructions how to proceed with the connection are described in [8].

V. CONCLUSION

This paper presents the possibilities of real-time data acquisition from MEMS microphones via the TDM interface using FPGA design. The main contribution of the paper is that the designed device uses the TDM interface, which can transmit data from up to 16 connected microphones on one line. The modules for acquiring and processing TDM data or generating clock signals can be used on various FPGA boards, creating a flexible device design that is not dependent on a specific hardware model. FPGA boards offer many GPIO pins that can be used to create multiple TDM buses, enabling the implementation of obtaining sound data from microphone arrays. We propose a possible solution for capturing and processing sound waves from multiple microphone arrays arranged at calculated nodes on a spherical mount for specific purposes. The possibility of connecting with MATLAB software for real-time sound capturing creates opportunities for various subsequent processing of the captured data.

The final system architecture of the designed device, which is shown in Fig. 7, consists of two main parts. The microphone unit provides signal processing of a specific channel, including its conversion into digital form, and the computer unit provides beamforming processing of the array of microphones as a whole using time domain filtering.

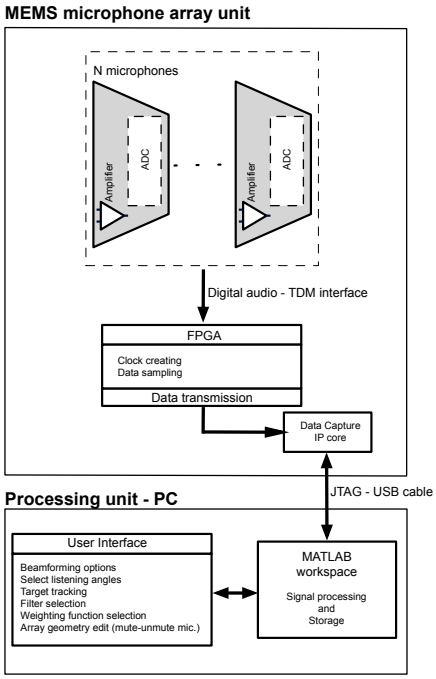


Fig. 7. System architecture.

Further research is necessary with respect to more efficient methods for real-time signal beamforming processing using a MATLAB with captured data from a spherical microphone array.

REFERENCES

[1] H. L. Van Trees, "Detection, Estimation, and Modulation Theory IV: Optimum Array Processing," John Wiley and Sons, Inc., 2003.

[2] D. P. Jarrett, E. A. P. Habets, P. A. Naylor, "Theory and Applications of Spherical Microphone Array Processing," vol. 9, Springer, 2017.

[3] J. Bryzek, "Impact of MEMS technology on society," in Sensors and Actuators A: Physical, vol. 56, 1996, pp. 1-9.

[4] D. Vagner, "Návrh sférického mikrofonního pole," B.S. thesis, Dept. Radio Eng., ČVUT, Prague, 2021.

[5] Invensense, Inc., "ICS-52000: Low-Noise Microphone with TDM Digital Output," DS-000121 data-sheet, rev 1.3, April 2017.

[6] Cirrus Logic, "Time Division Multiplexed Audio Interface: A Tutorial," AN301 User Manual, rev 1, Sept. 2006.

[7] Xilinx, Inc., "Clocking Wizard v6.0," PG065 LogiCORE IP Product Guide, April 2022, pp. 52-59.

[8] The MathWorks, Inc., "HDL Verifier™ Support Package for Xilinx® FPGA Boards," R2022b User's Guide, 2014-2022.

[9] Digilent, Inc., "Arty S7™," Reference Manual, Revised Oct. 2019. [Online]. Avaible: https://bit.ly/41OoL6r

Structure Analysis of the FRP Rebar Using Computer Vision Techniques

1st Juraj Lagiň
Institute of Concrete and Masonry Structures
Faculty of Civil Engineering
Brno University of Technology
Brno, Czech Republic
lagin.j@fce.vutbr.cz

2nd Šimon Bilík
Department of Control and Instrumentation
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
bilik@vut.cz

***Abstract*—In this paper we present a method to analyze the inner structure of the composite FRP rebar, namely the shift of the real centre of gravity with a respect to the geometrical centre of rebar and changes of cross-sectional characteristics. We propose an automated pipeline based on classical computer vision techniques and on the ratio between the glass fibres and epoxy filament in the analyzed cross-section to compute the shift vector of the real centre of gravity in respect to the geometrical centre together with the cross-section area and its principal moments. We discuss the achieved results over two cross sections in a different portion of the rebar and in the end, we suggest possible direction and improvements for our future work. We also made our code publicly available.**

***Index Terms*—FRP rebar, center of the gravity, image segmentation, image morphology**

I. Introduction

Composite reinforcement is a widely used alternative for reinforcing concrete structures. The FRP (fibre-reinforced-polymer) is known as high tensile strength reinforcement made by *pultrusion* technology. The reinforcement consists of fibres, matrix and surface treatment, which defines material characteristics of rebar. The straight FRP reinforcement is well-known and was described in many books or reviews, such as for example [1], [2]. However, for all-composite reinforced structures need to be designed indirect reinforcement as well. The first and oldest design code for bent FRP reinforcement was the JSCE 1997 [3]. The next generation of design codes is based on knowledge of this design approach - for example the American ACI 440.1R-15 [4] or Canadian CSA 806-12 [5] retrieved the approach of determination of ultimate tensile strength for bent FRP reinforcement, which depends on geometrical dimensions of rebar.

In fact, bent reinforcement is made by bending a straight reinforcement while the manufacturing process, which for thermosetting material leads to to a reduction of the ultimate tensile strength of bent rebar. The subject of this study is to further describe the influence of inner structure bent FRP reinforcement for its ultimate tensile strength. Our study uses computer vision techniques for describing differences in inner structure straight and bent FRP reinforcement.

II. Related Research

From point of inner structure view, the straight FRP rebars are quite widely described, as can be seen for example in [6], [7], [8]. The research of the inner structure of indirect FRP reinforcement was made just in scale visible by a bare eyed. In the studies [9], [10] could be clearly seen the effect of bending while manufacturing process, which leads to kinking fibres on the inner side of the bent portion of rebar and to a flattering of the cross-section of rebar as shown in Fig. 1. This effect causes a reduction of the ultimate tensile strength of bent FRP rebar, and it was already included in design code JSCE 1997 [3].

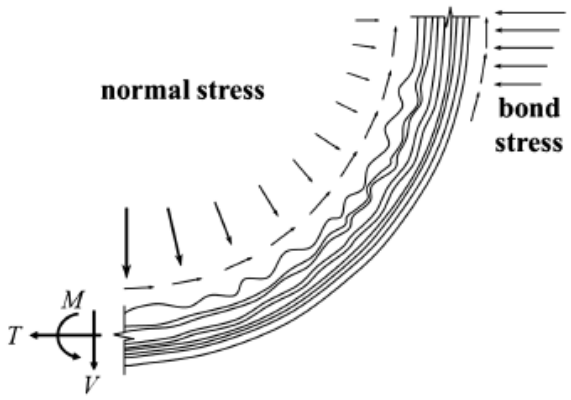


Fig. 1. Scheme of kinked kinked fibres along the bent portion of FRP rebar [10]

Computer vision techniques are today widely used in various applications including traffic [17], industrial inspection [15] or biology [14]. In this paper, we use classical morphological operations described for example in [13], Otsu method for segmentation threshold estimation [12] and the region properties analysis over a binary image.

III. Motivation

The main motivation of this research was to describe the differences of the inner structure in the straight and bent portion of FRP rebar. For the purpose of calculating internal forces

and stresses in structural engineering, structural mechanics is usually used. The summary is described in many literatures such as [16]. Normally, the ultimate tensile strength is defined as a force acting on a rebar area, expressed by the equation:

$$\sigma_u = \frac{N}{A}, \tag{1}$$

where σ_u is the ultimate tensile strength, *N* load force, and *A* is the area of the cross-section. As described in the previous capture, the manufacturing process of bent FRP rebar leads to a flattening circle cross-section. The changes cause translocation and shifting of load-bearing fibres to the inner surface of the cross-section. As is the fibre density several times higher, the real centre of gravity of the cross-section is shifted from the geometric centre of gravity. The distance between them is called eccentricity.

When the load force is not acting in the real centre of gravity, it creates an additional bending moment, calculated as:

$$M = N \cdot e, \tag{2}$$

where *N* is a load force and *e* is an eccentricity (distance) between the real and geometric centre of gravity.

If the cross-section is loaded by load force and bending moment, the tensile stress in the arbitrary point of the cross-section could be expressed as:

$$\sigma_u = \frac{N}{A} + \frac{M}{I_y}z - \frac{M}{I_z}y, \tag{3}$$

where *M* is the bending moment, I_y and I_z are moment of inertia of cross-section, z and y are the distance of the considered point from the real centre of gravity of cross-section. The moment of inertia can be calculated:

$$I_y = \int_A z^2\, dA, \text{ and } I_z = \int_A y^2\, dA, \tag{4}$$

where *A* is an area of cross-section and z is the distance of the considered point from the real centre of gravity of the cross-section.

Because the rotation of the cross-section in the image is random, the principal moments of inertia of the cross-section needs to be calculated. The principal moment of inertia also defines a principal axis *1* and *2*. The principal moment of inertia can be calculated as:

$$I_{1,2} = \frac{1}{2}(I_y + I_z) \pm \frac{1}{2}\sqrt{(I_y - I_z)^2 + 4{D_{yz}}^2}, \tag{5}$$

where D_{yz} is a deviation moment of the cross-section calculated as:

$$D_{yz} = \int_A yz\, dA, \tag{6}$$

In a cross-section loaded by a combination of tensile force and bending moment, it is necessary to find the point where the greatest tensile stress is applied. According to structural mechanics, this point can be found based on the neutral axis (see Fig 2), which is defined by the points 1_n and 2_n, calculated as:

$$1_n = \frac{i_2^2}{e_1}, \text{ and } 2_n = \frac{i_1^2}{e_2}, \tag{7}$$

where 1_n and 2_n are an inertia radius, expressed as:

$$i_1^2 = \frac{I_1}{A}, \text{ and } i_2^2 = \frac{I_2}{A}, \tag{8}$$

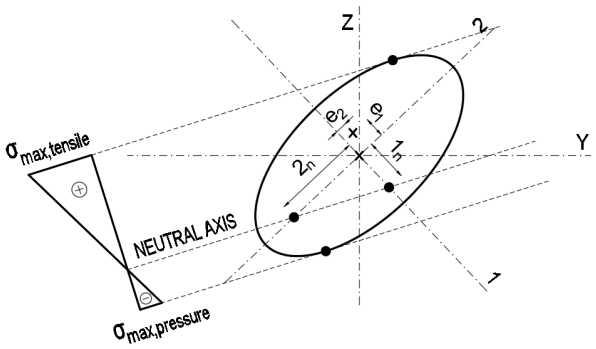


Fig. 2. Normal stress according eccentric tensile stress

In general, it could be assumed that cross-sections with a lower cross-sectional heights have a higher moment of inertia. This leads to higher tensile stress in critical locations of the cross-section in the same magnitude of the load. The aim of our study is to determine the changes in the area of cross-section, translation of the centre of gravity and influence of additional bending moment for the reduction of tensile strength of rebar.

IV. Materials and methods

In this section, we describe our data, its collection and the method used to find the shift of the center of the gravity based on the ratio between the fibers and filament masses.

A. Input data description

The samples for the study were cut from FRP reinforcement, made by E-glass load-bearing fibres, epoxy matrix and sand-coating surface treatment. The reinforcement was made by pultured method, nearly described in [11].

The radius of curvature was 30mm and the diameters of rebars *8mm* and *10mm*. Samples were taken from FRP rebar in straight portion (cut n.1) and in a bent portion (cut n.3) as shown in Fig. 3.

The samples used for our study (see Fig 4) were grinded by grinding discs with the roughness of surface *75um*, *35um* and *8um* after cutting. Then they were painted with a gold-plated layer and embedded in the vacuum chamber. Microscopical captures of the samples are shown in Fig. 5. The images of samples were obtained by electron microscope at 20 times enlarged.

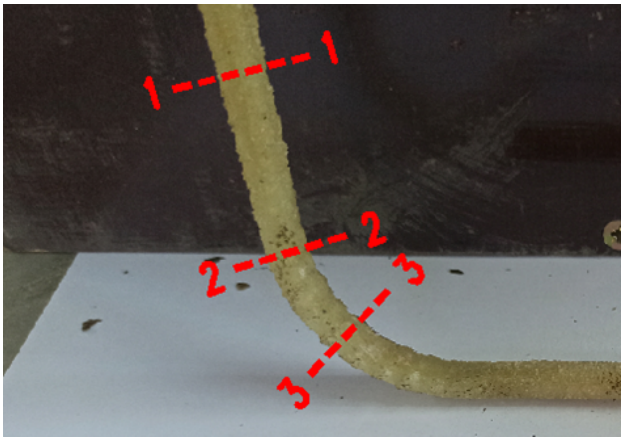


Fig. 3. Cut section of experiment samples

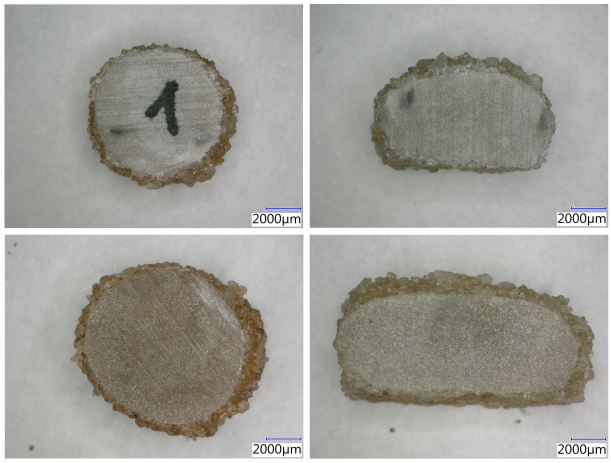


Fig. 4. Samples of experiment. From upper left to lower right: ⌀8_1-1, ⌀8_3-3, ⌀10_1-1, ⌀10_3-3

B. Center of the gravity detection

As a first step to detect the center of the gravity, we had to segment the image's ROI, which covers the central section of the cut without the bigger fibers on the side as could be seen in Fig. 5. The other areas, which should be removed are the darker grey spots and they are most likely caused by overheating the material while cutting and polishing.

To achieve such segmentation, we decided to use a cascade of two morphological reconstructions for a rough and fine segmentation. The rough segmentation as a first step is performed over the gray scale image using opening-close reconstruction with the disk kernel of a high diameter. It is followed by multiple Otsu-threshold and it results in a rough mask, which still contains parts of the side fibres.

The fine segmentation aims to remove the side fibres and the overheated parts on the cross section edges, which is done by unsymmetrical close-opening reconstruction with slightly stronger erosion using a disk kernel. This results into the fine mask and ROI segmentation, as shown in Fig. 6.

In order to compute the real centre of the gravity, we assume

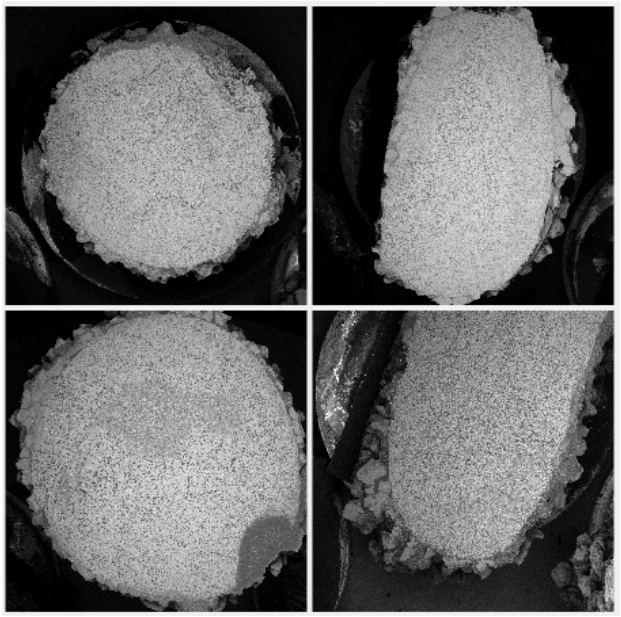


Fig. 5. FRP rebar cuts used in our experiment. From upper left to lower right: ⌀8_1-1, ⌀8_3-3, ⌀10_1-1, ⌀10_3-3

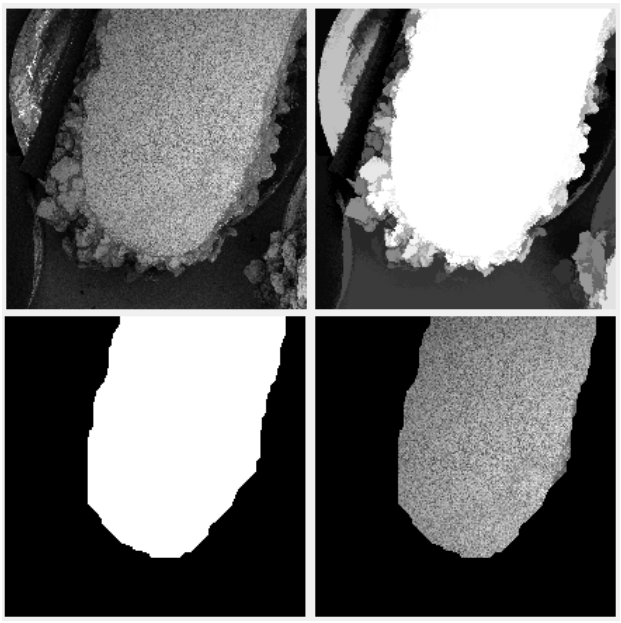


Fig. 6. Segmentation steps over the RFB rebar ⌀10_3-3 sample. From upper left to lower right: original image, rough segmentation, binary mask, segmented ROI

that the segmented ROI contains only glass fibres and epoxy filament. To separate them and the background, we again apply multiple Otsu-threshold and we assigned the corresponding values of density and tensile strength to the fibres and epoxy which are together with the ROI used to compute the weighted center.

To assign the corresponding values of densities, we compared the effect of a global and local approach. In the global

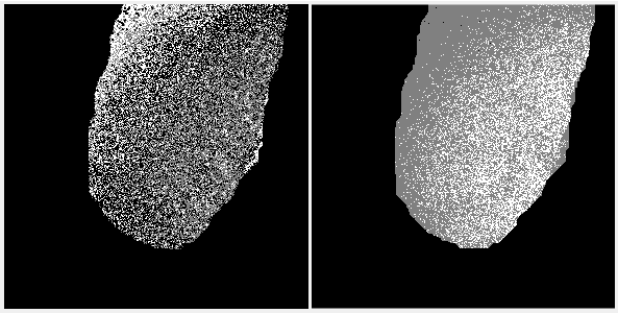


Fig. 7. Assigning the weights over the RFB rebar ⌀10_3-3 sample. 7x7 local assigning result on the left and global assigning result on the right.

approach, we considered the brightest pixels as glass fibres and all other pixels under the mask as the epoxy filament. This approach nevertheless did not bring satisfactory results, because the brightness of the cross-section changes with the position. Therefore we used a sliding window of size 7x7 under which we assign the glass fiber density to the pixel with a maximal value and the epoxy density to the pixel of minimal value. Pixels between this range are considered as invalid and we assign them a zero weight. A comparison of the global and local approach is shown in Fig. 7

Geometrical centre is computed in a similar way over the binary mask without assigning any weights. In the end, we compute a shift vector between the geometrical and real centre of gravity, which is also displayed in the original image. To compute the distance in *mm* from the distance in *pixels*, we use a scale contained in the header of the microscopic images.

C. Momentum and cross-section points calculation

To compute principal moments of inertia I_1, I_2 together with the intersections of their axis and eccentricity line with the cross-section's mask, we had to first compute the moments I_z, I_y, deviation moment D_{yz}.

To compute I_z, I_y and D_{yz}, we iteratively apply morphological erosion with the size of one pixel to the given mask and compute difference between the previous and eroded image. For every pixel of such obtained contour, we compute its contribution to the equations 4 and 6 until there were no remaining non-zeros pixels in the difference image.

Using those values, we compute the principal moments of inertia I_1, I_2 described with equation 5 and rotation of their axes with the respect to the centre of the gravity axes. Using the I_1, I_2 axis direction, we compute the intersection points with the given mask and rotate their coordinates with the respect to the I_1, I_2 axis. The same is applied for the intersection of the eccentricity line in the direction of the centre of the masses shift vector.

To find the intersection point coordinates, we test if the mask pixel under the given line has value of zero or one. After we find an edge, we save the pixel coordinates as an intersection point in a given direction.

V. Experimental results

As a first part of our experiment, we compare the effect of the sample's shape on the resulting shift vector. Obtained results are shown in Tab. I, vector size and angle are shown with respect to the Cartesian coordinate system with the beginning in the lower left image corner. The density of the glass fibres was set as 2600 kg/m ^{3} and the density of the epoxy filament as 1300 kg/m ^{3}.

TABLE I
Resulting shift vectors of the analyzed samples

Sample	Vector size $[mm]$	Vector angle $[^\circ]$
⌀8_1-1	0.024	-55.90
⌀8_3-3	0.248	-88.27
⌀10_1-1	0.098	-55.51
⌀10_3-3	0.211	56.07

We could see that the curvature of the rebar significantly affects the resulting shift vector size and angle in the comparison with the non-bended cross-sections - the resulting shift of the elliptic cross-sections *3-3* could be considered as approximately 10x higher than by the circular cross-sections *1-1*. Computations of the geometrical centre of the sample *10_3-3* seems to be misleading, because the sample was too big to fit in the microscope completely. The resulting center of the mass shifts, axes and intersection points of the analyzed samples are shown in Fig. 8.

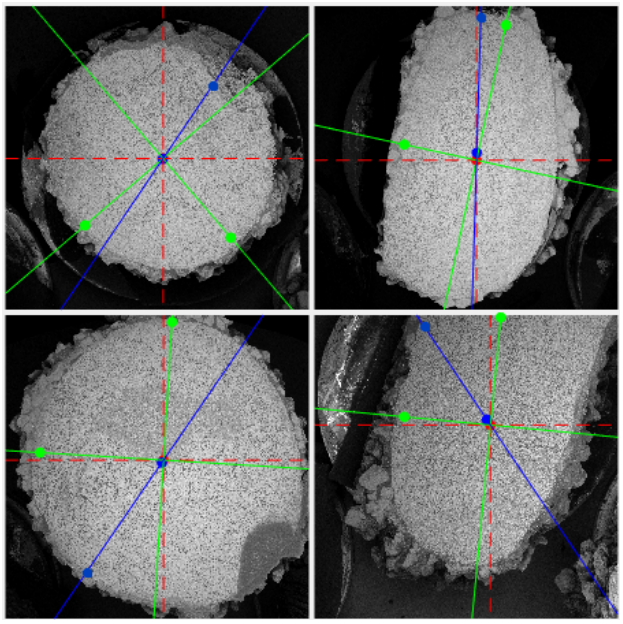


Fig. 8. FRP rebar cuts, their centers of the mass (red - original, blue - shifted), axis (red - original, green - principal moments) and cross-section points. From upper left to lower right: ⌀8_1-1, ⌀8_3-3, ⌀10_1-1, ⌀10_3-3

Predicted and obtained cross-sectional characteristics compared in the second part of our experiment are shown in Tab II:

where $A_{s,p}$ is a predicted area of cross-section according to declared diameter by the manufacturer, and A_s, I_1, I_2 are

TABLE II
RESULTING CROSS-SECTION CHARACTERISTICS

Sample	$\mathbf{A_{s,p}}[\mathbf{mm^2}]$	$\mathbf{A_s}[\mathbf{mm^2}]$	$\mathbf{I_1}[\mathbf{mm^4}]$	$\mathbf{I_2}[\mathbf{mm^4}]$
⌀8_1-1	50.27	37.45	118.60	106.31
⌀8_3-3	50.27	38.22	208.82	66.45
⌀10_1-1	78.54	58.41	311.91	248.31
⌀10_3-3	78.54	-	-	-

TABLE III
THE TENSILE STRESSES IN CROSS-SECTIONS

Sample	$N_{u,exp}$ [kN]	$\sigma_{u,exp}$ [MPa]	$\sigma_{u,C}$ [MPa]	$\sigma_{u,A}$ [MPa]	$\sigma_{u,E}$ [MPa]
⌀8_straight	68.02	1353.28	1815.94	1856.02	1862.16
⌀8_bent	26.58	528.71	694.14	757.21	840.59

obtained results of the area and principal moment of inertia. As part of the control, comparisons of the computed moment of inertia of a circular cross-section (but also simpler rectangular cross-sections) of the identical cross-sectional area were also processed. The observed deviations are low and comparing the measured and computed results, our proposed method could be considered to be functional.

From the obtained results shown in Tab II, we can see a clear decrease in the effective cross-sectional area, which nevertheless may be partly caused by a deviation of our method. The most significant difference can be observed by the moment of inertia, caused by the change in cross-sectional shape.

In our previous experiments carried out on a larger number of specimens, we investigated the strength of straight and bent rebars. The resulting tensile strength $\sigma_{u,exp}$ was proved to be related to the declared cross-sectional area. Based on our obtained values shown in Tab II, it was possible to apply the experimentally determined average failure force to the 8mm diameter members and then to calculate the stresses in points C, A, E etc. (see Fig 9).

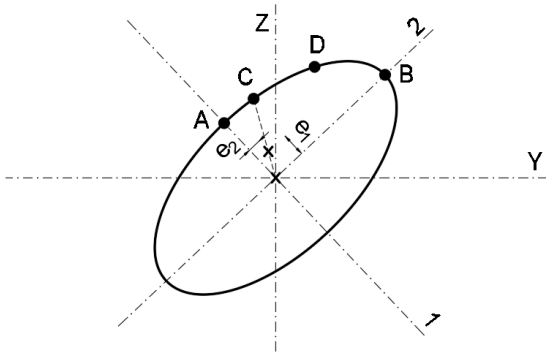


Fig. 9. Scheme of points considered for calculation of tensile stress

The tensile stresses $\sigma_{u,exp}$ and $\sigma_{u,C}$ were calculated by equation 1. In points A and E is also a bending moment, so the tensile stresses $\sigma_{u,A}$ and $\sigma_{u,E}$ could be calculated by equation 3. Those calculations were performed only on the samples *⌀8* because the sample *⌀10_3-3* was not fully displayed. The obtained results are summarized in Tab III:

The results clearly show an increase of the tensile stresses at the centre of gravity of the cross-section compared to the tensile stresses when considering the manufacturer's declared reinforcement diameter. Considering the applied eccentricity, an increase in the tensile stresses at the extreme point *E* was 2.55% for straight rebar and 21.10% for bent rebar. If we compare the maximum tensile stress in point E at straight (*⌀8_1-1*) and bent (*⌀8_3-3*) sections of the rebar, we can observe a difference of 54.86%. This difference indicates that the bending moment is not the only factor causing the reduction of the tensile strength of rebar and as was mentioned in the introduction of the paper, the overall reduction in tensile strength is also caused by the shrinkage of the fibres on the inner side of the bend, which causes that the fibres to lose their effectiveness. However, based on the results obtained in previous experiments, it can be argued that from the total experimentally determined tensile strength reduction of the reinforcement of 60.98%, approximately 6.12% can be attributed to the effect of bending moment and change in cross-sectional characteristics. Our code is available at [1].

VI. CONCLUSION

In this study, we analysed microscopic samples of the FRP rebar cuts in order to compute the centre of the gravity shift vector with the respect to the rebar's geometrical centre. To achieve this, we presented a method how to automatically segment the rebar's ROI and how to compute shifted centre of gravity based on the given weights and ratio between the glass fibres and epoxy resin. Based on the obtained results was found the following:

- We proved a decrease of the cross-section area based on a comparison of the experimentally obtained value and area computed from a manufacturer-declared diameter.
- We observed that the centre of gravity translations for samples *⌀10_1-1* and *⌀8_1-1* were insignificant and that they might be caused by a deviation of our method. Sample *⌀10_3-3* shows a translation of the centre of gravity in the direction of the outer surface. However the sample was not fully displayed, so the geometrical centre of gravity is not determined correctly;
  For the sample *⌀8_3-3* the translation of the centre of the gravity to the inner surface of the cross-section is several times higher in comparison to *⌀8_1-1*. It suggests, that the fibres of rebar are concentrated on the inner surface of the cross-section;
- In the case of straight reinforcement, the additional bending moment caused an increase in tensile stresses of 2.55%, which seems to be consistent with the measurement deviation mentioned above. However, for the bent

[1] https://github.com/boortel/FRP-Rebar-Structure-Analysis

rebars there was an increase up to 21.10% and it can be argued that the effect of the bending moment on the reduction of the tensile strength of the bent reinforcement has been demonstrated;

- By our previous experiments, the reduction of tensile strength was determined to be 60.98%. By comparing the tensile stresses at the point E of the cross-section of straight and bent reinforcement, it can be argued that the effect of bending moment on the overall tensile strength reduction is approximately 6.12%;
- Weight assignment shown at Fig. 7 should be further improved in such way, that we decrease the number of the pixels with the zero weights;
- Our experiments were performed on a low amount of samples and for determination of influence for reduction of tensile strength, experimental dataset should be increased;
- As a future experiment with a bigger dataset, we suggest to use machine learning methods as semantic segmentation to automatically segment the ROI and to assign the weights. Also, a method should be developed for the automatic calculation of the neutral axis, based on which the point with the maximum tensile stress along the cross section can be determined.

ACKNOWLEDGMENT

The completion of this paper was made possible by grant No. FEKT-S-23-8451 - "Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence, automation and measurement" financially supported by the Internal science fund of Brno University of Technology and by the grant "FW01010520 - Development of bent composite reinforcement for environmentally exposed concrete constructions" conducted on Institute of Concrete and Masonry Structures, Faculty of Civil Engineering - Brno University of Technology.

REFERENCES

[1] L. Hollaway, "Polymer Composites for Civil and Structural Engineering", 1993, 10.1007/978-94-011-2136-1.

[2] O. Abbood, S. Odaa and K. Hasan, "Properties evaluation of fiber reinforced polymers and their constituent materials used in structures - A review", *Material Today: Proceddings*, 2021, 10.1016/j.matpr.2020.07.636.

[3] Japan Society of Civil Engineers, "Recommendation for Design and Construction of Concrete Structures using Continuous Fiber Reinforcing Materials", *Research Committee on Continuous Fiber Reinforcing Materials*, 1997

[4] America Comittee Institute, "ACI 440.1R-15: Guide for the Design and Construction of Structural Concrete Reinforced with Fiber-Reinforced Polymer (FRP) Bars", 2015

[5] Canadian Standards Association, "CSA S806-12: Design and construction of biulding structures with fibre-reinforced polymers", 2012

[6] M. Robert ano B. Benmokrane, "Physical, Mechanical, and Durability Characterization of Preloaded GFRP Reinforcing Bars", 2010, 10.1061/ASCECC.1943-5614.0000104.

[7] A. Maligno, N. Warrior and A. Long, "Finite element investigations on the microstructure of fibre-reinforced composites", *Express Polymer Letters*, 2008, 10.3144/expresspolymlett.2008.79.

[8] I. Sevostianov, and V. Kushch, "Effect of pore distribution on the statistics of peak stress and overall properties of porous material", *International Jounral of Solids and Structures*, 2009, 10.1016/j.ijsolstr.2009.09.002.

[9] E. A. Ahmed, A. K. El-Sayed, E. El-Salakawy, B. Benmokrane, "Bend Strength of FRP Stirrups: Comparison and Evaluation of Testing Methods", *Journal of Composites for Construction*, 2010, doi.org/10.1061/(asce)cc.1943-5614.0000050.

[10] C. Lee,,M. Ko, Y. Lee, "Bend Strength of Complete Closed-Type Carbon Fiber-Reinforced Polymer Stirrups with Rectangular Section", *Journal of Composites for Construction*, 2014, doi.org/10.1061/(asce)cc.1943-5614.0000428.

[11] J. Lagiň, J. Prokeš, F. Girgle, V. Kostiha, D. Vašátko, O. Bezděk, O. Sedláček, P. Štěpánek, "The Effect of the Bent of FRP Rebar in its Short-Term Mechanical Propreties", *ICBMPT2023*, May 2023,

[12] Otsu, N. A threshold selection method from gray-level histograms. *IEEE Transactions On Systems, Man, And Cybernetics*. **9**, 62-66 (1979)

[13] Sonka, M., Hlavac, V. & Boyle, R. Image processing, analysis, and machine vision. (Cengage Learning,2014)

[14] Bilik, S., Kratochvila, L., Ligocki, A., Bostik, O., Zemcik, T., Hybl, M., Horak, K. & Zalud, L. Visual diagnosis of the varroa destructor parasitic mite in honeybees using object detector techniques. *Sensors*. **21**, 2764 (2021)

[15] Bergmann, P., Löwe, S., Fauser, M., Sattlegger, D. & Steger, C. Improving unsupervised defect segmentation by applying structural similarity to autoencoders. *ArXiv Preprint ArXiv:1807.02011*. (2018)

[16] Carpinteri, A. , Structural Mechanics Fundamentals *CRC Press, Taylor & Francis Group*. (2010)

[17] Zemcik, T., Kratochvila, L., Bilik, S., Bostik, O., Zemcik, P. & Horak, K. Performance Evaluation of CNN Based Pedestrian and Cyclist Detectors On Degraded Images. *International Journal Of Image Processing (IJIP)*. **15**, 1 (2021)

Direct Piezoelectric response miscroscopy of Bismuth ferrite thin films

Denis Misiurev, Holcman Vladimír
Department of Physics
Brno University of Technology
Brno Czech Republic
xmisiu00@vutbr.cz, holcman@vut.cz

***Abstract* —** The paper will provide a brief overview of one of the most popular multiferroic material bismuth ferrite, which was under intense study over the past decades. The reasons why the material earned such a great popularity in recent years will be described in detail. Piezoresponse Force Microscopy (PFM) would be used in order to investigate ferroelectric and magnetic properties of the material. In addition to ferroelectric and magnetic properties, the PFM method reveals the presence of piezo–active regions with different orientation based on their colorations. Samples of 100 nm $BiFeO_3$ material were deposited by using pulsed laser deposition technique on multiplayer substrates $Pt/Ti(TiO_2)/Si$ will undergo investigation by PFM. The investigation of produced samples is necessary to determine the topology of produce samples in terms of surface roughness across the investigated surface and potential effect of topology of deposited samples on layout of domains.

Keywords — Nanomaterials, multiferroics, ferroelectrics, bismuth ferrite, piezoresponce, polarization.

I. INTRODUCTION

Bismuth ferrite thin films have received significant attention in recent years due to their potential for various technological applications, particularly in the field of complex electronics. The multiferroic properties of bismuth ferrite, such as this high dielectric constant, temperature resistance, piezo response, make it an attractive material use in film films. Bismuth ferrite thin films have a lot of potential for using a variety of technological applications, particularly in the fields of electronics and energy harvesting. In the field of electronics, bismuth ferrite thin films have been exported for the use of a variety of devices, including high–frequency filters, resonators and oscillators.

Samples of 100 nm BFO $Pt/Ti(TiO_2)/Si$ thin films are deposited by Pulsed Laser Deposition (PLD) technique [1], which has been proven to be a valuable technique for deposition of high quality and homogeneous film films for various applications in material science and engineering.

$Pt/Ti(TiO_2)/Si$ substrate refers to three–layer material system, what consist of platinum on top of the $Ti(TiO_2)$ on the silicone [1]. The PT monolayer serves as a catalyst for various chemical reactions and provide electrical contact and ensures integration of $BiFeO_3$, while $Ti(TiO_2)$ layer acts as a barrier preventing diffusion of Si [1]. The Si layer (substrate) provides mechanical stability and acts as a based material for the multilayer structure [2].

Further research is needed to fully understand and optimize the properties of bismuth ferrite thin films, but they have an enormous potential to be utilized in the development of new technology in the future.

II. DEFINITION OF MULTIFERROIC MATERIALS

Multiferroic materials are materials, what exhibit both ferroelectric and magnetic properties. These properties arise from the same microscopic origin, which is the interaction between the magnetic spin and the electric polarization of the material [3]. One of the unique features of multiferroic materials is their magnetoelectric coupling, which refers to the influence of an applied magnetic field on the electric polarization and vice versa [3]. This allows for the manipulation of the magnetic and electrical properties of the material through a single stimulus. multiferroics can be broadly classified in the two subcategories: weak and strong coupling multiferroic materials [3]. Weak coupling of multiferroics have a weak magnetoelectric coupling, which means, what the magnetic and electrical properties are only weakly correlated (CuO, $MnWO_4$, $BiMnO_3$, $CoCr_2O_4$). Strong coupled multiferroics, on the other hand, have a strong magnetoelectric coupling and their magnetic and electric properties are closely related ($BiFeO_3$, $YmnO_3$, $TbMmO_3$, $EuTiO_3$) [4]. The study of multiferroics is a rapidly growing field and the researchers are actively working to develop new 2D multiferroic materials and understand their properties [4]. 2D multiferroics have attracted significant attention in recent years due to their ability of creating local ferroelectricity and huge potential to be used in a variety of technological applications, including data storage, energy convention and energy harvesting.

III. BISMUTH FERRITE MATERIAL

Bismuth ferrite is a multiferroic material, which means it is exhibits both magnetic and ferroelectric properties [4,5]. Bismuth ferrite exhibits rhombohedral structure [5], which is common among ferroelectric materials, and its ferroelectricity arises from displacement of oxygen ions in the crystal lattice [5]. Bismuth ferrite is also known for its large electrical polarization ~(100–120) μCcm^{-2}, high temperature stability, and high resistance to electric fatigue makes it a promising material for use in a high temperature, high power electronic devices. Bismuth ferrite has a complex magnetic structure with both

Internal Grant Agency (FEKT–S–23–8228)

antiferromagnetic and ferromagnetic components [6]. Its magnetic ordering temperature (Neel point, T_N) is approximately 640K [6] with (T_C) Curie temperature being 1120K [7], which is relatively high for a magnetic oxide. The magnetic properties of bismuth ferrite can be manipulated by applying the external magnetic field, electric field or temperature, make it extremely flexible material for magnetic application [7]. One of the most prominent magnetic properties of bismuth ferrite is large magnetic anisotropy, which is described by its preferentially alignment magnetic moments alongside of its specific crystal axis [8]. This property is extremely important for magnetic data storage, as it allows for stable magnetic domains to be formed and maintained. Bismuth ferrite is also demonstrate large magnetic coupling between its magnetic and ferroelectric domains, which is resulting in magnetoelectric coupling [8]. This means, the magnetic properties of the material can be influenced by applied electric field and vice versa. In addition to large electromagnetic compelling, the material also exhibits a spin cycloid orientation [8]. The spin cycloid orientation is a type of magnetic texture, what can be observed in multiferroic materials, which is characterized by spiral arrangement of magnetic spin in the material. Particularly in bismuth ferrite, the spin cycloid is a result of the interplay between the magnetic and electric properties of the material. The electric polarization of the material is influenced the magnetic structure and the magnetic structure, in turn, is influencing the electric polarization. This results in a spiral arrangement of the magnetic spins, what forms the spin cycloid [9]. The bismuth ferrite is highly desirable for various applications, particularly in the field of electronics, since the material allows manipulation by both electric and magnetic parts. Additionally, its magnetic properties can be used for magnetic data storage and magnetic field sensing.

IV. Piezoresponse Force Microscopy.

PFM stands for Piezoresponse Force Microscopy. The technique is used to study the electrical properties of ferroelectric and piezoelectric materials under the influence of oscillating electrical field. PFM is a type of scanning probe microscopy [9], there small cantilever with sharp tip on it is used to investigate the surface of the sample. In this technique the underlighting voltage is applied to the sample, and the resulting displacement or polarization response is analyzed [9]. The cantilever is utilized to measure the change in tip–sample interaction force, resulting from the samples response to applied electric field [9]. This interaction force can be related to electrical response of the sample be used to study a variety of properties including polarization, ferroelectric switching and the domain wall motion [10]. The method is a powerful tool for characterizing the properties of ferroelectric and piezoelectric materials at the nanoscale [10], and it has a number of advantages over other characterization techniques. The PFM is a non–destructive, non–contact method, which can be used to study both ferroelectric and piezoelectric properties of materials. In the context of the PFM, in–plane and out–of–plane refers to orientation of electric field resulting polarization response relative to surface of the sample [11]. In–plain refers to the measurement of the sample, when the electric field is applied in parallelly direction to the surface of the analized sample (bright region). Out–of–plane relates to the measurement of the polarization response of the sample, when the electric field is applied perpendicularly to the surface of the investigated sample (dark region) [11]. The combination of in–plain and out–of–plain allows recreation of a 2D piezoelectric map. The 2D piezoelectric map provides valuable information regarding the piezoelectric active regions across the sample, and also, the ability of switching their magnetic domain orientation, which is served to evaluate the overall magnetic properties of the investigated sample [12].

Direct PFM is often used to study the polarization behavior of ferroelectric and piezoelectric materials, including domain structure, domain wall motion/formation, polarizations switching and hysteresis [12]. The method is utilized to study the polarization properties of thin films, where the thickness of the material may not be sufficient to support bulk polarization behavior [12]. Direct PFM measure the polarization response of a ferroelectric or piezoelectric material in response to direct electric bias. The method is extremely important to study polarization behavior of thin films [12]. The thickness of deposited samples does not affect the magnetic properties. The factor which has significant impact on magnetic properties of bismuth ferrite, is the presence of oxygen vacancies. Oxygen vacancies create displacement of crystallographic structure of bismuth ferrite leading to ferroelectric properties of the material.

V. Results and Observations.

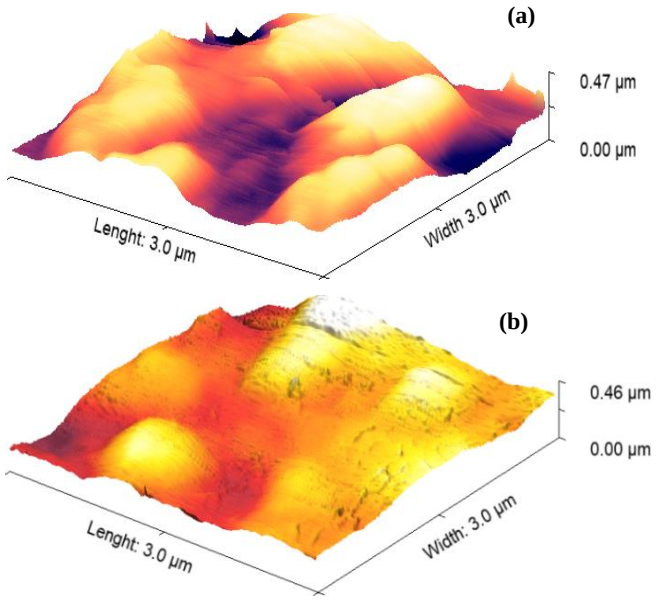


Fig 1. 3D surface of the analyzed 100 nm BFO (Pt/Ti/Si) sample.

The measurement using Piezoresponse Force Microscopy has been performed using NTEGRA Prim, which is owned by company NT–MDT. The surface has been analyzed with FMG01/Pt needle, which is specifically used for measurement of piesoresponce of bismuth ferrite thin films, thanks to platinum coating and monocrystal of Si. The analyzed area of surface is (3x3) μm^2, with peak height being around 472 nm (Fig 2).

The RMS roughness (*Sq*) of the entire sample is within 90 nm, where average roughness (*Sa*) of surface area is approximately 75 nm. Surface morphology of 100 nm BFO (Pt/Ti/Si) exhibits a small degree of inhomogeneity with several clusters of BFO material and indentations. A similar situation could be absorbed on the different analyzed surface area (Fig 3). However, the overall roughness of the surface is less, in comparison to the first analyzed area, with *Sa* is approximately 30 nm and *Sq* being around 60 nm. By comparison of the first and second analyzed areas we are capable to conclude, what average roughness of the surface stains is around 35 nm.

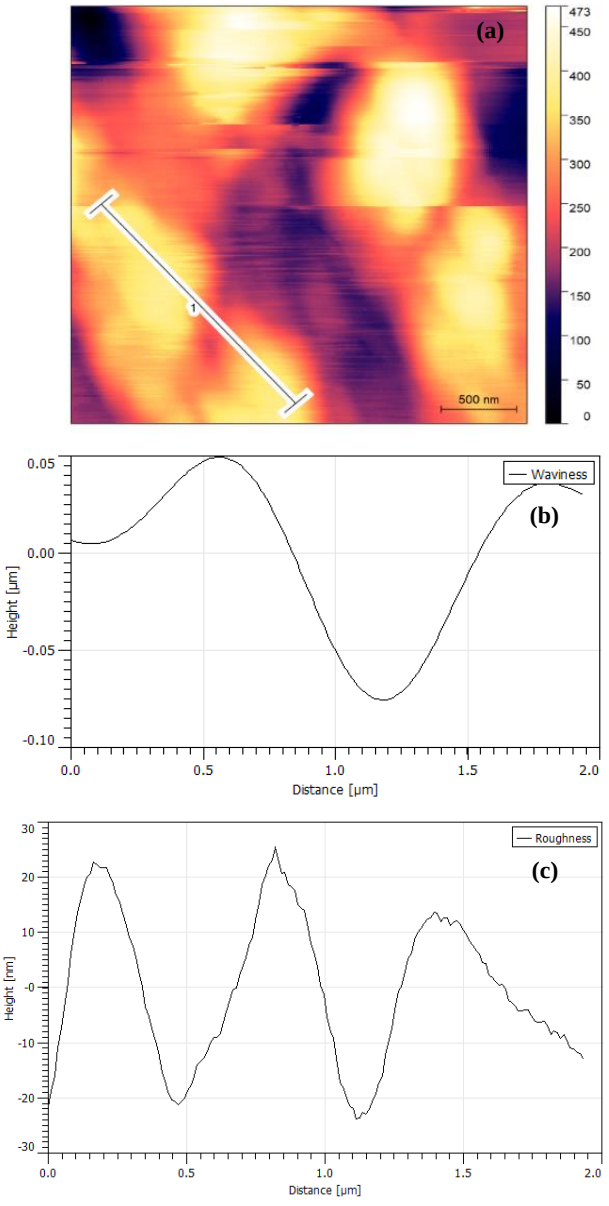


Fig. 2 AFM topology (a), waviness (b) and roughness (c)

In PFM, it is common to observe noise signals that can arise from various sources and factors. These noise signals can include thermal noises, electrical noises, mechanical noise and other factors. Thermal noise arises from the random motion of atoms and electrons within the material and can lead to fluctuation in the PFM signal. Electrical noise is caused by fluctuation in the applied voltage or current during the measurement process and can also be attributed to noise in the PFM signal. The mechanical noise, which is the most common cause of insufficient image quality, is attributed to vibrations or fluctuation in experimental setup or sample environment. Considering the area where the setup is located, the noise signal could be attributed to nearby vibration caused by many objects.

In summary, both analyzed areas show relatively smooth surface, which is achieved by PLD technique and reflect the quality of produced films.

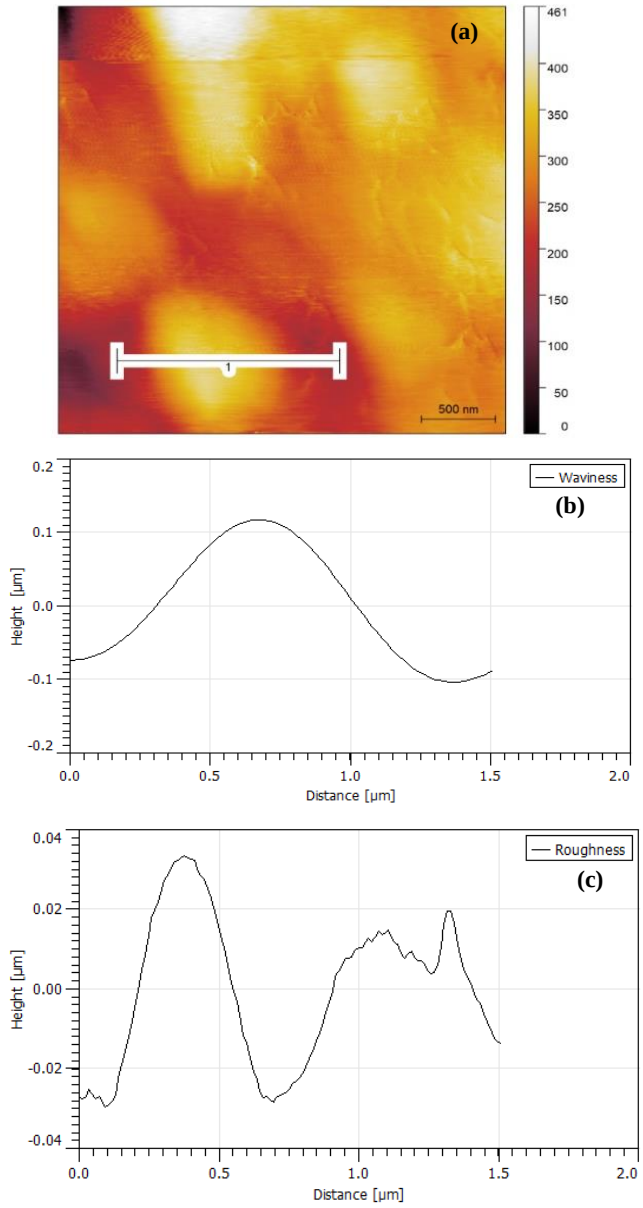


Fig. 3 AFM topology (a), waviness (b) and roughness (c) of different area

In Direct PFM analysis the voltage is applied directly to a bismuth ferrite sample and the resulting polarization response is measured using the cantilever. By analyzing the response across the surface of the sample, it is possible to identify the surface

regions are different piezoelectric properties and obtained valuable information about the piezoelectric behavior of the material. Direct PFM analysis reveals what sample, across the analyzed surfaces, shows piezo–active regions, which can be determined by bright and dark regions.

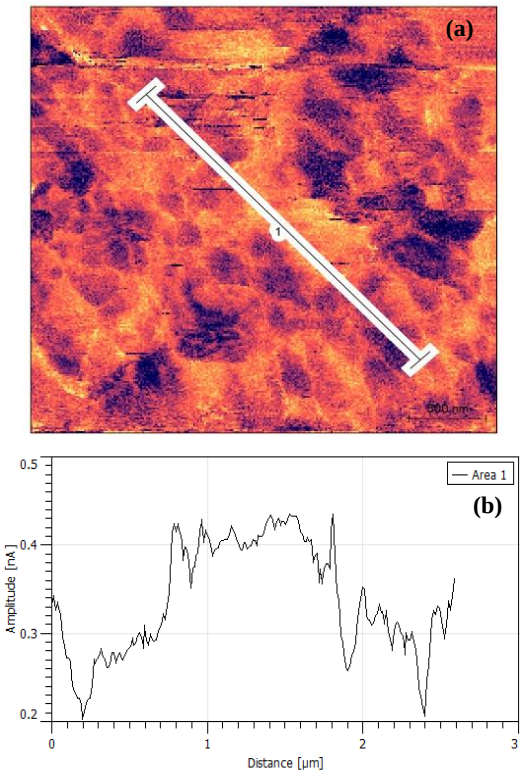


Fig 4 Literal PFM of investigated surface

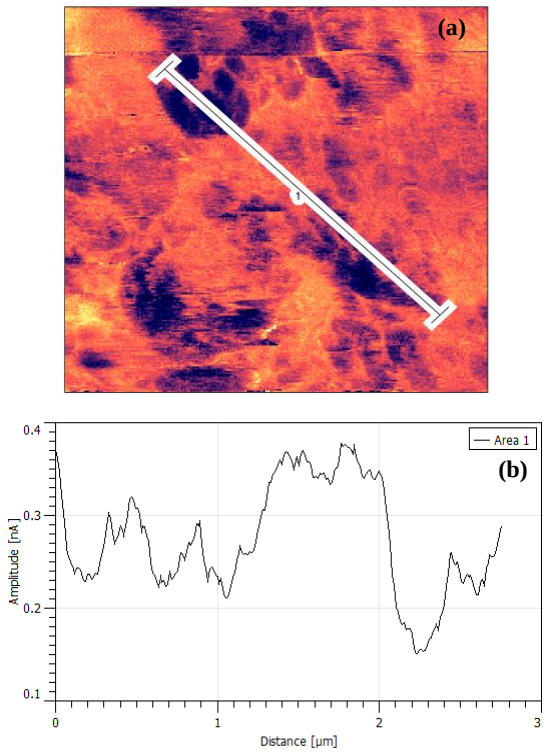


Fig 5 Literal PFM of different investigated surface

Fig. 4,5 indicate the mosaic layout of piezo–active regions with highly orientated domains in bright and dark reigns. Studies using direct PFM have indicate, what bismuth ferrite has a complex domain structure, with domains of different polarization direction. By comparison of topological and current distribution images, we can conclude, what grains of bismuth ferrite material are presumably more likely to create darker regions, which would be potentially described by various surface tension potential, thus different conductivity across the surface of the samples. In contrast, smother surface demonstrate formation of brighter regions. The application of electric field can result in the movement of domain walls, leading to a charge in overall polarization of the material.

ACKNOWLEDGMENT

This work was supported by the Internal Grant Agency of Brno University of Technology, grant No. FEKT–S–23–8228.

I would like to express my appreciation to doc. Mgr. Dinara Sobola Ph.D. for guidance and support during the research.

REFERENCES

[1] G. R. Fox, S. Trolier-McKinstry, S. B. Krupanidhi, and L. M. Casas, “Pt/ti/sio2/si substrates,” *J. Mater. Res.*, vol. 10, no. 6, pp. 1508–1515, 1995.

[2] Q. Yun, A. Bai, and S. Zhao, “Lattice distortion of holmium doped bismuth ferrite nanofilms,” *J. Rare Earths*, vol. 32, no. 9, pp. 884–889, 2014

[3] D. Misiurev, P. Kaspar, and V. Holcman, “Brief Theoretical Overview of Bi-Fe-O Based Thin Films,” *Materials (Basel).*, vol. 15, no. 24, 2022

[4] V. Bouquet *et al.*, “Influence of two-dimensional oxide nanosheets seed layers on the growth of (100)BiFeO3 thin films synthesized by chemical solution deposition,” *Thin Solid Films*, vol. 693, no. 100, 2020

[5] M. Pooladi, I. Sharifi, and M. Behzadipour, “A review of the structure, magnetic and electrical properties of bismuth ferrite (Bi2Fe4O9),” *Ceram. Int.*, vol. 46, no. 11, pp. 18453–18463, 2020

[6] C. C. Zhou, K. X. Jin, B. C. Luo, X. S. Cao, and C. Le Chen, “Photoconductivity in BiFeO3/La0.7Sr 0.3MnO3 heterostructure,” *Mater. Lett.*, vol. 64, no. 15, pp. 1713–1716, 2010

[7] S. U. Lee, S. S. Kim, M. H. Park, J. W. Kim, H. K. Jo, and W. J. Kim, “Effects of co-substitution on the electrical properties of BiFeO 3 thin films prepared by chemical solution deposition,” *Appl. Surf. Sci.*, vol. 254, no. 5, pp. 1493–1497, 2007

[8] H. Toupet, F. Le Marrec, J. Holc, M. Kosec, P. Vilarhino, and M. G. Karkut, “Growth and thermal stability of epitaxial BiFeO3 thin films,” *J. Magn. Magn. Mater.*, vol. 321, no. 11, pp. 1702–1705, 2009

[9] S. Sharma, A. Sharma, V. Gupta, N. K. Puri, and M. Tomar, “Comparison of Ferroelectric Photovoltaic Performance in

BFO/BTO Multilayer Thin Film Structure Fabricated Using CSD & PLD Techniques," *J. Electron. Mater.*, vol. 50, no. 4, pp. 1835–1844, 2021

[10] V. V. Lazenka *et al.*, "Effect of rare-earth ion doping on the multiferroic properties of BiFeO3 thin films grown epitaxially on SrTiO3 (1 0 0)," *J. Phys. D. Appl. Phys.*, vol. 46, no. 17, 2013

[11] D. Chen *et al.*, "Interface Engineering of Domain Structures in BiFeO3 Thin Films," *Nano Lett.*, vol. 17, no. 1, pp. 486–493, 2017

[12] Y. H. Chu, L. W. Martin, M. B. Holcomb, and R. Ramesh, "Controlling magnetism with multiferroics," *Mater. Today*, vol. 10, no. 10, pp. 16–23, 2007, 7021(07)

Modifications of single-tip cold-field emitters by nanolayer thin dielectric coatings of refractory oxides

1st Daniel Burda
Department of Physics
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czechia
Daniel.Burda@vut.cz

2nd Mohammad M. Allaham
Advanced Ceramic Materials
Central European Institute of Technology
Brno University of Technology
Brno, Czechia
allaham@vutbr.cz

3rd Alexandr Knápek
Electron and Plasma Technologies
Institute of Scientific Instruments of Czech Academy of Sciences
Brno, Czechia
knapek@isibrno.cz

***Abstract*— The aim of this research is to gain a more detailed understanding of the field emission characteristics of tungsten field emitters coated with a multilayers of refractory oxides such as Al2O3, HfO2, TiO2, which act as an additional potential barrier in metal-vacuum system. Introducing additional barriers to the emitter's metal vacuum interface system can alter its emission characteristics. Thin oxide barriers were prepared by a recipe involving use of low-temperature Atomic Layer Deposition. The parameters of grown layers were studied in Local Electrode Atom Probe and Field Ion Microscope. The emission parameters of the experimental emitters were studied in Field Emission Microscope.**

Keywords—cold field emission, nanolayer, tungsten, alumina, hafnium trioxide, electron tunneling

I. Introduction

It has been 85 years since Erwin Müller discovered field emission microscope in 1936 [1]. Since then, field electron emission theory has expanded greatly and many practical applications have emerged, such as electron guns, X-ray and microwave sources or field emission displays.

Field emitters can be divided into two main categories, single-tip emitters (STFEs) and large-area emitters (LAFEs), consisting of dozens of single-tip emitters in a matrix. Because of the tungsten durability and the simple and repeatable preparation of the tip, the polycrystalline tungsten wire-based STFE was chosen as a base emitter, on which a thin barrier coating is grown. In cold field electron (CFE) emissions, most electrons tunnel from electron states below the Fermi level. Tunneling is assumed through the Schottky-Nordheim (SN) barrier. The equation that forms the basis for field emission theory is called the Fowler–Nordheim equation. These equations can be utilized to extract emitter properties from the quantities measured in FEM. In this study we employ so-called *extended Murphy-Good equation* (EMG), details regarding the equation can be found in [2,3]

$$I(V) = \{A^{\mathrm{SN}}(\theta \exp \eta) V_{\mathrm{R}}^{-\kappa}\} V^{\kappa} \exp(-\eta V_{\mathrm{R}}/V) \qquad (1)$$

where A^{SN} is the formal emission area assuming SN barrier tunneling, V_{R} is a (constant) reference measured voltage [2] needed to pull the top of the SN barrier down to Fermi level [3], $\theta(\varphi)$ and $\eta(\varphi)$ are work function φ dependent scaling parameters, parameter $\kappa(\eta)$ is related: $\kappa = 2 - \eta/6$ [2]. By applying natural logarithm to (1):

$$\ln\{I/V^{\kappa}\} = \ln\{A^{\mathrm{SN}}(\theta \exp \eta) V_{\mathrm{R}}^{-\kappa}\} V^{\kappa} - \eta V_{\mathrm{R}}/V \qquad (2)$$

the equation becomes linear, $\ln\left\{\frac{I}{V^{\kappa}}\right\}\left(\frac{1}{V}\right)$, this form is called theoretical Murphy-Good plot.

Assuming the work function of the emitter interface is known, so-called orthodoxy test [3] based on (2) can be performed on measured experimental *I-V* characteristics. When an emitter is declared orthodox, the emitter related parameters: formal emission area A^{SN}, voltage conversion length ξ [2] can be extracted. The exact parameters and criteria of the orthodoxy test are further described elsewhere [4]. *Electron emission convention* is used.

II. Methodology

A. Electrochemical etching

The base tungsten emitter was prepared by the means of two-step electrochemical etching of polycrystalline tungsten wire (GoodFellow, 7440-33-7, 99,9+%, diameter of 0.3 mm). in 3M aqueous solution of NaOH. To ensure perfectly symmetrical emitter, tungsten wire connected as anode is submerged in electrolyte at the center of Pt ring connected as cathode, details on in-house developed setup can be found elsewhere [5]. Sharp tungsten tips with tip radius of 80 nm or less can be obtained.

To ensure the tip surface is clean before the next step, it is carefully submerged in 38% hydrofluoric acid (HF) for 20 minutes, the acid readily removes oxides, Na-based contaminants and organics. After the tip is removed from the acid, while transferring the tip into FEM or ALD chamber, oxygen from the air reacts with pristine W surface and up to 10 Å of tungsten trioxide is formed [6], the effects of the buried tungsten trioxide layer after covering it with 15 nm of Al_2O_3 were not apparent.

B. *Atomic layer deposition*

ALD is a HV deposition technique, the growth of layers is based on sequential, self-limiting surface reactions of two vapour phase precursors. ALD offers exceptional conformality, which allows to grow uniformly thin layer on very complex or sharp structures [6]. The nature of the layer formation allows for precise thickness control at the Angstrom level [6], as the overall thickness is controlled by number of cycles of sequential reactions of the two precursors. Also, tuning of film composition is achievable through the process temperature setting, such as increased crystallinity of the layers by increasing the process temperature [6].

Thin layer coatings of Al_2O_3 were prepared at CEITEC Nano facility in Ultratech-Cambridge Nanotech Fiji 200 ALD tool. The precursors used were trimethylaluminum (TMA) and water vapor, process temperature was set to 150 °C, the number of cycles was set to 150. Each one ALD cycle comprised of the steps: 1) injection of TMA, duration 0.06 s, 2) purge, duration 20 s, 3) injection of H_2O, duration 0.06 s, 4) purge, duration 20 s. The process resulted in 15 nm thin layer.

C. *Local electrode atom probe and field ion microscopy*

The local electrode atom probe Tomography (LEAP) field evaporates (with a voltage and laser pulse) a specimen in the form of a point. The ionized particles are collected in a position-sensitive, time-of-flight mass spectrometer. The technique can determine the atomic species with ppm sensitivity and position from within the original analysis volume with sub-nm accuracy [7].

D. *Field emission microscope measurement*

The measurement was carried out in the in-house developed FEM at ISI of the CAS. The emitter had been imagined in triode configuration, see Fig. 1. The extractor electrode with hole of 1 mm of diameter was precisely mounted just tenths of mm above the emitter tip to help draw off the electrons. Al-coated Ce:YAG scintillator was used as anode. The emission current was measured between the scintillator and the ground by Keithley 485 Pico ammeter.

Distance between the tip of cathode and the scintillator was set to 40 mm. During the active emission, the pressure inside the chamber remained lower than $8 \cdot 10^{-6}$ Pa.

After the etching in NaOH and HF cleaning, the emitter was loaded into FEM, *I-V* characteristics of clean W emitter were captured. Then, after the deposition of 15 nm thin layer of Al_2O_3, the very same emitter was loaded into FEM for the second time and the measurements were repeated.

TABLE I
VALUES OF FIRST EMISSION OF THE DIFFERENT EMITTERS

	V_{cat} (V)	V_{ext} (V)	I (nA)
Clean W	5030	662	0.14
Al_2O_3 coated	5030	977	5000

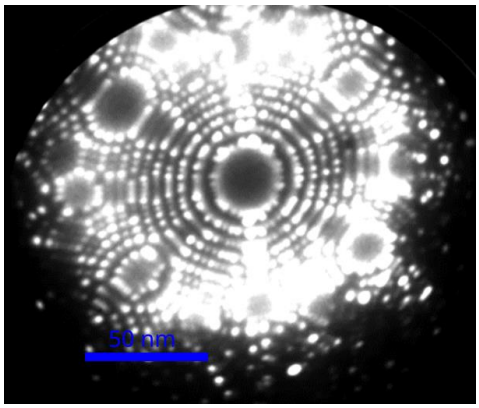


Fig. 1. Field ion microscope image of cleaned tungsten surface, of <110> plane, developed at applied voltage of 7000 V, Helium imaging gas, pressure 2.5e-4 mbar

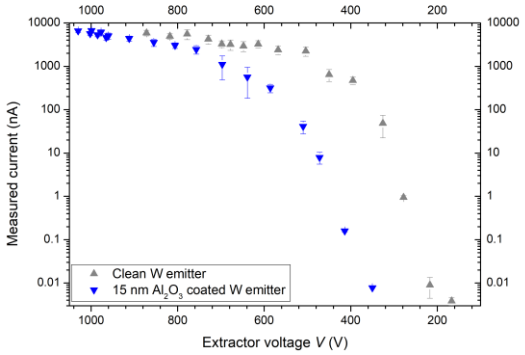


Fig. 2. *I-V* characteristics of clean and coated emitter, cathode voltage 5030 V.

III. RESULTS AND DISCUSSION

In order to protect the emitter from exploding during the measurement, the *I-V* characteristics shown in the Fig. 2 were captured in this manner: After the first ignition (Table 1), the extraction voltage had been increased, so that the ammeter would collect the current of no more than 10 μA.

Then the actual *I-V* measurement started, recording the decrease of current while decreasing the extraction voltage at the constant cathode voltage.

Based on the current-voltage values in Table 1, the coated emitter shows very strong *switch-on effect*, i.e. first emission does not begin at the lowest measurable current, but at a value often several orders of magnitude higher [8].

For the purpose of the *orthodoxy test*, work function of clean tungsten emitter is assumed to be 4.5 eV. Work function of Al_2O_3 is assumed to be 4.0 eV, as based on calculations from ultra-violet photoelectron spectroscopy of several nanometers thin Al_2O_3 layers [9]. Murphy-Good plots based on collected data are shown in the Fig. 3. Both of the emitters showed nontrivial behavior, passing the orthodoxy test when extraction voltage was set low and failing the test when extraction voltage was set high. This is fairly common in pristine tungsten emitters, at high extraction voltage and thus high measured emission current, slight macroscopic changes in electrostatic field at the emitter can occur and then the emitter can no longer

be properly described by the current theory. Extracted values from linear parts of Murphy-Good plots are shown in the Table 2.

The projection images of the Al_2O_3 coated emitter and their changes in respect to the change of extraction voltage are illustrated in the Fig. 4. The images resembled of the multi-spot emission patterns, fluctuations in the current were also visible as occasional bright spots. Unfortunately, the images were slightly off-center and the portion of electron beam touched the conductive scintillator holder ring that resulted in parts of the images on the left side not visible, which renders more advanced determination of emission active planes inconclusive.

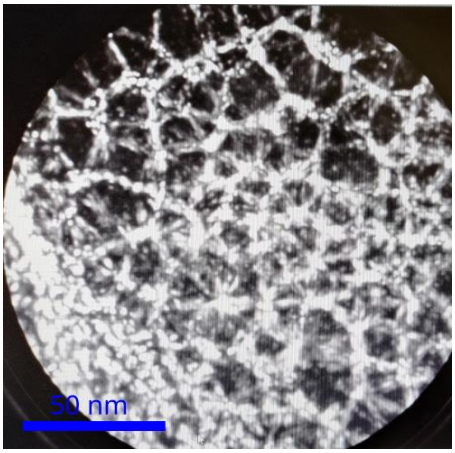


Fig. 3. Al_2O_3 oxide layer structure observed in Field ion microscope,camera settings: exp 0.5s, ISO 800, gain 0, f1.8 lens

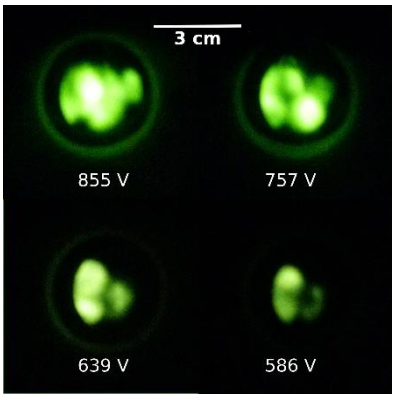


Fig. 4. Evolution of emission patterns in projection images of Al_2O_3 coated emitter. Four values of extractor voltage V_{ext}, constant cathode voltage V_{cat} = 5030 V, exposure: f/2, 1/250 s, ISO 1600.

TABLE II
EXTRACTED VALUES OF EMITTER PARAMETERS FROM LINEAR PARTS OF MURPHY-GOOD PLOTS

	Result	A^{SN} (nm^2)	ζ (nm)
Clean W	Passed	63.7	75.9
Al_2O_3	Passed	369.3	158.4

IV. CONCLUSIONS

A Presented research shows the possibilities of fabrication of multilayer coating of very thin Al_2O_3 / HfO_3 oxide stacked field emitter using ALD. The deposition technique enables for fine tuning of the thickness of the coating. The multilayer coatings, especially the structure of amorphous Al_2O_3 layers, were observed in FIM. Similar to other coated emitters, the emitter exhibits strong *switch-on effect*. The orthodox behavior can be observed when the Al_2O_3 coated emitter is operated in low current, i.e. extraction voltage is low, the emitter can then be described by cold field emission theory.

ACKNOWLEDGMENT

CzechNanoLab project LM2018110 funded by MEYS CR is acknowledged for the financial support of the measurements/sample fabrication at CEITEC Nano Research Infrastructure. The research was supported by the Czech Academy of Sciences (RVO:68081731). The research described in this paper was supported by the Internal Grant Agency of Brno University of Technology, grant No. FEKT-S-20-6352 and by grant No. CEITEC VUT/FEKT-J-23-8307.

REFERENCES

[1] M. Drechsler, “Erwin Müller and the early development of field emission microscopy”, *Surface Science*, vol. 70, no. 1, pp. 1-18, 1978.

[2] R. G. Forbes, “The Murphy–Good plot: a better method of analysing field emission data”, *Royal Society Open Science*, vol. 6, no. 12, Dec. 2019.

[3] M. M. Allaham, R. G. Forbes, A. Knápek, and M. S. Mousa, “Implementation of the orthodoxy test as a validity check on experimental field emission data”, *Journal of Electrical Engineering*, vol. 71, no. 1, pp. 37-42, Feb. 2020.

[4] M. M. Allaham, R. G. Forbes and M. S. Mousa, “Applying the Field Emission Orthodoxy Test to Murphy-Good Plots”, *Jordan Journal of Physics*, vol. 13, no. 2, pp. 101-111, Aug. 2020.

[5] A. Knápek, J. Sýkora, J. Chlumská, and D. Sobola, “Programmable set-up for electrochemical preparation of STM tips and ultra-sharp field emission cathodes”, *Microelectronic Engineering*, vol. 173, pp. 42-47, 2017.

[6] A. Warren, A. Nylund, and I. Olefjord, “Oxidation of tungsten and tungsten carbide in dry and humid atmospheres”, *International Journal of Refractory Metals and Hard Materials*, vol. 14, no. 5-6, pp. 345-353, 1996.

[7] T. F. Kelly and D. J. Larson, “Local Electrode Atom Probes”, *Materials Characterization*, vol. 44, no. 1-2, pp. 59-85, 2000.

[8] C. S. Athwal and R. V. Latham, “Switching and other nonlinear phenomena associated with prebreakdown electron emission currents”, *Journal of Physics D: Applied Physics*, vol. 17, no. 5, pp. 1029-1043, May 1984.

[9] W. Song and M. Yoshitake, “A work function study of ultra-thin alumina formation on NiAl(110) surface”, *Applied Surface Science*, vol. 251, no. 1-4, pp. 14-18, 2005



YUNEX
TRAFFIC

Become a Traffic Transformer





Měníme dopravní průmysl

Kopeme za tým Yunex Traffic, globálního lídra v oblasti inteligentních dopravních systémů. V Brně dlouhodobě pracujeme na revolučních technologiích pro mobilitu budoucnosti. Vyvíjíme řešení pro řízení dopravy, vedoucí k vyšší plynulosti, bezpečnosti, efektivitě a šetrnosti k životnímu prostředí. Výsledky naší práce vnímají obyvatelé ve městech po celém světě.

- 3300 kolegů celosvětově
- 80 kolegů v Brně a Praze
- 110 měst, kterým pomáháme



Hledáme:

- SW Developery, Testery, UX-designéry
- Studenty, kteří u nás budou řešit bakalářské nebo magisterské práce
- Studenty na PhD. výzkum

Nabízíme:

- Propojení SW vývoje, IoT a dopravní problematiky
- Zázemí globální společnosti s lidským přístupem
- Moderní technologie a jejich aplikaci v každodenním životě našich měst
- Práci v mezinárodním prostředí – spolupracujeme s týmy v Německu
- Hybridní spolupráci a práci z domova
- Účast na školeních, konferencích a komunitních setkáních
- Až 33 dní dovolené
- Pestrý program benefitů
- Moderní kanceláře kousek od centra Brna s dobrým parkováním

A jako bonus:
„Když to zrovna situace dovolí, rádi se potkáváme u piva nebo dalších aktivit (fotbálek, hry, simracing) a také rádi navštěvujeme naše kolegy v Německu."

Kontaktujte nás:
Gabriela Golcová
in @gabriela-golcová
+420 777 193 568
gabriela.golcova.ext@yunextraffic.com

Přidej se k nám!

www.yunextraffic.cz

Design of all-inorganic lead-free cesium bismuth iodide perovskite

1st Tatiana Pisarenko
Department of Physics
Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, the Czech Republic
ORCID: 0000-0002-3527-7099

2nd Stevan Gavranović
Institute of Physical and Applied Chemistry
Faculty of Chemistry
Brno University of Technology
Brno, the Czech Republic
ORCID: 0000-0002-0176-8703

***Abstract*—The aim of this work is the design, synthesis, and subsequent analysis of a novel environmentally friendly $Cs_3Bi_2I_9$ perovskite. $Cs_3Bi_2I_9$ is a promising material for optoelectronic applications such as solar cells and photodetectors due to its excellent light absorption properties in the visible and near-infrared regions. The paper, after successful synthesis, is mainly devoted to the non-destructive characterization of the perovskite material using energy-dispersive electron microscopy, Raman spectroscopy, and X-ray photoelectron microscopy. The data obtained confirm the successful design of the perovskite structure. These results thus serve as a crucial step in the subsequent advanced development of a lead-free sensor.**

***Index Terms*—perovskite, synthesis, characterization, analysis, $Cs_3Bi_2I_9$, lead-free**

I. Introduction

Perovskite is a term used to refer to a specific crystal structure that is widely found in different materials, including minerals, ceramics, and semiconductors. Perovskites are a family of materials with trivalent compounds with the general formula ABX_3, where A and B represent cations of different sizes—A is usually an alkali or alkaline earth metal, B is typically a transition metal. X represents an anion that is negatively charged. These materials have many interesting properties such as piezoelectric effect, high thermal conductivity, high electron mobility, exceptional optical properties, and can act as a tuneable electronic structures. Various materials with perovskite structures have been discovered, and the structure has proven to be an essential building block for many technological applications, including solar cells, LEDs, catalysis, and magnetic materials [1].

As an ideal structure, A atoms are located at the corners of the unit cell, B atoms occupy the center of the unit cell, and X atoms are at the center of each face of the unit cell. The resulting crystal structure is cubic, with the B atom surrounded by X atoms arranged in an octahedral fashion. This arrangement is critical for stability and allows many different materials to adopt perovskite structures [2].

Research described in the paper was financially supported by the Ministry of Education, Youth and Sports of the Czech Republic under the project CEITEC 2020 (LQ1601), by the Internal Grant Agency of Brno University of Technology, grant No. FEKT-S-23-8228 and grant No. FCH/FEKT-J-23-8288. Part of the work was carried out with the support of CEITEC Nano Research Infrastructure supported by MEYS CR (LM2018110).

Lead-free perovskite is a new material that is more stable than traditional perovskite, which contains lead [3], [4]. It also has higher efficiency in converting light energy into electricity and can be used in a wide range of applications. In addition, lead-free perovskite is easier to fabricate and has higher cost-effectiveness than traditional materials. $Cs_3Bi_2I_9$ perovskites are a type of metal-halide perovskite material that has recently gained attention from scientists and engineers due to a number of advantages compared to traditional photovoltaic materials. The material is also used in optoelectronic devices, such as lasers and photodetectors, and in nanowire-based devices. It can also be used in LEDs and other display technologies. The material is more cost-effective and efficient than traditional semiconductor materials, and it can be prepared in solution form, allowing for more straightforward device fabrication. The material also exhibits good thermal stability and is optically transparent, making it suitable for use in thin-film solar cells [5].

However, their commercial viability is limited by their susceptibilities to environmental degradation, such as humidity and light. The primary mechanisms of degradation of metal-halide perovskites are due to water absorption, charged carrier recombination, and photo-oxidation. Water absorption leads to the formation of hydroxide ions, which can create structural defects in the material. Charged carrier recombination occurs due to the presence of ionized species in the material, which can lead to the formation of free radicals [6]. Lastly, photo-oxidation occurs when the material is exposed to light, which can create oxidative species in the material [7].

II. Materials & Methods

The section describes two primary parts—fabrication and subsequent instrumentation to examine the resulting crystal and evaluate the success of the synthesis.

A. Design and synthesis

$Cs_3Bi_2I_9$ single crystals were prepared using two solution methods: hydrothermal (HT) and inverse temperature crystallization (ITC). In the first method, cesium iodide (99.9 % CsI) and bismuth (III) iodide (99.9 % BiI_3) were dissolved (molar ratio 3:2) in hydroiodic acid (57 % HI, EMSURE®) and heated

up to $200\,^{\circ}\mathrm{C}$ in autoclave reactor. This solution was then cooled down to prepare the crystal seeds (slow cooling $1\,^{\circ}\mathrm{C/h}$) and perovskite powder (fast cooling overnight). In the second method, the single crystal seeds were placed into the solution of $Cs_3Bi_2I_9$ in N,N-dimethylformamide (anhydrous, $99.8\,\%$ DMF) and slowly heated up using an oil bath and hotplate (Thermo Scientific™ SuperNuova+™ with PT1000 temperature probe) to increase its volume. All chemicals were purchased in Sigma Aldrich and used without further purification.

B. Material investigation

Several spectroscopic non-destructive methods were used for material analysis and characterization to evaluate the successful synthesis and quality of the obtained perovskite sample.

1) Scanning electron microscopy (SEM) and energy dispersive spectroscopy (EDS): The primary objective of using SEM was to accurately examine the sample to provide a detailed understanding of the overall surface structure. For this purpose, a Lyra3 microscope was used (Tescan, Brno, CZ). The optical microscope provided a quick overview of the state of the sample. However, the SEM was used for much more detailed investigation in the order of microns and hundreds of nanometers. This was mainly for the analysis of the formation of crystal structures and the evaluation of the successful growth of the crystal, and also for the possible detection of degradation processes.

No preparation or coating of the conductive layer was required prior to observation due to the sample being partially composed of alkali and heavy metals. Therefore, an accelerating voltage of $10\,\mathrm{kV}$ was used, as it was not expected that there would be charge accumulation on the surface.

In addition to the conventional observations, the X-Max 50 EDS detector (Oxford Instruments, Abingdon, UK) was used. For perovskites, EDS is an essential tool for studying their chemical composition using X-ray spectroscopy. During interaction with the material, electrons emit characteristic X-rays which are recorded by the detector. Based on the intensity of these signals, it can be determined which elements are present in the sample. This is important because perovskites are often doped with different elements to achieve desired properties, such as improved thermal stability or optical conductivity. This technique also allows a high-resolution image as a map of the material to be obtained, which is very useful for visualizing microstructures. The interaction volume for EDS and the X-rays that are emitted from the sample is higher than for detecting topographical information using secondary electrons from a conventional SEM. Therefore, it was appropriate to increase the accelerating voltage to $15\,\mathrm{kV}$. However, the spectral region of interest was cut down to the most important and characteristic peaks.

2) Raman spectroscopy: Raman spectroscopy of perovskites can be very useful for the study of new materials. This non-invasive method works by applying a laser beam to the surface, which interacts with the molecules in the sample. These molecules then vibrate, producing a characteristic Raman spectrum. The structural fingerprint of the sample and information about intermolecular interactions and vibrations can thus be reliably determined within a few micron units of the material depth.

The confocal Raman imaging system alpha300 R (WITec, Ulm, DE) was selected. A red laser with a wavelength of $633\,\mathrm{nm}$ was used as the light source. The laser power was set to $1\,\mathrm{mW}$ before the objective. Relatively low power was chosen concerning material stability. The spectral noise can be controlled quite reliably by the laser-sample interaction time and the spectra accumulation. In this case, the integration time was $7\,\mathrm{s}$ and 6 accumulations were performed. Objective magnification was set to $100\times$, which reduces spot-size quite a bit. The main spectral region of interest for $Cs_3Bi_2I_9$ lies in the $40\,\mathrm{cm^{-1}}$ to $200\,\mathrm{cm^{-1}}$ band.

3) X-ray photoelectron spectroscopy (XPS): This is the most accurate method used in this work. It uses high-energy X-rays to induce photoelectrons from surface atoms [8]. XPS is a frequently used technique for the elemental analysis of perovskite materials. It allows the non-invasive study of the bonding character of elements to assess their ability to react with other materials, a study of the oxidation state of cations, functional groups, and many other parameters of perovskites. This technique is also helpful for assessing the quality of perovskite layers in terms of homogeneity and changes in chemistry due to exposure to the atmosphere. The instrument used was the AXIS SupraTM X-ray photoelectron spectrometer (Kratos Analytical Ltd., Manchester, UK). A wide spectrum in the range $1200\,\mathrm{eV}$ to $0\,\mathrm{eV}$, high-resolution spectra of the C1s region (where the entire spectrum was calibrated according to the $\mathrm{C-C}$ chemical state at $\sim 284.8\,\mathrm{eV}$), Cs4p region, Bi4d region, and I3d region were recorded. Prior to the measurements itself, etching using Ar^+ ion sputtering at $1500\,\mathrm{eV}$ energy for $30\,\mathrm{s}$ was applied.

III. Results

The image acquired of successfully grown large size and well-shaped perovskite single crystal sample is shown in Fig. 1. In agreement to AMX_3 perovskite, $Cs_3Bi_2I_9$ follows the MX_6 octahedral generic structure, with $[BiI_6]^{3-}$ octahedral pairs sharing a common face to form $[Bi_2I_9]^{3-}$ dioctahedral clusters. A gap between dioctahedrons is filled with Cs^+ cations to form an isolated zero-dimensional (0D) of $Cs_3Bi_2I_9$ molecular salt configuration [9].

The Fig. 2 clearly shows the hexagonal prismatic structure from the SEM where several areas are covered. The first part a) defines the surface preview, a larger region of the crystal where significant growth can be observed. The image in part b) identifies the cross-section and subdivision of each layer. On closer inspection, the fine detail of the crystal is recognized in part c). The porous surface resembling fractures can be observed on deeper examination. Their detail is shown in the high-resolution image in part d). These cracks may have been caused by degradation and subsequent oxidation processes [10].

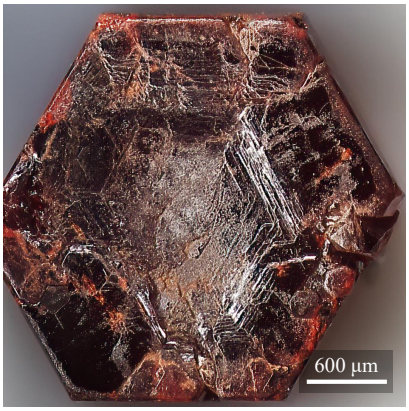


Fig. 1. Preview of the successfully grown perovskite crystal from the optical microscope.

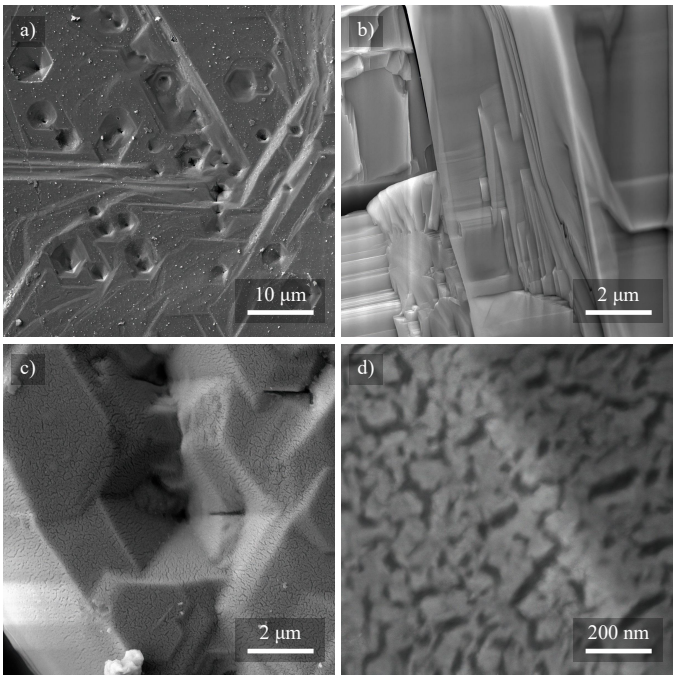


Fig. 2. Four images of one sample using the SEM. The broad surface of the sample examined from the top is shown in the picture a). Its record from the cross-section is taken in the picture b). A detailed view of the octagonal prism can be seen in the picture c). The porous structure in very high magnification is shown in the picture d).

Despite having an analysis proved structural composition, $Cs_3Bi_2I_9$ has highly unusual ferroelastic phase transitions and optical characteristics that are hard to fully comprehend given its molecular nature [11].

The curves obtained by Raman spectroscopy are in good agreement with an earlier study by Vakulenko et al. [12]. Results in Fig. 3 show that the largest region of interest is concentrated up to $200\,\mathrm{cm}^{-1}$. No other peaks were found at higher values.

From the analysis of the series vibrational modes, the two most prominent peaks occur at $104.8\,\mathrm{cm}^{-1}$ and $146.5\,\mathrm{cm}^{-1}$ in the form of a symmetric stretch of the bridge Bi-I bonds and

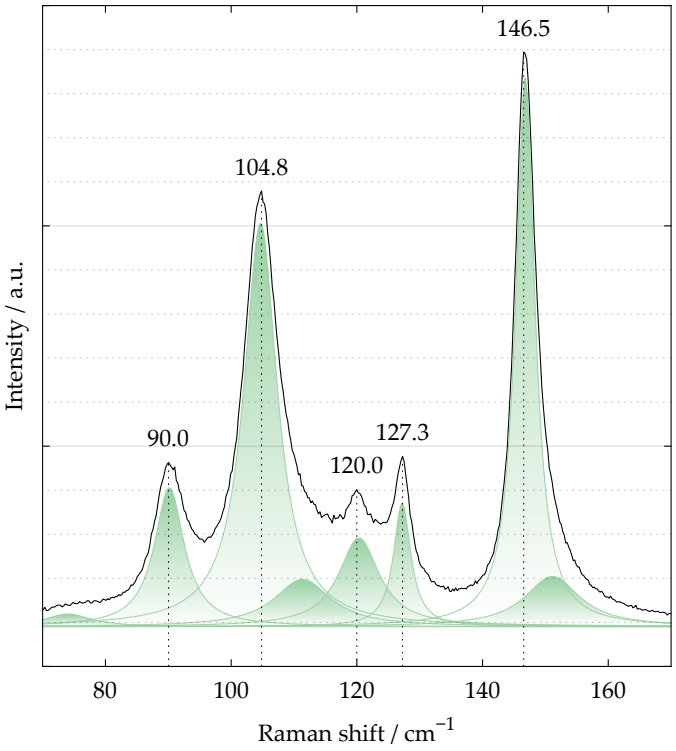


Fig. 3. Raman spectra of $Cs_3Bi_2I_9$ perovskite. The corresponding fits to the spectrum are also carefully modeled.

terminal Bi-I bonds respectively [13]. Supporting overlapping peaks are found at $90.0\,\mathrm{cm}^{-1}$, $120.0\,\mathrm{cm}^{-1}$ and $127.3\,\mathrm{cm}^{-1}$. The majority of occurring vibrational modes are the Bi-I stretching [14] as summarized in Tab. I.

TABLE I
THE ASSOCIATED PEAKS OF THE RAMAN SPECTRUM.

Raman peak / cm^{-1}	Description	Vibration	Symmetry
90.0	bridge Bi–I, asymmetric stretch	ν_7	E″
104.8	bridge Bi–I, symmetric stretch	ν_2	A_1^1
120.0	terminal Bi–I, asymmetric stretch	ν_{10}	E′
127.3	terminal Bi–I, asymmetric stretch	ν_{15}	E″
146.5	terminal Bi–I, symmetric stretch	ν_1	A'_1

The most prominent bands of the EDS spectra were within $5\,\mathrm{keV}$ as shown in Fig. 4. The material composition shows the presence of all the major elements $Cs_3Bi_2I_9$. Since the scan was performed as a map, the spectrum is accumulated from a broad area of the material, yet there were increases in oxygen in different parts of the crystal. This is quite significant information. Although in this measurement, there was an increase of $7.05\,\mathrm{wt}\%$, in others, it was up to $20\,\mathrm{wt}\%$. Oxygen may play a primary role in the aging of this structure. The ratio between Cs, Bi, and I mainly remained at similar values, with the combination of cesium and bismuth predominating at close to equal rates and iodide at lower rates.

The elemental material composition from XPS provided several interesting findings and proved the $Cs_3Bi_2I_9$ perovskite formation. In the wide-spectrum in Fig. 5, the presence of

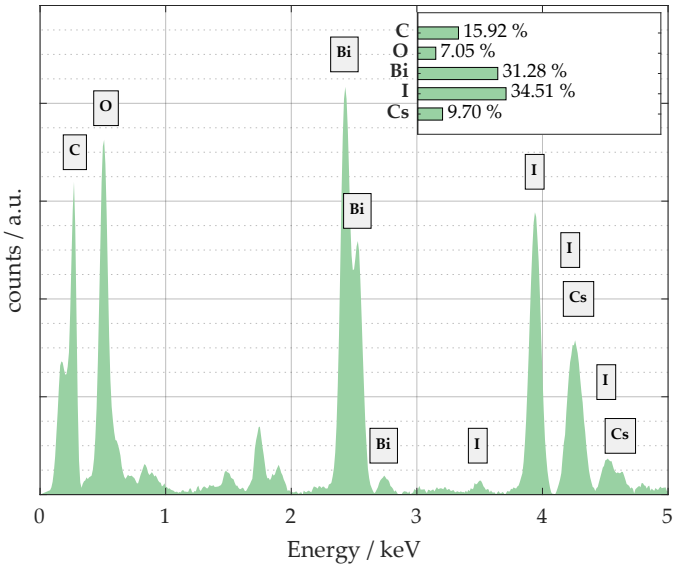


Fig. 4. Survey spectra from EDS where the all most significant peaks confirm the ideal composition of $Cs_3Bi_2I_9$ sample.

several characteristic bands is quite generous. Carbon peak C1s, cesium peak Cs3d, bismuth peak Bi4f, and iodide peak I3d were chosen for our purposes. These spectra can be found in Fig. 6. Fitting was performed at the regions where the overlapping peaks appeared.

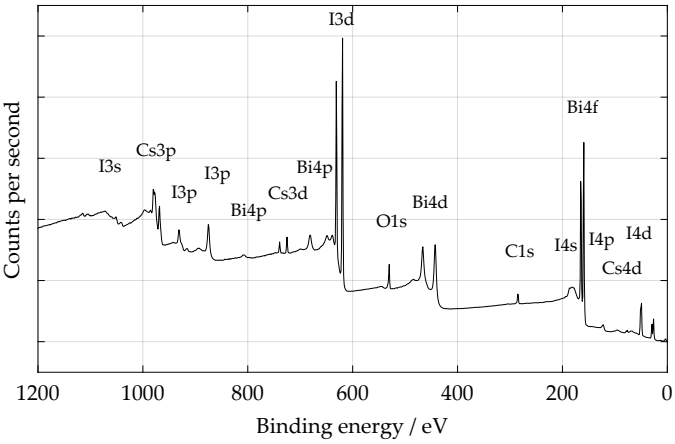


Fig. 5. Wide XPS spectra of the fabricated sample serving as a measurement summary. Several characteristic peaks for cesium, bismuth, and iodide can be identified.

The core-level binding energies of measured cesium Cs3d were recognized at 738.5 eV and 724.6 eV with a spin-orbit splitting of 13.9 eV relevant to the Cs^+ state. The lower energy at 738.5 eV is the cause of higher valence anionic units $(Bi_2I_9)^{3-}$, which are not directly bonded to one another in 0D−$Cs_3Bi_2I_9$ [15]. The bismuth Bi4f spectra were characterized at the binding energies at 164.4 eV and 159.0 eV with measured energy separation of 5.4 eV, which is typical for Bi^{3+} species [16]. As evident from the fit, these two main peaks have started to deconvolute into doublets in which one is composed at 162.5 eV and the second at 157.2 eV, and can be related to Bi^0 state. The second doublet peaks are separated by 5.4 eV. Iodide band I3d binding energies at 619.5 eV and 631.0 eV can be attributed to $3d_{5/2}$-I and $3d_{3/2}$-I, respectively, indicating the presence of a monovalent iodine atom with a comparable coordination and an energy separation of 11.5 eV [17]. A typical carbon spectrum of C1s defines three chemical states at 284.8 eV for the C − C component, 286.2 eV for the C − O − C component and 289.1 eV for the O − C = O component. The latter shows a very low abundance in spectra, and due to the inaccuracy of the fit, the actual value may vary within tenths of eV. The C1s peak, by default, represents adventitious carbon contamination in the range of units of nm on the surface, which can, however, be minimized by argon sputtering. During argon sputtering, however, reduction of Bi^{3+} and the formation of metallic bismuth can occur [15].

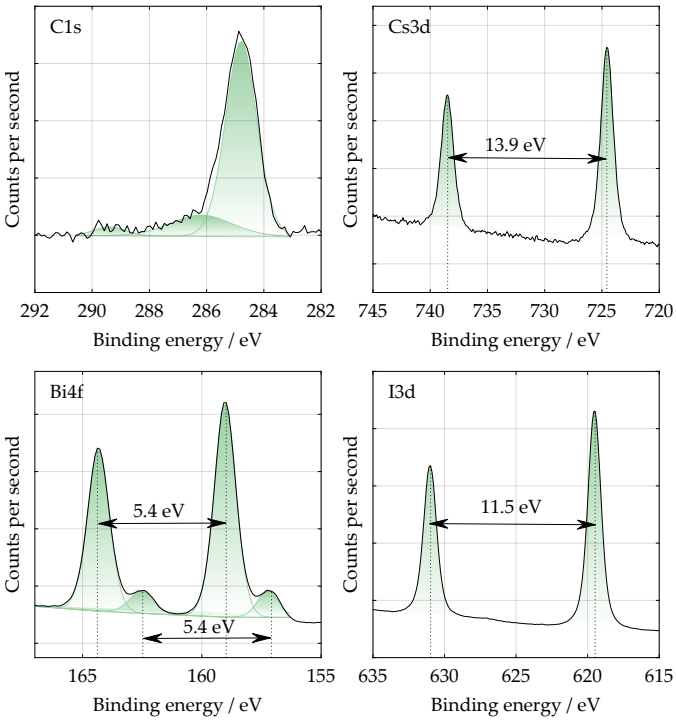


Fig. 6. High-resolution XPS spectra of C1s, Cs3d, Bi4f, and I3d.

IV. Conclusion

In conclusion, the perovskite structure is a vital building block for many materials with various technological applications. Its unique properties have made it a subject of intense research in various fields. With ongoing research, perovskite-based materials have the potential to bring about transformative changes in the technology industry. While more research is needed to improve the material's performance further, $Cs_3Bi_2I_9$ perovskites have a bright future in the optoelectronics industry.

After successful sample fabrication, all measurements and analyses performed independently confirmed the formation of $Cs_3Bi_2I_9$ crystal and its chemical fingerprint. EDS showed that

the ratio of Cs and Bi was very similar. Raman spectroscopy revealed an almost identical occurrence of peaks characteristic for this lead-free perovskite and a predominance of the Bi-I stretching vibrational modes. A similar situation was reported for XPS, where typical ranges of energy separations were found for specific high-resolution regions. SEM showed perfect crystal growth. However, a porous structure was discovered at high magnification, which can lead to imperfections in the crystal and a higher risk of an oxidation process. This is yet another motivation for the further development of this material.

Acknowledgment

Authors would like to thank doc. Mgr. Dinara Sobola, Ph.D. and Ing. Nikola Papež, Ph.D., for their valuable comments and guidance through this work.

References

[1] S. Gavranović, J. Pospíšil, O. Zmeškal, V. Novák, P. Vanýsek, K. Částková, J. Cihlář, and M. Weiter, "Electrode Spacing as a Determinant of the Output Performance of Planar-Type Photodetectors Based on Methylammonium Lead Bromide Perovskite Single Crystals," *ACS Applied Materials and Interfaces*, vol. 14, no. 17, pp. 20 159–20 167, 5 2022. [Online]. Available: https://pubs.acs.org/doi/abs/10.1021/acsami.1c24362

[2] F. Bai, Y. Hu, Y. Hu, T. Qiu, X. Miao, and S. Zhang, "Lead-free, air-stable ultrathin $Cs_3Bi_2I_9$ perovskite nanosheets for solar cells," *Solar Energy Materials and Solar Cells*, vol. 184, pp. 15–21, 9 2018.

[3] Y. Liu, Z. Xu, Z. Yang, Y. Zhang, J. Cui, Y. He, H. Ye, K. Zhao, H. Sun, R. Lu, M. Liu, M. G. Kanatzidis, and S. F. Liu, "Inch-Size 0D-Structured Lead-Free Perovskite Single Crystals for Highly Sensitive Stable X-Ray Imaging," *Matter*, vol. 3, no. 1, pp. 180–196, 7 2020.

[4] R. Wang, J. Wang, S. Tan, Y. Duan, Z. K. Wang, and Y. Yang, "Opportunities and Challenges of Lead-Free Perovskite Optoelectronic Devices," *Trends in Chemistry*, vol. 1, no. 4, pp. 368–379, 7 2019.

[5] Y. Hu, S. Zhang, X. Miao, L. Su, F. Bai, T. Qiu, J. Liu, and G. Yuan, "Ultrathin $Cs_3Bi_2I_9$ Nanosheets as an Electronic Memory Material for Flexible Memristors," *Advanced Materials Interfaces*, vol. 4, no. 14, p. 1700131, 7 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/admi.201700131

[6] D. W. Dequilettes, K. Frohna, D. Emin, T. Kirchartz, V. Bulovic, D. S. Ginger, and S. D. Stranks, "Charge-Carrier Recombination in Halide Perovskites," *Chemical Reviews*, vol. 119, no. 20, pp. 11 007–11 019, 1 2019. [Online]. Available: https://pubs.acs.org/doi/abs/10.1021/acs.chemrev.9b00169

[7] Z. Cui, Y. Wu, S. Zhang, H. Fu, G. Chen, Z. Lou, X. Liu, Q. Zhang, Z. Wang, Z. Zheng, H. Cheng, Y. Liu, Y. Dai, B. Huang, and P. Wang, "Insight into a strategy to improve charge carrier migration in lead-free bismuth-based halide perovskite for efficient selective oxidation of thioanisole under visible light," *Chemical Engineering Journal*, vol. 451, p. 138927, 1 2023.

[8] T. Pisarenko, N. Papež, D. Sobola, Ş. Ţălu, K. Částková, P. Škarvada, R. Macků, E. Ščasnovič, and J. Kaštyl, "Comprehensive Characterization of PVDF Nanofibers at Macro- and Nanolevel," *Polymers*, vol. 14, no. 3, 2 2022. [Online]. Available: https://doi.org/10.3390/polym14030593

[9] Y. Zhang, Y. Liu, Z. Xu, H. Ye, Z. Yang, J. You, M. Liu, Y. He, M. G. Kanatzidis, and S. F. Liu, "Nucleation-controlled growth of superior lead-free perovskite $Cs_3Bi_2I_9$ single-crystals for high-performance X-ray detection," *Nature Communications*, vol. 11, no. 1, pp. 1–11, 5 2020. [Online]. Available: https://www.nature.com/articles/s41467-020-16034-w

[10] M. Li, S. Xu, L. Wu, H. Tang, B. Zhou, J. Xu, Q. Yang, T. Zhou, Y. Qiu, G. Chen, G. I. Waterhouse, and K. Yan, "Perovskite $Cs_3Bi_2I_9$ hexagonal prisms with ordered geometry for enhanced photocatalytic hydrogen evolution," *ACS Energy Letters*, vol. 7, no. 10, pp. 3370–3377, 2022. [Online]. Available: https://doi.org/10.1021/acsenergylett.2c01856

[11] K. M. McCall, C. C. Stoumpos, S. S. Kostina, M. G. Kanatzidis, and B. W. Wessels, "Strong Electron-Phonon Coupling and Self-Trapped Excitons in the Defect Halide Perovskites $A_3M_2I_9$ (A = Cs, Rb; M = Bi, Sb)," *Chemistry of Materials*, vol. 29, no. 9, pp. 4129–4145, 5 2017. [Online]. Available: https://pubs.acs.org/doi/full/10.1021/acs.chemmater.7b01184

[12] O. V. Vakulenko, V. O. Gubanov, S. V. Kun, F. V. Motsnyi, E. Y. Peresh, and V. A. Terekhov, "Raman diagnostics of new types of $A_3B_2X_9$ layered crystals," in *Optical Diagnostics of Materials and Devices for Opto-, Micro-, and Quantum Electronics*, vol. 3359. SPIE, 4 1998, pp. 351–354. [Online]. Available: https://doi.org/10.1117/12.306241

[13] A. Nilă, M. Baibarac, A. Matea, R. Mitran, and I. Baltog, "Exciton–phonon interactions in the $Cs_3Bi_2I_9$ crystal structure revealed by Raman spectroscopic studies," *physica status solidi (b)*, vol. 254, no. 4, p. 1552805, 4 2017. [Online]. Available: https://doi.org/10.1002/pssb.201552805

[14] S. Öz, J. C. Hebig, E. Jung, T. Singh, A. Lepcha, S. Olthof, F. Jan, Y. Gao, R. German, P. H. van Loosdrecht, K. Meerholz, T. Kirchartz, and S. Mathur, "Zero-dimensional $(CH_3NH_3)_3Bi_2I_9$ perovskite for optoelectronic applications," *Solar Energy Materials and Solar Cells*, vol. 158, pp. 195–201, 12 2016. [Online]. Available: https://doi.org/10.1016/j.solmat.2016.01.035

[15] S. Devasia, S. Shaji, D. A. Avellaneda, J. A. A. Martinez, and B. Krishnan, "In situ crystallization of 0D perovskite derivative $Cs_3Bi_2I_9$ thin films via ultrasonic spray," *Journal of Alloys and Compounds*, vol. 893, p. 162294, 2 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092583882103704X

[16] R. Waykar, A. Bhorde, S. Nair, S. Pandharkar, B. Gabhale, R. Aher, S. Rondiya, A. Waghmare, V. Doiphode, A. Punde, P. Vairale, M. Prasad, and S. Jadkar, "Environmentally stable lead-free cesium bismuth iodide (Cs3Bi2I9) perovskite: Synthesis to solar cell application," *Journal of Physics and Chemistry of Solids*, vol. 146, p. 109608, 11 2020. [Online]. Available: https://doi.org/10.1016/j.jpcs.2020.109608

[17] D. Phuyal, S. M. Jain, B. Philippe, M. B. Johansson, M. Pazoki, J. Kullgren, K. O. Kvashnina, M. Klintenberg, E. M. Johansson, S. M. Butorin, O. Karis, and H. Rensmo, "The electronic structure and band interface of cesium bismuth iodide on a titania heterostructure using hard X-ray spectroscopy," *Journal of Materials Chemistry A*, vol. 6, no. 20, pp. 9498–9505, 5 2018. [Online]. Available: https://pubs.acs.org/doi/10.1021/acsaem.0c01129

Real-Time Autonomous Vehicle Sensor Performance Assessment in Adverse Weather Conditions

1st Stanislav Svědiroh
Department of Control and Instrumentation
Brno University of Technology
CEITEC - Central European Institute of Technology
Brno, Czechia
stanislav.svediroh@ceitec.vutbr.cz

2nd Luděk Žalud
Department of Control and Instrumentation
Brno University of Technology
CEITEC - Central European Institute of Technology
Brno, Czechia
ludek.zalud@ceitec.vutbr.cz

***Abstract*—The future of the automotive industry appears to be intricately linked to Advanced Driver Assistance Systems (ADAS) and various levels of Automated Driving Systems (ADS). Over the years, numerous companies have incorporated sensors into their vehicles, however, none have yet achieved the development of a completely robust and self-aware system capable of operating safely in adverse weather conditions. To guarantee safety, the vehicle must possess an awareness of its environment and the current performance of its sensors. This includes the ability to detect not only current weather conditions such as rain, fog, haze, and snow, but also smoke, soiling from various sources, and extreme lighting conditions such as glare or low light. It is crucial for the vehicle to detect these conditions in real-time without delaying decision-making systems. This study summarises the effects of various environmental threats on commonly used sensors in ADAS or ADS and proposes algorithms to detect degrading sensor performance, which can then be integrated into the sensor fusion framework utilised in the creation of the vehicle's local map. The ultimate aim of such a system is to accurately detect and report sensor degradation, enabling subsequent sensor fusion and path-planning algorithms to modify the vehicle's behaviour and minimise unreasonable risk.**

***Index Terms*—ADAS, ADS, Adverse Weather, Sensor Performance Assessment**

I. Introduction

For the autonomous vehicle to remain safe it needs to be able to have an understanding of its own limits. That means to continually evaluate the outputs of its sensors and look for signs of degraded performance when compared to the nominal properties. This can be, in some situations, computationally very expensive and some possible signal degradations due to the adverse weather conditions would be very hard to even expose. Typically, such a vehicle has sensors of various modalities such as RGB Cameras, IR Cameras, LiDARs, Radars, Ultrasonic sensors, etc. Each of these sensors have different strengths and are more or less susceptible to different weather-imposed threats. There have been numerous papers that estimated the effect of some weather phenomena on sensor performance in isolation. Filgueira et al. [1] has measured the influence of different rain intensities on LiDAR sensors. They've analysed the points returned from various surfaces common when driving on the structured road and plotted the range, intensity and number of points change with respect to the rain intensity. From the measurements, it is clear that wet surfaces and rainfall causes the decrease in returned signal intensity and in the number of detected points due to that. Carballo et al. [2] has tested multiple LiDARs in the weather chamber with decreased visibility due to the homogeneous fog from 10 m to 80 m, precipitation of 30 mm/h and 80 mm/h, and glare caused by 6 kW xenon light. Results show that dense fog can cause a lot of fake detections around the car as the intensity of the echo from the fog is greater than the intensity returned from real objects. Experiments with the strong visible light source also show that such interference can cause missing points.

When it comes to RGB cameras, the effect of weather on their performance can be easily imagined since the visual information is something we, humans, are used to processing all the time. Fog, smoke, rain and snow all have the same effect on the 2D camera image - the edges become less sharp and colours less saturated. This consequently leads to less detections or smaller detection confidence or accuracy [3]. Strong light like glares can easily saturate the image pixels causing clipping and a loss of information the same way as low light does on the other side of the lightness values. At the same time, RGB or IR cameras can be easily incapacitated by a fallen leaf, mud or even a single water drop on its lens.

Radars are one of the sensors that are more invariant to weather conditions, but the output data are too sparse to create a reliable environmental map ready to be used in autonomous vehicles.

Goelles et al. [4] presents a survey of Fault Detection, Isolation, Identification and Recovery (FDIIR) methods for automotive perception sensors. Their classification of sensor faults related to weather can be split into two groups. *Layer on sensor* (dirt, water, ice, etc.) and *unfavourable environmental*

The work has been performed in the project ArchitectECA2030: Trustable architectures with acceptable residual risk for the electric, connected and automated cars, under grant agreement No 877539/8A20002 [14]. The work was co-funded by grants of Ministry of Education, Youth and Sports of the Czech Republic and Electronic Component Systems for European Leadership Joint Undertaking (ECSEL JU). The work was supported by the infrastructure of RICAIP that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857306 and from Ministry of Education, Youth and Sports under OP RDE grant agreement No CZ.02.1.01/0.0/0.0/17_043/0010085.

conditions (dust, smoke, rain, fog, hail, snow, etc.). Other types of sensor faults can be their sub-component failing, mounting issues, crosstalk or security hack. They've also categorised FDII for perception sensors into several methods based on what is the examined sensor output compared with. For this paper following methods are relevant:

- Monitoring the sensor output,
- Comparison to other sensor of same type
- Comparison to other sensor of different type

Surveys are also carried out to aggregate the literature and quantify the effects of weather on ADS sensors such as [6].

The paper is structured as follows. Section two describes the datasets, used sensors, and data fusion framework we've been working with to implement sensor performance assessment algorithms. Section three proposes an architecture for real-time sensor performance evaluation. Section four discusses RGB camera weather susceptibility and algorithms used to detect its degraded performance. Following section discusses the same for IR camera sensors and finally section six talks about LiDARs. Radars are not included since they can be considered very resilient in realistic weather conditions [5]. Lastly, we discuss the achieved results and identify possible ways to improve robustness of the framework.

II. Dataset, sensors and framework used

Throughout this paper, we are primarily using dataset collected at Brno University of Technology called *Brno Urban Dataset* and its *Winter Extension* [8] [9]. We also use LIBRE Dataset [2].

The Brno Urban Dataset comprises 4x RGB 1920x1200 px 10 fps cameras, two of them facing forward, two of them covering the sides. An IR camera FLIR Tau2 with the resolution of 640x512 px and 30 fps. The cameras are mounted so they all share part of their field of view with the adjacent ones so the detections and key points can be matched between sensors. The dataset also contains two side-mounted Velodyne HDL-32 LiDARs slightly tilted around forward-pointing axis to cover the close surroundings of the vehicle. One Livox Horizon LiDAR and mmWave AWR1642 radar are also mounted in the front of the vehicle to increase point density where it is needed the most. To provide positioning and time, IMU and GNSS are also present on the test vehicle.

The sensor fusion framework used for the sensor performance assessment is called *Atlas Fusion* [7]. The current intention of this framework is to pass raw sensor data incoming through the common pipeline and pass it to a *fail checker* which then redirects it to the appropriate *dedicated sensor fail checker* that outputs the sensor reliability score which is used, alongside the raw sensor data, in sensor fusion algorithms to create a local map. This would only test each sensor in isolation without taking advantage of the knowledge gained from other sensors.

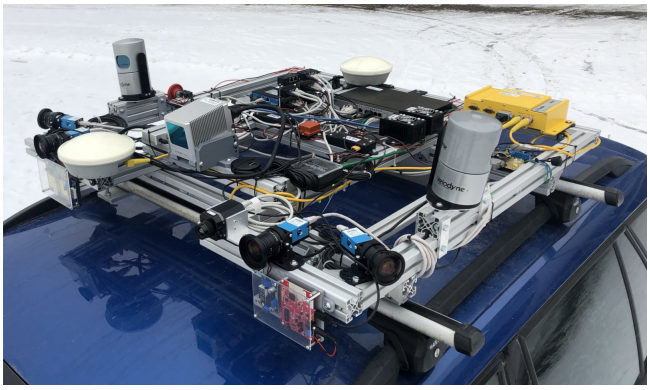


Fig. 1. Detailed view of the sensory framework used in Brno Urban Dataset [9]

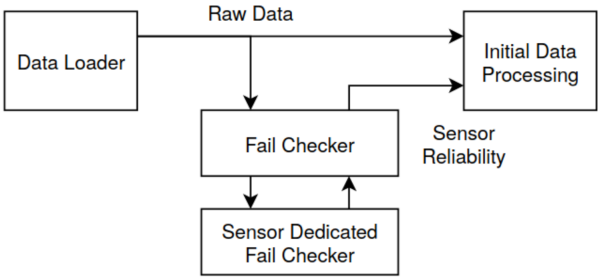


Fig. 2. Fail checker pipeline of Atlas Fusion framework [7]

III. Proposed performance assessment architecture

In the spirit of FDII discussed in [4] and architecture proposed in [10] we propose an architecture that not only verifies the sensor performance in isolation, but also compares the sensor output with sensors of the same and different modalities. While doing so, it is also building a *Environment condition assessment* model that accumulates and persists all information extracted to build an understanding about the current weather, road surface state, ego vehicle status, etc.

This enables us, in theory, to use every sensor mounted on the vehicle to validate every other sensor, thus not relying on the previously proposed fail checker layer to be absolutely robust. From the built *Environment condition assessment* module we can then make quick presumptions about each sensor performance based on the current weather status.

Figure 3 shows such a system in a block representation. The data loader, data processing and fusion backbone stays the same, but the *Sensor dedicated performance assessment* is now updating *Environment condition assessment* with every call. Instead of just being called when passing the raw data in, it is now also invoked with detections/key points etc. that have been extracted from the raw data.

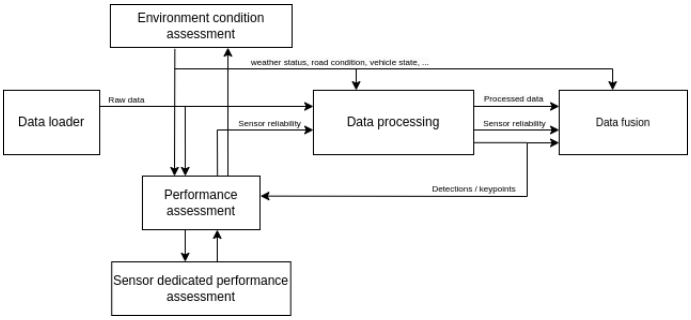


Fig. 3. Proposed sensor performance assessment pipeline

IV. RGB Camera performance assessment

Decreased visibility, low light, glare (from the Sun or artificial sources) and occlusion / soiling all affect camera performance greatly. Decreased visibility, which is the common result of all weather phenomena like rain, snow, fog, smog or smoke, has the effect of blurring image edges (low pass filter) and desaturating the scene. Ideally, the vehicle would be able to measure *Meteorological Optical Range* (MOR) or visibility in other words, which is measured in metres. Its value only depends on the surrounding air transparency and doesn't consider ambient light level. That means visibility can still be high on a clear night. Algorithm that would measure (or estimate with high precision) visibility from a RGB camera image without having reference objects in the known distance in the frame at all times is an ill-posed problem. Many research groups are working towards defining such an algorithm, because its direct measurement would add even more cost to the vehicle. We are still considering multiple options, but for the time being we are converting a small region of interest around *Vanishing Point* (VP) to a frequency domain and calculating power spectrum slope inspired by [11].

For the robust VP detection, many methods have been proposed with [12] being very promising. Ultimately, due to performance considerations, we settled on our own IMU-guided method where by default VP is located in the centre of the image and is shifted to either side when the vehicle is turning. The resulting power spectrum slope is then compared with furthest detections of the front LiDAR to estimate visibility. Figure 4 shows visualisation of visibility *score* calculation. Since our implementation is not yet properly calibrated, we're mapping the calculation result into the $\langle 0, 1 \rangle$ range where 1 means great visibility.

Cameras can also suffer from translucent or opaque soiling. Translucent soiling, again, results in a blurred image. We split the RGB image by a regular grid into 64 parts and we expanded the previously mentioned algorithm of calculating power spectrum slope to be calculated on each of these parts. This results in a heatmap that depicts whether high frequency information is present in that image part. When, for a defined time period, there are low values detected and other image sensors that are possibly covering that region do detect higher frequencies, this image spot can be flagged as soiled.



Fig. 4. visibility score calculation visualization

For an opaque soiling and glares, we run histogram analysis on an image split by a regular grid into 144 parts. When the histogram is clipping the low luminance values, after a defined period, the image part can be flagged as soiled or occluded. On the other hand, when high luminance values are clipped in the histogram representation, the spot is flagged as glared. To aid the glare detection, we're calculating the sunrise and sunset times with the date and position data from IMU and GNSS to calculate the sun position w.r.t the vehicle heading. Figures 5 6 show the occlusion (blur) and glares being detected in a single frame.

As proposed, the performance assessment is extracting data from a single RGB image, but it is also taking advantage of the common field of view of other RGB cameras and using LiDAR data to estimate visibility score. This system is inherently more robust than it would be when only working with one sensor data frame.



Fig. 5. The top heatmap showing high frequencies being detected in the image

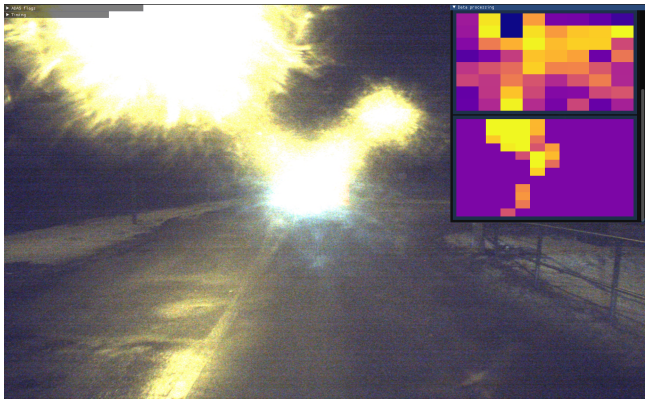


Fig. 6. The bottom heatmap showing glares being detected in the image

V. IR Camera performance assessment

Infrared cameras share a lot of common properties with their RGB counterparts. They can be completely blocked by an opaque object or blurred out by a transparent one. Since the sensor is not sensitive to radiation in the visible spectrum, it is not suffering from glares. The IR radiation can also penetrate fog or even dense smoke to a certain degree, but it gets partly scattered and absorbed, so the received power is smaller and the image can appear blurry. Some materials that are transparent in the visible spectrum are not in *Long Wave Infrared* (LWIR). Glass, water and some plastic materials can absorb most of this radiation, which is why such camera could not be installed behind the car's windshield. High humidity levels can also, like with the RGB camera, cause fog condensation on the lens and almost completely block all incoming radiation. Since LWIR spectrum is absorbed by water, heavy rain can corrupt a lot of information normally present in the image. Checking the *environmental condition assessment* module for rain or humidity status can help with classifying IR sensor performance.

Looking at the temperature variations in the whole image, or even better just around neural network detections of cars and living creatures made from RGB images will tell us a lot. When a car or a person is detected just in front of the ego vehicle, there should be some temperature variation in the IR frames too. Figure 7 shows RGB camera frame Yolo detections matched between frontal cameras reprojected into the IR camera frame. When the temperature variation in those spots is too low, the camera lens may be soiled.

VI. LiDAR performance assessment

LiDAR scanners help us understand the scene geometry better. While it is possible to reconstruct the scene depth by only using cameras with overlapping fields of view, as we already discussed, cameras are very susceptible to almost all adverse weather conditions. They can even be completely blinded by a single water drop. On the other hand, LiDAR 3D scanners are more robust when it comes to weather. They work with the light spectrum just between the visible and infrared light (850-950 nm or 1550 nm). 1550 nm scanners



Fig. 7. RGB YOLO Detections reprojected to the IR camera frame

have the advantage of higher optical power being allowed and considered eye-safe, so they can penetrate fog or smoke more easily. Their disadvantage when compared to more prevalent 850 - 950 nm is higher water extinction coefficient thus higher susceptibility to rain and wet surfaces. Since both wavelengths are in the band of low solar irradiance, they can be practically unaffected by sunlight and glares. [2] mentions missing points when a strong 6 kw xenon light source is pointed directly into the LiDAR, but as stated, such illuminance is not expected even at the horizon.

When the fog is dense, the scanners tend to fail as the strongest echo returned is now not from the object itself, but some particle along the way. This causes false detections close to the camera in case of dense fog, or where there could be nothing in the case of local smoke patch (car exhaust) Some scanners can process multiple echos so it can be defined, whether the strongest return or last return will be processed. Last return can be beneficial in smoke conditions.

As these sensors generate hundreds of thousands of 3D points per second, the real-time validation options are limited due to the sheer amount of data. The main parameters of a point cloud are measurement accuracy, point density and measurement range - all of those are affected by adverse weather conditions. Measurement accuracy is something that cannot be simply verified as that would require objects of known position and size in the scene. Low point density in certain regions of interest can indicate specific issues. For example, when there is a low number of points in front of our test vehicle detected by the Livox Horizon sensor, we can assume that the road surface is wet or iced. This is because the light is scattered away from the sensor and only a small part is reflected back. Voids in front of and behind the vehicle can be indicative of a snow swirl [13]. Lastly, measurement range decreases considerably below the manufacturer claims when the visibility decreases - so this too can be used to indicate

weather status.

Comparing LiDAR scans to data of different modality is also possible. After filtering out the ground plane, resulting points can be grouped together by *Euclidean Cluster Extraction*. Those point clusters represent possible obstacles. To reduce the amount of data needed to describe those obstacles even more, they can be converted to the bounding boxes in 3D space. Comparing RGB camera detection positions to those bounding boxes conclusions can be made about how many obstacles are ”shared” between those modalities which tells us more about the quality of LiDAR scans. Figure 8 shows how RGB camera detections made by YOLO neural networks are matched with Lidar detections. Matched lidar detections are coloured with the colour representative of the YOLO class (blue for cars).

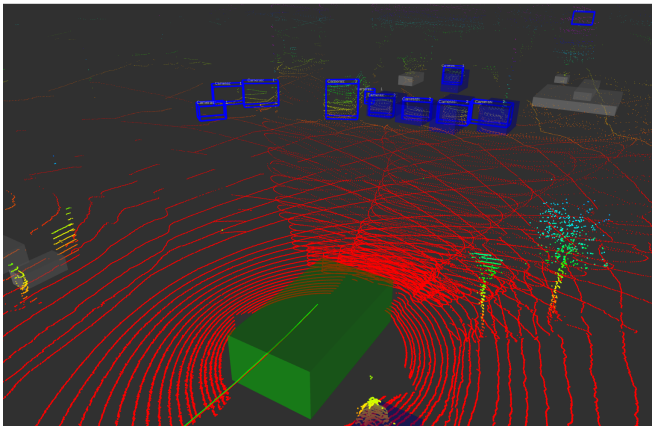


Fig. 8. Lidar (colored boxes) detections being matched with RGB camera frustums (colored outlined boxes)

VII. Conclusions and further work

In this paper, we've proposed an improved architecture for real-time sensor performance assessment. Our main focus is to implement algorithms capable of building a probabilistic model of the environment called *Environment condition assessment*. This model should ultimately contain a very precise description of the current weather status so the vehicle is able to presume performance of its sensors without having to inspect every sensor output precisely. This is only possible because we have a pretty good understanding of how individual sensors react to adverse weather conditions. Our goal is to expand our dataset (*Brno Urban Dataset* [8]) to include more situations of low visibility. At the same time, we are trying to test our performance assessment algorithms on various publicly available datasets to cover a wider spectrum of possible situations. Atlas Fusion framework currently doesn't use the performance metadata in the sensor fusion stage, so we only print out the status information as shown in figure 9. In the future, we would like to take this metadata into account so that bad weather doesn't affect sensor fusion output.

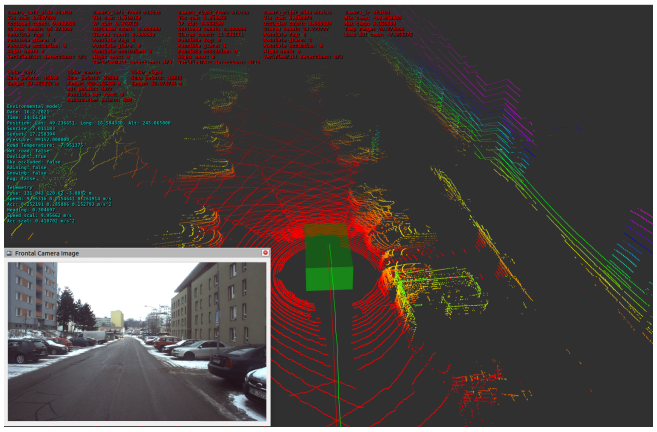


Fig. 9. Performance assessment outputs being shown alongside the local map

References

[1] A. Filgueira, H. González-Jorge, S. Lagüela, L. Díaz-Vilariño, and P. Arias. Quantifying the Influence of Rain in LiDAR Performance. Measurement 95 (January 1, 2017): 143-148. https://doi.org/10.1016/j.measurement.2016.10.009.

[2] A. Carballo, J. Lambert, A. Monrroy, D. Wong, P. Narksri, Y. Kitsukawa, E. Takeuchi, S. Kato, and K. Takeda. LIBRE: The Multiple 3D LiDAR Dataset. In 2020 IEEE Intelligent Vehicles Symposium (IV), 1094-1101, 2020. https://doi.org/10.1109/IV47402.2020.9304681.

[3] F. Reway, W. Huber, and E. Parente Ribeiro. Test Methodology for Vision-Based ADAS Algorithms with an Automotive Camera-in-the-Loop. In 2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES), 1-7, 2018. https://doi.org/10.1109/ICVES.2018.8519598.

[4] T. Goelles, B. Schlager, and S. Muckenhuber. Fault Detection, Isolation, Identification and Recovery (FDIIR) Methods for Automotive Perception Sensors Including a Detailed Literature Survey for Lidar. Sensors 20, no. 13 (January 2020): 3662. https://doi.org/10.3390/s20133662.

[5] S. Zang, M. Ding, D. Smith, P. Tyler, T. Rakotoarivelo, and M. A. Kaafar. The Impact of Adverse Weather Conditions on Autonomous Vehicles: How Rain, Snow, Fog, and Hail Affect the Performance of a Self-Driving Car. IEEE Vehicular Technology Magazine 14, no. 2 (June 2019): 103-111. https://doi.org/10.1109/MVT.2019.2892497.

[6] Y. Zhang, A. Carballo, H. Yang, and K. Takeda. Perception and Sensing for Autonomous Vehicles under Adverse Weather Conditions: A Survey. ISPRS Journal of Photogrammetry and Remote Sensing 196 (February 1, 2023): 146-177. https://doi.org/10.1016/j.isprsjprs.2022.12.021.

[7] A. Ligocki, A. Jelinek, and L. Zalud. Atlas Fusion – Modern Framework for Autonomous Agent Sensor Data Fusion. arXiv, April 20, 2022. https://doi.org/10.48550/arXiv.2010.11991.

[8] A. Ligocki, A. Jelinek, and L. Zalud. Brno Urban Dataset - The New Data for Self-Driving Agents and Mapping Tasks. In 2020 IEEE International Conference on Robotics and Automation (ICRA), 3284-3290, 2020. https://doi.org/10.1109/ICRA40945.2020.9197277.

[9] A. Ligocki, A. Jelinek, and L. Zalud. Brno Urban Dataset: Winter Extension. arXiv, July 29, 2021. https://doi.org/10.48550/arXiv.2106.02952.

[10] T. Ö. Şahin, S. Hörmann, B. Schäufele, and F. Kuhnt. Automated Vehicle System Architecture with Performance Assessment. In 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), 1-8, 2017. https://doi.org/10.1109/ITSC.2017.8317862.

[11] R. Spinneker, C. Koch, S. Park, and J. Jeongsuk Yoon. Fast Fog Detection for Camera Based Advanced Driver Assistance Systems. In 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 1369-1374, 2014. https://doi.org/10.1109/ITSC.2014.6957878.

[12] C. N. Khac, Y. Choi, J. H. Park, and H. Jung. A Robust Road Vanishing Point Detection Adapted to the Real-World Driving Scenes. Sensors 21, no. 6 (January 2021): 2133. https://doi.org/10.3390/s21062133.

[13] M. Jokela, M. Kutila, and P. Pyykönen. Testing and Validation of Automotive Point-Cloud Sensors in Adverse Weather Conditions. Applied Sciences 9, no. 11 (January 2019): 2341. https://doi.org/10.3390/app9112341.

[14] J. Reckenzaun, S. Solmaz, T. Goelles, M. Hilbert, D. Weimer, P. Mayer, A. Chromy, U. Hentschel, N. Modler, M. Toth, M. Hennecke. Robust perception systems for automated, connected, and electrified vehicles: advances from EU project ArchitectECA2030. Transport Research Arena (TRA2022), Lisbon, Portugal, 2022, (To appear at Elsevier Transportation Research Procedia in 2023).

Model of Distribution System for Energy Testbed

1st Antonín Bohačík
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
Antonin.Bohacik@vut.cz

2nd David Kohout
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
David.Kohout@vut.cz

***Abstract*—This paper focuses on the design of a distribution system model for the Energy Polygon already created at Brno University of Technology. The main contribution of this article is to present an expansion model for the Energy Polygon to represent the full energy chain from production to consumption, including commonly used communication protocols, technologies, and devices. During the design process, emphasis was placed on creating a complex system that would emulate the behavior of the real distribution system as a whole.**

Currently, the Energy Polygon implements protocols IEC 60870-5-104 and GOOSE (Generic Object Oriented Substation Event) from standard IEC 61850. The original model will be expanded with multiple modules implementing other protocols from IEC 61850, charging station protocol OCPP (Open Charge Point Protocol), protocols used for renewable resources like Modbus and Sunspec, and OPC (Open Platform Communications), or Smart Metering protocol DLMS (Device Language Message Specification).

***Index Terms*—Distribution System, DLMS, Energy Polygon, IEC 60870-5-104, IEC 61850, OpenMUC, SCADA**

I. Introduction

The electric power transmission system is one of the most important parts of any modern state. It can be classified as the critical infrastructure that is essential for the operation of the state. Therefore, it is important to ensure that this system and all the used technologies are modernized accordingly. The modernization process is dependent on creating and deploying new Information and Communication Technologies. To help with the development and testing, it is appropriate to use the so-called polygons. Polygons are a testing environment that can be used to test new technologies and devices, to simulate particular processes, to emulate devices, or to test the security and safety aspects of the elements or the whole system.

This article presents an extension model design of the distribution system that would represent a functional sample for testing power equipment. The model builds on the already created Energy Polygon at Brno University of Technology (BUT) to achieve the full energy chain from the point of generation to the point of consumption. This extension model was designed to be highly modular with the possibility to easily add new expansion modules.

A. Related Work

The creation of testing polygons is a common practice in many universities, but they are also created in companies focused on creating specific devices. Testing polygons provide an environment that is suitable for online or offline testing, as a specific device cannot be disconnected in a real system without disrupting the normal operation of the system.

In terms of implementation, we can divide polygons into (i) physical, (ii) virtual, and (iii) hybrid. Creating a fully functional physical polygon is very time, space, and cost-consuming. Therefore, some institutions lean towards virtual polygons, which are largely implemented using specific software (e.g., Matlab) or equipment (e.g., RTDS – Real Time Digital Simulator). Their clear advantage is their cost, which can be in the order of tens of percent compared to an equivalent physical polygon. The most commonly used solutions are hybrid polygons, which contain some physical components, but the basic physical units are simulated (e.g., current and voltage level) and passed to the physical device. Devices for generating physical units or emulating whole devices are, for example, RTDS, and Opal-RT. With such devices, simulations can be generated for many polygons in both real and simulated time. The advantage of real (physical) simulation is its illustrative nature and the accuracy of the data outputs. Another advantage is the possibility of real-time interaction, which can be used, e.g., for student or staff training. The disadvantage is paradoxically the time consumption of the simulation itself.

Standalone polygon can be used for testing new devices, which are not ready for deployment into the real operation. In this way, the device can be tested in pseudo-real environment without the possibility of breaking or crashing the real system. This can also reveal any operational or structural defects, which can be resolved before the real deployment. The same procedure can be used to test the ever increasing demands on the cybersecurity. The device is tested with multiple security testing techniques to reveal any possible vulnerabilities.

Depending on their design and type, some polygons can be implemented with virtual components, such as a prototype device that has not yet been actually constructed. From the perspective of the development process, this is the best possible approach, as it allows us to discover additional requirements for the device and better tailor the hardware for the final system according to the measured needs.

Based on the mentioned categorization, a survey of existing Energy Polygon was conducted. A comparison of the polygons found is presented in Table I. For the found polygons, an analysis of the implemented properties was performed. The first searched parameter was whether it contained transmission

TABLE I
COMPARISON OF EXISTING ENERGY POLYGONS

Institutions	Ref.	Transmission system	Distribution system	Security analysis	Energy protocols
Aalborg Univ	[1]	✗	✓	✓	DNP3, IEC 61850 (GOOSE, SV), Modbus
Brno University of Technology*	[2]	✓	✗	✓	IEC 60870-5-104, IEC 61850 (GOOSE)
Florida State Univ	[3]	✗	✓	✓	IEC 61850 (GOOSE)
Iowa State Univ	[4]	✓	✓	✓	DNP3, IEC 61850, IEEE C37.118, Modbus
North China Electric Power University	[5]	✓	✗	✗	DNP3, IEC 60870-5-104
Pacific Northwest National Laboratory	[6]	✗	✓	✗	DNP3, IEEE C37.118, IEC 61850 (GOOSE, MMS)
Texas A&I Univ	[7]	✓	✗	✓	DNP3, IEC 61131-3
Univ Western Macedonia	[8]	✗	✗	✓	IEC 60870-5-104

**Testbed discussed in this paper.*

and distribution system logic, or which of these two parts were implemented. Subsequently, their purpose was examined, which could be purely for generation communication data or for deeper cybersecurity analysis. Polygons that implement advanced security mechanics, contain specific cybersecurity tools, or contain educational scenarios for understanding cyberattacks and cybersecurity of devices or entire systems fall into this category. The last parameter lists the energy communication protocols that have been implemented within the polygons.

B. BUT Energy Polygon

The Energy Polygon of the Czech Republic Transmission System was created at BUT for educational, simulation and testing purposes (see Figure 1). This polygon can simulate pseudo-real data traffic between individual stations and the monitoring center, that is based on open-source software for SCADA (Supervisory Control And Data Acquisition) systems called OpenMUC[1].

Currently, the implemented communication is emulated using the open-source library *lib60870-2.3.0*[2], which implements the IEC 60870-5-104 protocol (IEC 104 in short). This library also contains multiple mechanisms for securing data traffic using TLS tunnels according to the IEC 62351 standard. Each station in the polygon is implemented using a Raspberry Pi 3B+ single-board computer (RPi for short), which serves as the border communication gateway at substations or power plants. These stations act as simulators to generate values for physical units such as transformation voltage, current, and frequency, as well as temperature of the transformer cooling liquid or ambient temperature. All values are generated using a pseudo-random function, and their range can be defined in the configuration file. Currently, the ranges are set in accordance with the real ranges of the stations defined in the *Kodex přenosové soustavy ČR*[3] (translated – Code of the Transmission System of the Czech Republic).

All mentioned values are periodically send to the monitoring center using the IEC 104 protocol. In addition, the stations are designed to adjust data communication in case of operational events, such as undervoltage, overvoltage, or system failure. The stations are also able to respond to commands sent by the monitoring center, such as commands to safely disconnect stations. To display the most important data, all stations are equipped with an information display IIC I2C OLED (see Figure 1). Finally, all stations also include an internal logging system used for reverse analysis of transmitted data and simulated values.

The entire polygon contains more than 50 elements that must be continuously checked for proper functionality. The RPi server stations are periodically queried for their state or they can be irregularly queried in the form of user remote commands. These messages cannot be considered as valid data communication. For this reason, a separate network has been created within the polygon, which allows for communication with individual stations without direct intervention in the simulated data communication. This separate network can be used to change the station's internal parameters without interfering with data communication or to run predefined operational scenarios. The management of this separate management network is provided by a web application based on the open-source web framework Django[4].

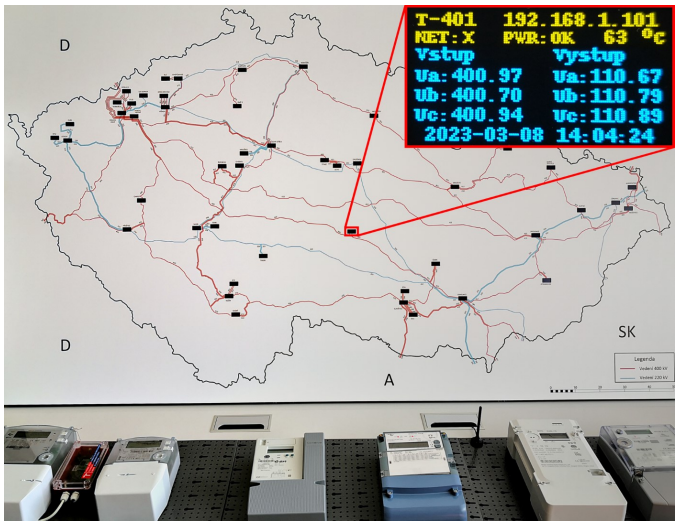


Fig. 1. BUT Energy Polygon

[1]Available from https://www.openmuc.org/
[2]Available from https://www.mz-automation.de/
[3]Available from https://www.ceps.cz/cs/kodex-ps
[4]Available at https://www.djangoproject.com

II. Electrical Power System

The electrical power system is a complex network of electrical devices that are connected to handle the generation, transmission, and consumption of electrical energy. It can be divided into three main components: generation sources, transmission system, and distribution system [9]. Generation sources refer to the different methods used to generate electricity such as thermal, nuclear, or renewable energy sources [10]–[12]. The transmission system is responsible for transferring high amounts of electrical energy over long distances, while the distribution system supplies energy to nearby households or industrial enterprises. The main goal of the electrical power system is to ensure reliability and safety of operation, optimize electricity production and transmission, integrate renewable energy sources, management, and regulation of voltage in the electrical grid [13]. The voltage levels discussed in the purposed model are based on electrical infrastructure of the Czech Republic (the model use only representative levels 110 kV, 22 kV and 230 V).

A. Electrical Power Transmission

Electrical power transmission is used to transmit large amounts of electricity over long distances at Very High Voltage (VHV, usually over 110 kV) with minimal energy losses. This system includes transformer stations, VHV power lines and other devices that enable the efficient and reliable transmission of electricity from generation sources to distribution networks and large industrial plants. VHV distribution networks are usually located near the production power sources and are used to distribute electricity to power lines. The transformer stations then allow the voltage to be changed and adapted for VHV transmission. This system is very important for the reliable and safe operation of the electric power system, and its proper management and maintenance is key to ensuring the stability and quality of electricity supply. In the Czech Republic, the transmission system is managed by ČEPS (translated – Czech Electricity Transmission System), which ensures operation and maintenance of the transmission system.

B. Electric Power Distribution

The electric power distribution is used to distribute electricity over shorter distances and at a lower voltage than in the transmission system. The distribution system includes High Voltage (HV) and Low Voltage (LV) distribution networks, transformation stations, distribution transformers, power lines and other equipment that enable the distribution of electricity to households, small and medium-sized industrial plants and other consumers. LV distribution grids are located in places where electricity is distributed directly to appliances, such as households and smaller industrial plants. The transformer station in the distribution system serves to transform the voltage from high to low voltage. Distribution transformers then serve to adjust the voltage for consumers. In the Czech Republic, the distribution systems are mainly managed and controlled by the companies ČEZ and E.ON, which are among the largest energy companies in Europe.

C. Smart Metering

The distribution system includes the transmission of electricity to the customer, and the point where electricity is supplied to the customer is the electricity meter, which measures mainly consumption. Old electricity meters are currently being replaced with new smart meters, which have two-way communication capabilities. This is used to manage and control the distribution system more precisely and in real-time. It also good for customers to see their consumption in real-time, and to have the ability to make their consumption more effective. In the future, smart meters can control even appliances with higher electricity consumption (e.g., heating, water heating, even charging vehicles) to effectively use electricity during times of surplus energy. Various algorithms and applications are also used in smart metering, making it possible to optimize energy consumption and reduce energy costs. The main standard used in smart meters is DLMS (Device Language Message Specification).

III. Design of the Distribution System

The distribution system model was designed with the goal of enabling effective communication among all parts of the system to achieve an accurate simulation of a real distribution network. As part of the model's design, a comprehensive analysis of the actual electrification system in the Czech Republic was conducted. However, the distribution system's topology was not considered during the model's design. The primary purpose of the model is to simulate data communication among devices (not the distribution of electrical energy).

The proposed model of the distribution system is shown in Figure 2, which represents a simplified energy chain from the point of production (power plant) to the point of consumption (customer, company). The model is divided into three basic parts according to the simulated voltage level: VHV – transmission (A), HV – distribution (B) and LV – end (C) system. Each part represents a sub-system that contains specific devices and communication protocols. Since the model cannot be created based on energy requirements, all voltage levels (except for the 230 V level) are replaced by service communication (indicated by orange dashed lines). This communication is separate from data communication and is used to transfer simulated values between individual devices, which would be measured in real systems.

In order to monitor the entire system, the OpenMUC software will be used in the proposed model. Its special feature is modularity and support for a wide range of communication protocols. This software will represent the supervisory SCADA system as well as in the realized Energy Polygon. Therefore, the existing software will only be expanded with additional functional parts. Each part will represent the supervision of a specific part according to individual voltage levels. Communication between the devices and the monitoring system will serve not only for monitoring, but also for remote control of the device using a data network.

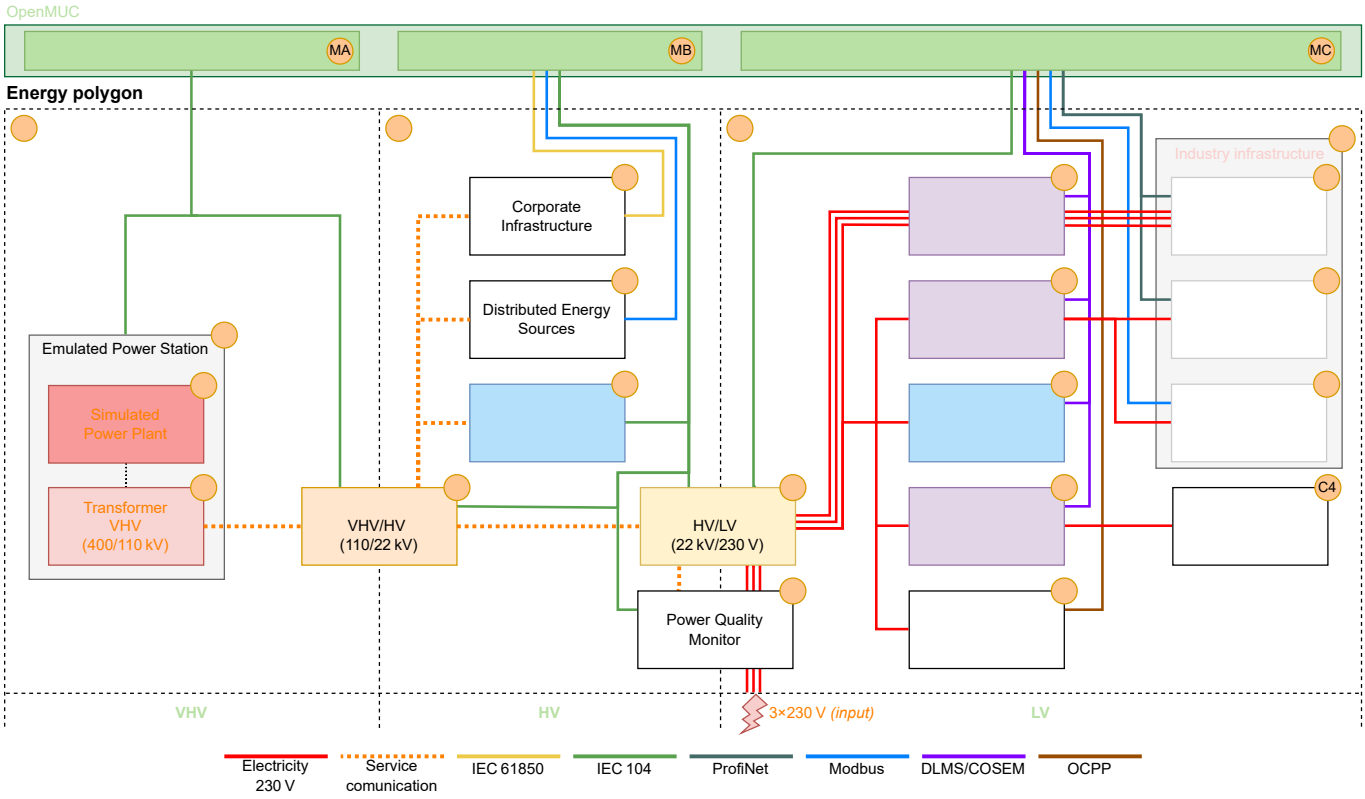


Fig. 2. Distribution system diagram

A. VHV Section

This first part (A) has already been implemented and described in the article [2]. However, as part of the implementation, there will be minor modifications to the software part of the emulated stations. The main change will be in the implementation of service communication to a lower layer, more precisely to the VHV/HV transformer (AB). This communication will represent the input quantities for the subsequent simulation of the transformation.

B. HV Section

The first higher part of the distribution system (B) starts after the output of the VHV/HV transformer (AB) and ends with the input of the HV/LV transformer (BC). This transformer will have a virtual counterpart (B3) using an RPi device, similar to the emulated power station (A1). This will represent the boundary communication element of the station for communication with the monitoring center (OpenMUC) (MB). Additionally, it will process service data from subordinate functional blocks or physical devices. The goal of processing this service data is to ensure a similar behavior of the system without actual energy transmission, as in a real distribution network. A notable feature of this transformer will be its connection with a physical transformer (BC), which will be connected to a real electrical network and the HV input side will be partially emulated with RPi. This connection will ensure the actual transfer of energy to the LV part and the creation of a fully functional physical system. This section will also include one power quality monitor MEg44PAN (B4), which will measure the quality of the output voltage on the individual phases. Furthermore, the modularity of the system allows for easy expansion with additional functional blocks.

Corporate Infrastructure (B1) represents the device at the boundary of larger enterprise systems directly connected to the 22 kV system. In the Czech Republic, companies such as Škoda Auto, Třinecké železárny, and Unipetrol RPA in Litvínov fall under this category. The ABB device (REF/REC/RER615) for measuring 3-phase voltage or an industrial controller from Siemens (SICAM A8000, S7-1200) can be used in this block. To establish reliable data communication, a signal generator will be used to feed individual inputs of the measuring devices. These devices support a wide range of industrial protocols, including IEC 104, DNP3, Modbus RTU/TCP, OPC, IEC 61850, ProfiNet, and ProfiBus, all of which are supported within OpenMUC.

Distributed Energy Sources (B2) represents sources of energy (wind, solar, or hydropower), which, unlike other blocks, supplies energy to the virtual network. The service communication is used to inform the VHV/HV transformer (BC) how much energy is supplied to the system. The simulated energy value is pseudo-random (depending on the time of day and the simulated weather). A whole range of energy devices can be used for this block, e.g., Siemens Scalance M875-2 device supporting, in addition to classic communication protocols

such as SNMP, HTTPS, SSH, as well as industrial protocols Modbus/TCP, ProfiNet, ProfiBus. The second of many suitable devices for this block is the Hirschmann MSP40 device, which supports protocols Modbus/TCP or ProfiNet.

Virtual transformer HV/LV (B3) represents a simplified model of the entire LV section. It thus allows creating different system load patterns without the need for physical devices.

C. LV Section

The last designed part (C) includes all devices connected to the output of the HV/LV transformer (BC). Similar to the previous part, this part is divided into individual functional blocks, where each one represents different element of a real system. All elements of the system communicate with the monitoring center using specific communication protocols to ensure independent monitoring.

Smart meters (3-phase (C1) and 1-phase (C2)) are being added to the polygon because they are a part of the distribution network and their implementation is required by legislation. The main communication protocol used for monitoring and remote management is the DLMS protocol. Through these electricity meters, industrial polygons (D) will be connected to the electricity. Virtual smart meter (C3) will also be implemented to provide similar functionality to a virtual HV/LV transformer (B3), which emulates the load of the LV part. An RPi device or another programmable device can be used to represent this block [14]. Load simulators can be connected to these devices to achieve a more realistic emulation.

Smart charging stations (C4) are becoming more and more common thanks to the electromobility trend. This block will be implemented using the Terra AC wallbox from ABB supporting the OCPP or MQTT communication protocols or the EVBox Elvi device supporting the OCPP, Modbus TCP, or MQTT protocols. Other suitable devices include ChargePoint Home Flex or Ensto EVBASIC.

Industrial Infrastructure (D) represents a simplified factories that incorporate the already implemented BUT industrial polygons: smart brewery (D1) [15], wastewater treatment plant (D2) [15] and robotic arms (D3) [16]. All of these polygons have already been implemented and virtualized in the OpenMUC system. As part of the extension's implementation, a special web interface will be created for monitoring of all devices.

IV. Conclusion

The article presents the design of a distribution system model that will be implemented as an extension to the existing Energy Polygon created at BUT. The purpose of the expansion module is to create a suitable environment for testing new devices and for collecting communication data, including communication data from industrial and energy equipment used in a real environment. Although these data are available, they lack a greater degree of complexity. The wide range of communication protocols used, such as IEC 104, IEC 61850, DLMS, Modbus, OPC, or S7, allows the creation of complex datasets suitable for further processing, for example, using AI. Another possible use is for educating students or employees interested in energy management. The designed polygon can also be used to create operational, service, or cyber scenarios that present common or less common events that can occur within the energy system.

Although the article only proposes an expansion model for the distribution system, it will be implemented in the experimental workplace. Subsequently, the compatibility of the module with the existing infrastructure of the Energy Polygon will be tested, and additional modifications will be considered to improve the overall functionality of the system.

Acknowledgment

This research is a part of the project reg. no. VK01030109 supported by Ministry of the Interior of the Czech Republic in program Open call in security research 2023–2029.

References

[1] K. Gupta, S. Sahoo, B. K. Panigrahi, F. Blaabjerg, and P. Popovski, "On the assessment of cyber risks and attack surfaces in a real-time co-simulation cybersecurity testbed for inverter-based microgrids," *Energies*, vol. 14, no. 16, p. 4941, 2021.

[2] A. Bohačík, "Scénáře polygonu České distribuční sítě," in *Proceedings I of the 27th Conference STUDENT EEICT 2021*, april 2021.

[3] K. Ibtissam, M. S. Abdelrahman, A. Alrashide, and O. A. Mohammed, "Assessment of protection schemes and their security under denial of service attacks," in *2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*. IEEE, 2022, pp. 1–6.

[4] G. Ravikumar, B. Hyder, J. R. Babu, K. Khanna, M. Govindarasu, and M. Parashar, "Cps testbed architectures for wampac using industrial substation and control center platforms and attack-defense evaluation," in *2021 IEEE PESGM*. IEEE, 2021, pp. 1–5.

[5] H. Zhang, D. Ge, J. Liu, and Y. Zhang, "Multifunctional cyber-physical system testbed based on a source-grid combined scheduling control simulation system," *IET Generation, Transmission & Distribution*, 2017.

[6] T. Becejac, C. Eppinger, A. Ashok, U. Agrawal, and J. O'Brien, "Prime: a real-time cyber-physical systems testbed: from wide-area monitoring, protection, and control prototyping to operator training and beyond," *IET Cyber-Physical Systems: Theory & Applications*, 2020.

[7] H. Huang, P. Wlazlo, Z. Mao, A. Sahu, K. Davis, A. Goulart, S. Zonouz, and C. M. Davis, "Cyberattack defense with cyber-physical alert and control logic in industrial controllers," *IEEE Trans. Ind. Appl.*

[8] P. Radoglou-Grammatikis, P. Sarigiannidis, I. Giannoulakis, E. Kafetzakis, and E. Panaousis, "Attacking iec-60870-5-104 scada systems," in *2019 IEEE World Congress on Services (SERVICES)*. IEEE, 2019.

[9] S. Sahoo, "Power and energy management in smart power systems," *Artificial Intelligence-based Smart Power Systems*, pp. 349–375, 2023.

[10] M. R. Maghami and A. G. O. Mutambara, "Challenges associated with hybrid energy systems: An artificial intelligence solution," *Energy Reports*, vol. 9, pp. 924–940, 2023.

[11] S. Upadhyay and A. K. Nayak, "Challenges and remedies of grid-integrated renewable energy resources," *Renewable Energy Systems: Modeling, Optimization and Applications*, pp. 379–402, 2022.

[12] S. Palanisamy, L. Shanmugasundaram, and S. Chenniappan, "Energy storage systems for smart power systems," *Artificial Intelligence-based Smart Power Systems*, pp. 99–114, 2023.

[13] S. Denysiuk, B. Basok, and V. Opryshko, "Optimization of energy processes in local power supply systems with variable operating modes," in *Power Systems Research and Operation: Selected Problems II*, 2022.

[14] D. Kohout and P. Mlýnek, "Testing smart meters with custom application," in *2022 IEEE International Carnahan Conference on Security Technology (ICCST)*, 2022, pp. 1–6.

[15] O. Pospíšil, P. Blažek, K. Kuchař, R. Fujdiak, and J. Mišurec, "Application perspective on cybersecurity testbed for industrial control systems," vol. 21, no. 23, pp. 1–38, december 2021.

[16] E. Holasova and R. Fujdiak, "Deep neural networks for industrial protocol recognition and cipher suite used," *2022 IEEE International Carnahan Conference on Security Technology (ICCST)*, pp. 1–7, 2022-9-7. [Online]. Available: https://ieeexplore.ieee.org/document/9896532/

Workload Tool for Identifying Communication Bottlenecks in Industrial Networks

Eva Holasova* and Jan Pospisil
Dept. of Telecommunications, Faculty of Electrical Engineering and Communications,
Brno University of Technology, Technicka 12, 616 00 Brno, Czech Republic
*eva.holasova@vut.cz

***Abstract*—Industrial networks are used to control and manage critical operations. Because of the need for high reliability and availability of these devices, computational and communication limits must be identified. A large number of available tools are used to flood a selected node without obtaining feedback. In this work, a tool capable of generating a gradually increasing workload and identifying bottlenecks in terms of the amount of communication is presented. From the output of this tool, it is possible to define severity levels with a link to other devices to prevent the device from crashing.**

***Index Terms*—Communication limits, Device reliability, Industrial networks, Network testing**

I. Introduction

To find the communication limits of industrial networks and systems, a systematic approach and an overall knowledge of the critical elements of the network/structure under consideration are required. Conventional Operational Technology (OT) networks are different from Information Technology (IT) networks. This is not only due to the different communication protocols [1], the number of individual devices communicating, but also due to the different requirements for security services. In OT networks, availability is prioritized, followed by integrity and, in third place, confidentiality. The order is thus completely opposite in relation to IT networks, where confidentiality is given the highest priority. This difference is due to the different priorities or objectives of the IT and OT networks. The main objective of the OT network is to supervise and manage a defined process using technology in the industrial and critical infrastructure. Availability is thus often the most critical (delays can cause critical conditions in production). For this reason, it is necessary to look at the security of these networks in terms of the implemented security corresponding to the network, the technologies, and the processes used. Based on the knowledge of the capabilities of the network, it is thus possible to define different levels of criticality (severity levels of data load and also individual nodes, eg. industrial gateway [2]). In the event of an imminent danger (e.g. earthquake), or due to a network anomaly caused by a fault or an attacker, it is necessary to detect the imminent risk as soon as possible and to perform a safe shutdown of important components in order to ensure safety against loss of availability of the monitored component [3]. These states may correspond directly to the measures defined. Thus, in the case of low load, the monitored/controlled process is not affected, but in the case of higher levels, the process is automatic and especially safely shut down in order to prevent greater damage. Thus, the aim of this work is to build up a system that will detect the communication limits of each device under test to obtain a more detailed knowledge of the network. Due to the knowledge of the limits of the network in terms of the overall composition and individual components, this work aims to design a device/system to identify these locations. Based on this knowledge, it is then possible to define severity levels that will correspond to the load of the components. These levels can then be directly linked to response systems that either prevent further degradation, reroute traffic, perform safe shutdowns, etc.

II. Related Works

Current state-of-the-art focuses mainly on penetration testing only from the perspective of causing a denial of service and does not take into account the limits or boundary loads of individual network elements. Thus, it is often not possible to determine from the testing performed what maximum load is possible and on which network elements so that connectivity is not lost. For this reason, the most relevant literature was selected for comparison with the objectives of this work.

Richard Sharpe, et al. presented in their paper [4] evaluation of an Industry 4.0 in relation to the implemented security. In their paper, they have mentioned the most critical parts from the security point of view. The security recommendation has been also mentioned.

Paper [5] created by Irfan Siddavatam, et al. describes the created testbed running over an industrial protocol Modbus. Their paper described Modbus protocol structure and also cover three types of cyber-attack and four related error states.

Paper [6] by Eric Gamess, et al. focuses on Modbus protocol performance evaluation under normal operation and with an attack in progress. They also done selected Modbus tools benchmarking and provided output data visualization.

Panagiotis Radoglou Grammatikis, et al. described in their paper [7] Modbus related cyber-attacks using smod tool. They also presented a network traffic monitoring module to recognize denial of service attacks.

Eric Gamess, et al. presented in paper [8] performance evaluation of denial of service attacks. They also examined the different response times depending on the spacing of the messages sent to trigger denial of service.

Paper [9] by Tongxin Li, et al. focuses on the denial of service detection method in industrial networks based on the digital twin. They also presented a strategy based on Intrusion Detection Systems (IDS) like Snort to detect these types of anomalies. Using this strategy, they described the attack method, while tuning the Snort rule.

Eduardo Vásquez and Helder Castrillón described in their paper [10] a large number of potential attacks on industrial systems. Based on the defined vulnerabilities, they conducted research and identified the most common reason for successful attacks on industrial networks caused by exploiting the default settings of individual elements in the network.

Parul Sindhwad and Faruk Kazi created paper [11], in their paper, they divided attacks targeting industrial networks into three types: physical, logical, and network. The different attacks were then demonstrated on a wastewater treatment testbed.

Ayesha Rahman, et al. demonstrated in their paper [12] denial of service attacks and protection mechanisms in relation to industrial networks. The impact of individual attacks is shown in Factory IO simulation scene.

The Table I compares the above-mentioned articles in terms of the defined criteria. The criterion Increasing levels describes whether the article takes into account the gradual increase in the impact of the test tools (in terms of data generated), also whether the goal is to detect a critical state (limit) or just to cause a denial of service condition. Furthermore, the criterion used is the visualization of the obtained results, the identification of the most critical element (weka point), as well as the industrial protocol used. Another parameter is a correlation with security, i.e. creating a correlation between the obtained information and the used security tools or the created recommendation. The last parameter is the focus on denial of service.

TABLE I
RELATED WORK COMPARISON FROM THE VIEW OF FINDING LIMITS.

Publication	Year	Increasing levels	Finding limits	Visualization	Finding "weak point"	Industrial Protocol	Correlation with security	Denial of Service
[4]	2019	NO	YES*	NO	YES*	-	YES*	NO
[5]	2019	NO	NO	NO	NO	Modbus	YES*	YES
[6]	2020	NO	NO	YES	NO	Modbus	YES*	NO
[7]	2020	NO	NO	NO	YES*	Modbus	YES*	YES
[8]	2020	NO	NO	YES	NO	Modbus	YES*	YES
[9]	2021	NO	YES*	NO	NO	Modbus	YES*	YES
[10]	2021	NO	NO	YES	NO	Modbus	YES*	NO
[11]	2022	NO	NO	NO	NO	Modbus	YES*	YES
[12]	2022	NO	NO	NO	NO	Modbus	YES*	NO

* indicates incomplete coverage of the point.

It is clear from the table that a large number of works focus on causing denial of service in penetration testing without identifying individual network deficiencies. However, from a penetration testing perspective, the current state-of-the-art analysis has shown that there is no identified threshold at which denial of service or other negative impacts within the network occur. Therefore, the goal is to develop a comprehensive tool for testing an industrial network demonstrated on the industrial Modbus protocol.

III. INDUSTRIAL NETWORKS

The needs of industrial control systems (ICS) and automation applications are supported by industrial networks, which are specialized communication networks. These networks connect a variety of equipment and systems, such as sensors, actuators, controllers, and other pieces of industrial devices, and they give this equipment and systems a way to interact and exchange data. The specific needs of ICS applications, which frequently involve real-time control and monitoring of crucial systems and processes, must be taken into account while designing industrial networks. High reliability, secure, and able to sustain high-speed data transfer rates networks is necessary for this. Industrial networks come in a wide variety of forms, each with special capabilities and traits. Ethernet/IP, Modbus TCP, PROFINET, and DeviceNet are some of the industrial protocols that are most frequently used.

In contemporary industrial settings, the usage of industrial networks has grown in significance as businesses look to increase productivity, cut costs, and improve the security and dependability of their operations. However because industrial networks are frequently connected to the internet and other external networks, they also present new security problems because they are open to dangers like cyberattacks and other hazards. Organizations must thus take action to safeguard their industrial networks, including putting in place network segmentation, access controls, and other security measures.

Purdue model: The Purdue model is a well-known and frequently used reference model for the architecture of ICS. Many organizations and industries have since used it as a framework for organizing and controlling ICS infrastructure. This model provides an accessible and organized framework for comprehending and controlling ICS infrastructure, and it may be used to identify potential security risks and vulnerabilities at various system levels. It has been developed for application in a number of different fields, including cybersecurity and information technology, and is widely used in the industrial control and automation industries.

The Purdue model divides the ICS infrastructure into a series of hierarchical levels, where each level corresponds to a distinct aspect of the system's functionality and management. Field devices, such as sensors and actuators, are at the lowest level and are in charge of data collection and fundamental control operations. The control networks and devices, which perform more complex control operations and provide data to higher levels of the system, are located above this. The distributed control systems (DCS) and supervisory control and data acquisition (SCADA) systems, which offer real-time

control and monitoring of the ICS, are represented by the middle level of the Purdue model. These systems are in charge of controlling the transmission of information and orders between field equipment and the system's higher levels. The enterprise level, which includes the business and management systems that connect with the ICS, is at the top of the Purdue model. This level is in charge of supervising the ICS's overall operation and making choices using the data and knowledge provided by the system's lower levels.

IV. Modbus Protocol

In industrial networks, Modbus is a communication protocol that is frequently used. It was developed as a way of data transfer between network devices and programmable logic controllers (PLCs). The Modbus protocol is a client-server protocol, a client device initiates communication, and one or more server devices reply to the client's commands. A server device receives a request from a client device and responds with the requested data. The information is typically presented as binary values or analog signals, such as temperature or pressure readings. The Modbus protocol comes in two different versions: Serial Modbus and Modbus over Ethernet. The Modbus protocol defines two serial transmission modes, Modbus RTU and Modbus ASCII. The mode determines in which format the data is transmitted. Each unit must support RTU mode, ASCII mode is optional. All units on a bus must operate in the same transmit mode and communication is ensured by using an RS-232 or RS-485 serial interface. Modbus/TCP is an Ethernet-based protocol that uses Transmission Control Protocol/Internet Protocol (TCP/IP) for network communication.

Modbus RTU: Modbus RTU is a simple binary protocol that uses serial communication. This protocol primarily uses an RS-232 or RS-485 serial interface for communications. Modbus RTU messages are a 16-bit structure with a Cyclic-Redundant Checksum (CRC). The simplicity of these messages ensures reliability. Because of its simplicity, the fundamental 16-bit Modbus RTU register structure can be used to store unrelated data such as queues, floating points, tables, and ASCII text. The frame or packet of this protocol includes a server device address, function code, special data depending on the function code, and CRC. A checksum is a component of error checking that is used to guarantee the accuracy of the data. Fig. 1 shows the structure of the Modbus RTU message.

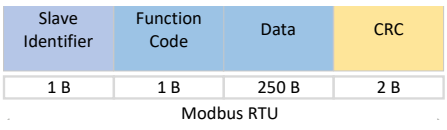


Fig. 1. Modbus RTU protocol structure.

Modbus/TCP: Modbus/TCP is a more complex protocol that is designed for use in modern industrial automation systems. Modbus TCP/IP is typically used in larger systems that require multiple devices to communicate with each other. It uses the same format of packets as Modbus RTU, but the data is sent via Ethernet using TCP/IP. For this purpose the packet contains a Transaction Identifier to uniquely identify each request, Protocol Identifier will always be 00 00, which corresponds to the Modbus protocol, length identifies the number of bytes in the message, a unit identifier to uniquely identify the Slave device, function code, the data field, and the CRC field is not here. As a result, it is quicker and more dependable than Modbus RTU and enables communication over greater distances. Fig. 2 shows the structure of the Modbus/TCP message.

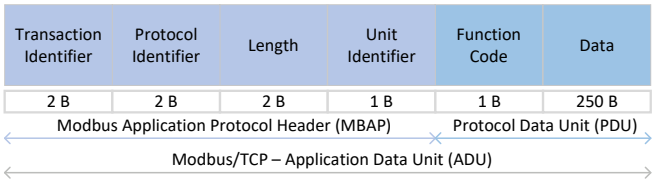


Fig. 2. Modbus TCP protocol structure.

Modbus/TCP Security: Modbus/TCP Security protocol is a security variant of the Modbus/TCP protocol utilizing Transport Layer Security (TLS). Security using TLS is provided by cipher suites. Cipher suites are settings for encryption schemes. An example of a cipher suite is TLS_AES_128_GCM_SHA256, which involves the use of an Advanced Encryption Standard cipher with a key size of 128 bits in Galois/Counter mode and uses a Secure Hash Algorithm with the size of 256 bits for the hash function. The use of cipher suites depends on the capabilities of the devices and is agreed upon when establishing the connection between the devices.

Modbus function codes: Modbus protocol divides data types into four specific structures - Coil (1-bit, read/write), Discrete Inputs (1-bit, read-only), Input Registers (16-bits, read-only), Holding Registers (16-bit, read/write). For this purpose, Modbus defines a set of function codes that are used to perform different types of operations. Each function code is represented by a unique byte value in the Modbus message. The size of each function code in Modbus protocol is 1 byte, which allows for efficient data transfer between Modbus devices. However, due to the specific operation being carried out, some function codes may include extra bytes in their request or response messages. Table II shows the most commonly used function codes.

TABLE II
Most commonly used Modbus function codes.

Name	#	Size	Note
Read Coils	01	1 B	read the status of discrete coils
Read Discrete Inputs	02	1 B	read the status of discrete inputs
Read Holding Registers	03	1 B	read the values of holding registers
Read Input Registers	04	1 B	read the values of input registers
Write Single Coil	05	1 B	write a single discrete coil
Write Single Registers	06	1 B	write a single holding register
Write Multiple Coils	15	1 B	write multiple discrete coils
Write Multiple Registers	16	1 B	write multiple holding registers

V. Industrial Evaluation Testing Tool Design

Based on the analysis, the available tools for testing the Modbus protocol were selected as Shodan, Modbus Doctor, ModbusPal, QModMaster, and smod. Each of these tools provides different outputs, but none of them is able to perform network evaluation and identify critical points at which a crash occurred or identify delays at a particular load. For this reason, a new tool was proposed. The basic objectives of the tool include the ability to identify stations in the network, determine communication protocols and ports, the ability to perform the extraction of communication memory blocks (coils and registers in the case of the Modbus), and perform a test write to individual memory blocks. The key requirement for the tool is considered to be the increasing load on the selected station in terms of data messages sent to the selected station. By gradually increasing the load in terms of long-term utilization or data traffic bursts, the communication capabilities of the device can be determined. For the relevant data obtained, it is useful to correlate the data obtained from the data traffic measurements (amount, size, and type of messages sent) can be used with the data obtained from the device under test (e.g. memory and processor usage).

The block diagram of the tool reflecting the above-mentioned requirements presented on the Modbus protocol is shown in Fig. 3. The proposed tool first identifies the stations, scans these stations in terms of open ports, and selects the appropriate slave station identifier. Subsequently, a block of reading data from the device (if selected) is performed, first extracting data from the device and then gradually increasing the load in terms of generated data on the target device. This element can be directly defined or increased until the device under test crashes or the computational capacity of the device under test is maximized. The individual information, which can include the number of rejected messages, the number of messages waiting in the queue, and the average time between request and response are logged. This point is critical for the correct evaluation of the acquired data. It is not necessary that the device under test crashes, but from the point of view of maximum response time to a query it is possible to define a critical load level for which certain parameters are met. In the next section, the identical process is performed with the write operation. The unique information is again stored, and the transfer parameters are evaluated and presented graphically to the user.

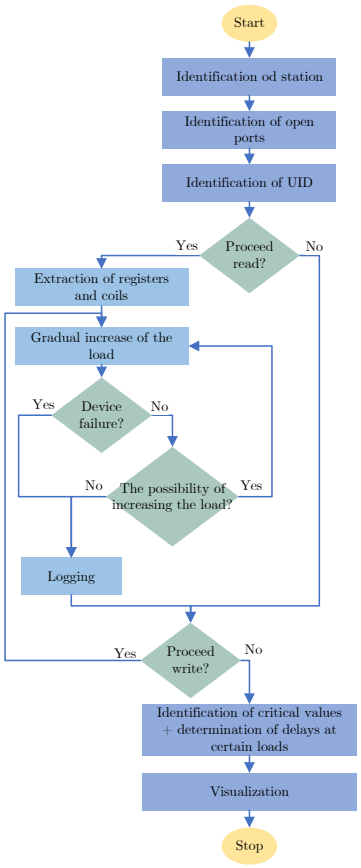


Fig. 3. Block diagram of the proposed tool used to identify the computational and communication capacities of the device under test.

Experimental Usage of the Proposed Tool

The proposed tool has been already implemented (read, write, and evaluation blocks) using the Python programming language. Parallelization was used to exploit the impact testing with large amounts of data. The operations setCoilValue, setCoilValues, getCoilValue, getRegisterValue, setRegisterValue, and getInputRegisters are implemented. In the experimental scenarios, the testing device was implemented using the PyModbus library in an Ubuntu 20.04 virtualized machine where the memory size was set to 2 GB and 1 AMD Ryzen 5 2600, 3.4 GHz processor was used. The results of the first scenario are shown in Fig. 4. The data generator sequentially generated six separate clusters of data where approximately 400 packets per second were generated during the cluster. In total, 36 000 packets were generated during the test (i.e., 6000 packets per function code). The average packet size was 83 B, and the average rate was 0.78 packets per second. CPU utilization was negligible during the test. In comparison to the first scenario, where the individual operations were run sequentially, the second scenario performed a simultaneous commands execution. The flow of the second scenario is shown in Fig. 5. The total of 60,000 packets were generated, with an average size of 83 B and an average rate of 1.49 packets per second. However, during the second scenario, the device crashed, causing a communication interruption. The figure shows the increase in packets sent in blue and the increase in CPU utilization of the device under test in red. Approximately 8 seconds after the test started, the CPU utilization increased to 90 %. This condition persisted for approximately 30 s and then a crash occurred. Experimental results showed the ability of the presented tool (although not in the final version) to be functional and usable in practice. From the testing performed, it is possible to determine the utilization of 800 packets per second as a critical value requiring immediate activation of failsafe countermeasures to prevent denial of service.

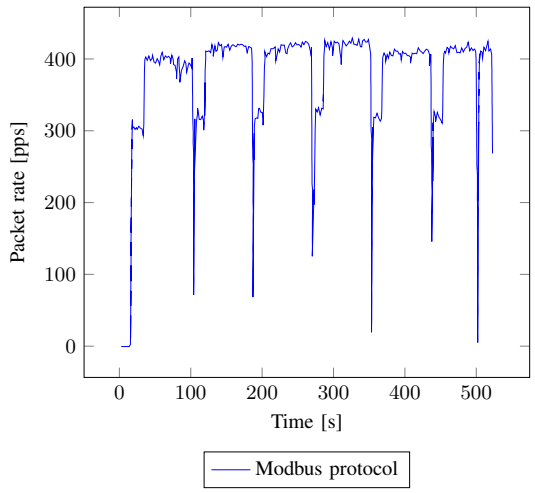


Fig. 4. Use of communication line in sequential testing – first scenario.

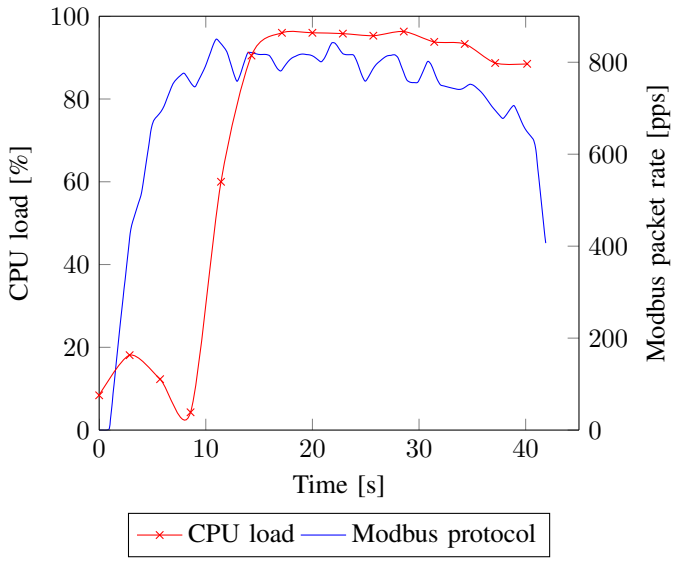


Fig. 5. Communication link utilization and CPU utilization during stress testing – second scenario.

VI. Conclusion

The aim of this work was to present a tool for identifying critical points (communication bottlenecks) in terms of network load within the industrial network infrastructure. The purpose of the tool is to provide feedback in terms of computational performance and response at certain communication load levels. The predefined tool was demonstrated using the Modbus industrial protocols. The basic requirements for the resulting tool were established and the main blocks of the tool were proposed. The tool was then partially implemented and the results were presented. The results of the tool were divided into two test scenarios, where the first scenario performed step-by-step testing of the implemented operations. The second scenario performed a 60 000 packets impact overload of the system under test, where the system subsequently crashed after 30 s of overload. The CPU utilization of the device under test was also monitored during the testing. Thus, the presented tool demonstrated the ability to identify the critical point. Based on the output of the program, the operation can be further handled to prevent critical conditions with subsequent denial of service.

The output of the program can be linked to other security mechanisms in order to divert traffic to ensure availability for critical operations. The tool also allows testing detection mechanisms in the network.

Acknowledgment

The presented research is a part of the project reg. no. FW06010490, financially supported by the Technology agency of the Czech Republic under the 6. TREND program.

References

[1] S. Jaloudi, "Communication protocols of an industrial internet of things environment," *Future Internet*, vol. 11, no. 3, 2019. [Online]. Available: https://www.mdpi.com/1999-5903/11/3/66

[2] V. G. Găitan and I. Zagan, "Experimental implementation and performance evaluation of an iot access gateway for the modbus extension," *Sensors*, vol. 21, no. 1, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/1/246

[3] I. Mugarza, J. Parra, and E. Jacob, "Cetratus," *2018 IEEE 13th International Symposium on Industrial Embedded Systems (SIES)*, pp. 1–8, 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8442088/

[4] R. Sharpe, K. van Lopik, A. Neal, P. Goodall, P. P. Conway, and A. A. West, "An industrial evaluation of an industry 4.0 reference architecture demonstrating the need for the inclusion of security and human components," *Computers in Industry*, vol. 108, pp. 37–44, 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0166361518304950

[5] I. A. Siddavatam, S. Parekh, T. Shah, and F. Kazi, "Testing and validation of modbus/tcp protocol for secure scada communication in cps using formal methods," *Scalable Computing: Practice and Experience*, vol. 18, no. 4, pp. 313–330, 2017-11-24. [Online]. Available: http://www.scpe.org/index.php/scpe/article/view/1331

[6] S. Jaloudi, "Communication protocols of an industrial internet of things environment," *Future Internet*, vol. 11, no. 3, 2019. [Online]. Available: https://www.mdpi.com/1999-5903/11/3/66

[7] P. Radoglou-Grammatikis, I. Siniosoglou, T. Liatifis, and A. Kourouniadis, *2020 9th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, 2020.

[8] E. Gamess, B. Smith, and G. F. III, "Performance evaluation of modbus tcp in normal operation and under a distributed denial of service attack," vol. 12, no. 2, pp. 1–21, 2020-03-31. [Online]. Available: http://aircconline.com/ijcnc/V12N2/12220cnc01.pdf

[9] T. Li, Y. Wang, C. Zou, Y. Tian, L. Zhou, and Y. Zhu, "Research on dos attack detection method of modbus tcp in openplc," *Journal of Computer and Communications*, vol. 09, no. 07, pp. 73–90, 2021. [Online]. Available: https://www.scirp.org/journal/doi.aspx?doi=10.4236/jcc.2021.97007

[10] E. Vásquez and H. Castrillón, "Scientific literature review for the search of vulnerabilities and pentesting tools for scada system," *Telematics and Computing*, pp. 90–103, 2021. [Online]. Available: https://link.springer.com/10.1007/978-3-030-89586-0_7

[11] P. Sindhwad and F. Kazi, "Exploiting control device vulnerabilities," *2022 32nd Conference of Open Innovations Association (FRUCT)*, pp. 270–279, 2022-11-9. [Online]. Available: https://ieeexplore.ieee.org/document/9953826/

[12] A. Rahman, G. Mustafa, A. Q. Khan, M. Abid, and M. H. Durad, "Launch of denial of service attacks on the modbus/tcp protocol and development of its protection mechanisms," *Int. J. Crit. Infrastruct. Prot.*, vol. 39, no. C, dec 2022. [Online]. Available: https://doi.org/10.1016/j.ijcip.2022.100568

Towards a modular multi-platform ROS2-based navigation system for mobile robots

1st Tomas Horelican
Department of Control and Instrumentation, Faculty of Electrical Engineering and Communication
Brno University of Technology
Brno, Czech Republic
0000-0003-4067-3296

***Abstract*—This paper provides an overview of a proof-of-concept modular system architecture built on top of the *ROS2* framework. As support builds up, the whole system should easily be adaptable to different robotic platform types with little to no modifications necessary. High-level path planning and autonomous navigation tasks make use of the open-source *Navigation2* package, which provides implementations of several well known algorithms and can be further configured to suite even more specific needs. Low-level robot and hardware control tasks are performed by custom developed packages which strive to perpetuate the modularity paradigm as much as possible. A description of the individual elements will be provided, and several experimental scenarios will be presented showcasing successful deployment of the system on two conceptually different robotic platform types.**

***Index Terms*—mobile robot, autonomous navigation, obstacle avoidance, localization, SLAM mapping**

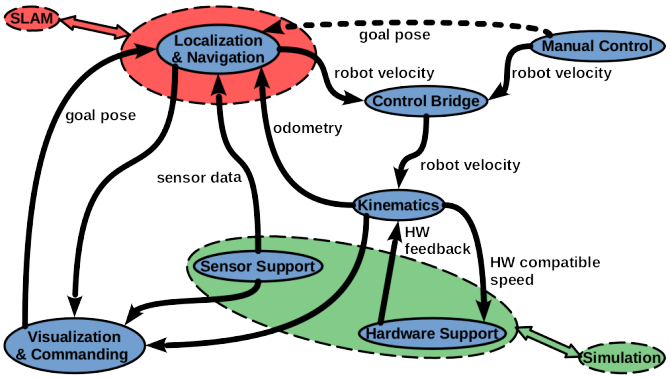


Fig. 1: System architecture and interconnection.

I. Introduction

Mobile robotic platforms can serve as highly convenient tools for exploration or mapping of unknown dangerous or otherwise humanly inaccessible environments. Depending on the platform and its equipment execution, these robots can also be sent out to perform various interactions with its surroundings. In order to execute any of these more complex tasks, the robot must first be able to effectively navigate and localize itself within the operating area. A lot of the time, this deficiency can be circumvented with a reliable remote control or telepresence solution, where the human operator is able to compensate for the high-level operations, as was proposed in [1]. However, in scenarios where long time delays eliminate any possibility of real-time manual control or the use of human operators becomes inefficient, a certain level of autonomy must be established instead. The work presented in [2] strives to further advance this notion by automating many of the procedures during a radiation contaminated environment survey.

Nevertheless, the intrinsic capabilities of the individual robotic platforms still desire for more environmentally-aware navigation solutions. There are already many well known and reliable algorithms that can be deployed onto a robot to help it solve the autonomy-related tasks. Their practical implementation onto a real-world robot often requires a laborious design and integration process to fit onto the specific platform. After that, the solution might not be easily reusable and porting it to a different kind of robot might as well lead to a complete redesign.

The *Robot Operating System (ROS)* framework introduced a consistent development workspace structure, within which different modules (called *nodes*) can be implemented and easily interconnected into a desired architecture. Each participating endpoint must abide to a well defined message interface, through which it interacts with all of the other participants. This enabled rapid integration of available, and continuously maintained, open-source packages together with custom designed *nodes* into a highly modular and reusable system. One example of such an open-source package is the *Navigation Stack*, which harbors a wide range of useful navigation and mapping tools utilizing many of the state-of-the-art algorithm implementations. The next generation *ROS2* framework builds upon the previous concepts with an emphasis on better reliability, performance and safety, as described by the authors in [3]. Following with the *Navigation2* package by [4].

This article will bring forth a follow-up on previously presented work in [5], supplementing it with the latest advancements. More focus will also be placed on other lower-level parts of the system that have not been described previously.

II. System Design and Structure

As displayed in Fig. 1, the whole system can be broken down hierarchically into several separate layers:

1) *hardware support*
2) *kinematics*
3) *sensor support*
4) *motion control bridge*
5) *autonomous localization and navigation*
6) *manual control*

All of these constituent elements are always running directly on a robotic platform (in case of a physical robot) and exchange data between each other through the standardized *ROS2 message* interface. In case of a simulated scenario, the *hardware* and *sensor support* layers are essentially replaced by the used *ROS2*-compatible simulation environment such as *Gazebo*, with the rest remaining equal. Other types of supporting or directly replacing layers (ex. *simulation* and *SLAM mapping*) can be connected to the whole system as well, such as:

1) *RViz visualization and commanding*
2) *basic simulation*
3) *Simultaneous Localization and Mapping (SLAM)*

A. Individual Component Description

This section will elaborate on each of the layers enumerated in the beginning of the chapter.

a) Manual control: This layer is used to directly interact with a gamepad interface and translate its inputs to linear and angular velocity motion commands in the robot frame for the *motion control bridge* layer, to enable taking manual control of the robot.

b) Autonomous localization and navigation: This performs all of the path planning and obstacle evasion tasks using output from the *sensor support* layer and desired goal pose inputs from a *commanding* layer. Direct velocity motion commands in the robot frame are sent to the *motion control bridge* layer. This section is heavily utilizing the *robot_localization* [6] and *Navigation2* package tool-sets, all of which has already been described more closely in [5], together with the specific algorithms and configuration used. A main addition is that the *Adaptive Monte Carlo Localization (AMCL)* and *SLAM Toolbox* packages are now fully utilized as well.

c) Motion control bridge: This is just a simple multiplexing interface that takes the velocity motion inputs from higher levels and chooses which will be passed down to the lower level.

d) Sensor support: This is responsible for connecting the robot sensor base to the rest of the system (ex. the *localization and navigation* layer). A more detailed description will be left for a different moment.

e) Kinematics: This layer receives the standardized velocity motion commands in the robot frame, passed through the *motion control bridge* layer, and converts them down to the desired control commands for the *hardware support* layer, depending on the utilized kinematics model type and robotic platform architecture design. At the same time, feedback from the *hardware support* layer is being used to provide basic odometry pose estimations.

All of the layers up until now were able to, in some way or another, interact using the same standardized message type being passed down. However, from this point, the messages will have to start diversifying, as different kinematic models may produce different outputs and simultaneously require different forms of feedback. This is being further complicated by the various possible robot hardware combinations requiring different interaction approaches a well.

For example, differential and omnidirectional motion models will differ in the calculations they execute. A two-wheeled robotic platform will necessitate different interactions than a three or four-wheeled one (which by itself might be differential or omnidirectional). The specific hardware control device on the platform might expect individual commands for each wheel, or already perform some of the calculations and require just one centralized type of data input. These are only some particular examples of complications that should be alleviated by a carefully designed interplay between the *hardware support* and *kinematics* layers in order to offer a frictionless integration process.

This can still be managed as, conceptually, a single *node*, within which the actual calculations may vary. To accomplish that, the *ROS2 plugin* interface is utilized. Each new kinematics (or hardware specialized) model can be implemented as a *plugin* that can be dynamically loaded during the *node* startup. This concept was in great detail explained by [7], were they utilized the *plugin* interface for various types of costmap operations, which depend on the source of incoming data. The authors of [8] presented a way to generalize the different kinematic calculations into a single consistent ruleset. Following this, to provide odometry pose estimates, the base *node* will always operate with incremental translational and rotational changes of the robot center frame (in standard units). It is then the job of the particular *plugin* to extract and unite this data from the *hardware support* layer in any applicable way. Similarly, when the *node* receives a standardized motion control command, the *plugin* will be responsible for translating that message into the *hardware support* layer compatible message structure (still using standard units). From there, it will be further translated into the specific hardware commands.

f) Hardware support: This is the lowest level and it mediates access for the various hardware elements that the robot platform is composed of. It is going to be the most specialized one since it has to account for a wide range of devices, protocols, and their combinations, and be able to interact at many different levels. Support for new devices will have to continuously be added and maintained as more of them are exposed to the system. It is responsible for translating messages from the *kinematics* layer directly into the controllers or actuators (accounting for gear-ratios, step increment resolutions, etc.), and also providing direct feedback up into the *kinematics* and/or *visualization* layers, essentially creating a generalized abstraction over the hardware components.

At present, this layer is designed as a monolithic library composed of several different tools for communications or

hardware control, which can be interconnected into any number of *nodes* based on the utilized platform, its design and requirements. The outermost endpoint directly interacting with the *kinematics* layer then translates and presents all the requested data in a unified format defined by the used kinematics model *plugin*. The tools in this library are not all strictly bound to just robot motion control tasks, but should assemble abstractions for any necessary hardware interaction.

III. Experiment Results

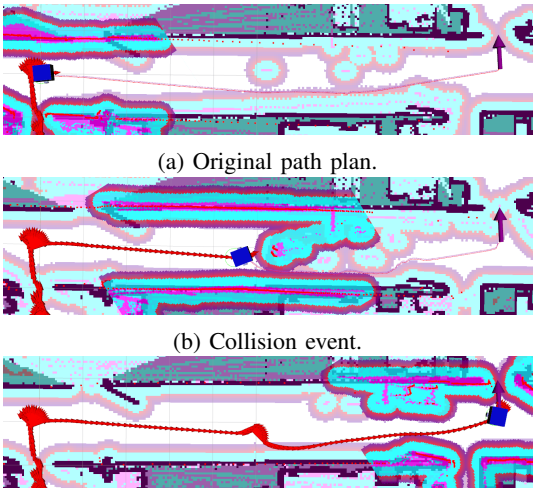
The system has already been deployed onto different robots, which were in some specific form also successfully used as part of a European grant project deliverable. This section will therefore serve as a more general proof-of-concept demonstration as it relates to the architecture design presented in chapter II, and demonstrate the deployability onto various robot types.

Two different robots were used for the experiments in this section. Specific details of the physical platform designs used are not relevant for the purposes of this article. The most important characteristic that has to be differentiated is the robot chassis type, which was:

1) a two-wheeled differential drive,
2) and a four-wheeled omnidirectional drive.

The second important attribute is the used sensor equipment, which can also vary platform-to-platform and, therefore, it closely relates to the feasibility of the integration process. It is important to emphasize that the purpose of this article was not to design any specific sensor system suitable for a given navigation task (that is left for a different occasion), but rather showcase a compound navigation system, which can with as little effort as possible integrate an already existing sensor base present on the robot. Of course, some basic realistic constraints have to be respected on both sides of the equation, where even the sensor system must first be capable of providing enough environmental data for the navigation task to be possible. Environmental feedback for obstacle avoidance and/or localization tasks of the two robots used in this article was provided using a single *Velodyne HDL32E* lidar sensor on the differential chassis, and a set of 6 *Livox MID70* lidar sensors on the omnidirectional chassis robot. The ground footprint diameters encompassing the two platforms are 50 cm and 64 cm, for the differential and omnidirectional robot respectively. Both of these robots were using the same map and conceptually the same high-level algorithms were utilized for navigation.

At the time of writing of this paper, only the differential platform was able to fully utilize its lidar for global map localization using an *AMCL* algorithm. The omnidirectional robot had to rely solely on its odometry calculations, which is visible from the worsened localization performance. A more reliable navigation solution, fusing a higher modality of sensor data can as well be implemented. However, the task of designing a navigation algorithm for any given platform is, once again, out of the scope of this paper and is reserved for another time.

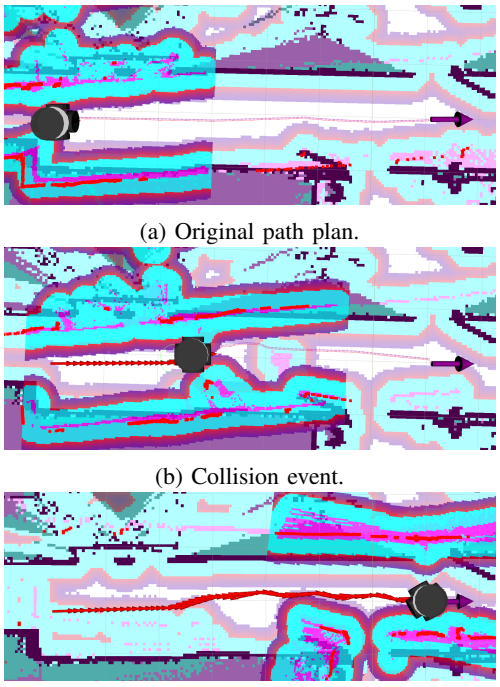


(a) Original path plan.

(b) Collision event.

(c) Goal reached.

Fig. 2: Differential drive robot evading an incoming obstacle in a narrow hall.



(a) Original path plan.

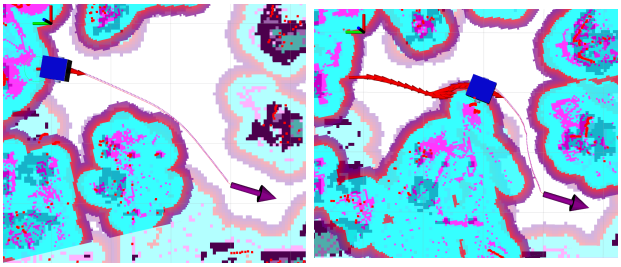
(b) Collision event.

(c) Goal reached.

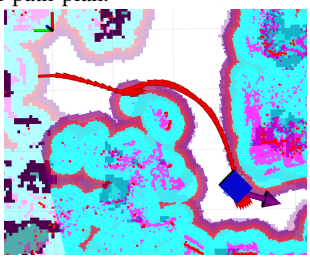
Fig. 3: Omnidirectional drive robot evading an incoming obstacle in a narrow hall.

Figures 2 and 3 comparatively show the behavior of both robots when faced with an incoming & passing by human obstacle.

Similarly, figures 4 and 5 display how the robots were able to adapt their planning and account for an extending obstacle area. In figures 3c and 5c, the position drift of the omnidirectional robot after performing evasion maneuvers is clearly visible. Nonetheless, both platforms were able to reach

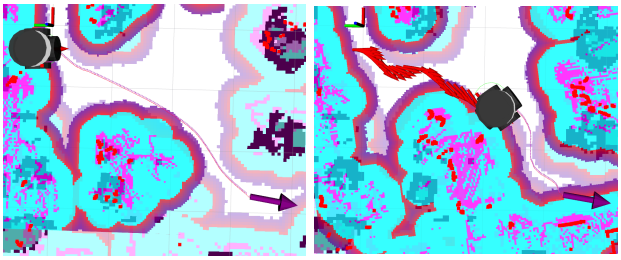


(a) Original path plan. (b) Collision event.

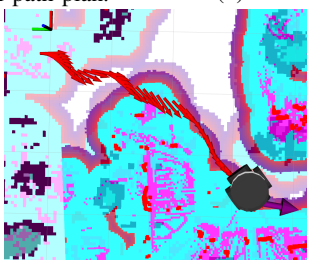


(c) Goal reached.

Fig. 4: Differential drive robot evading an extending obstacle in a wide room.



(a) Original path plan. (b) Collision event.

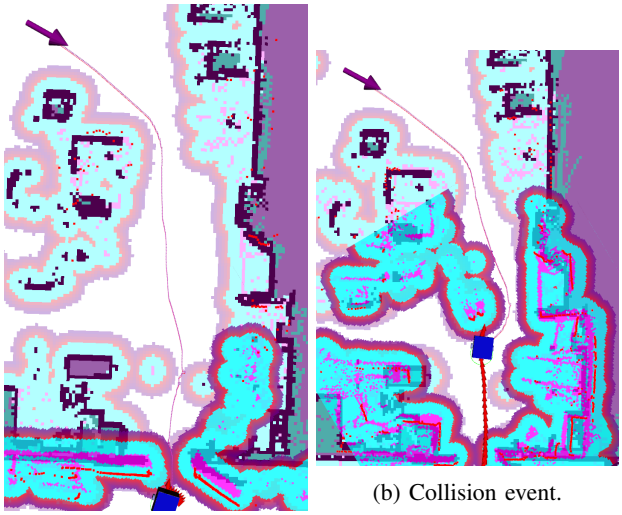


(c) Goal reached.

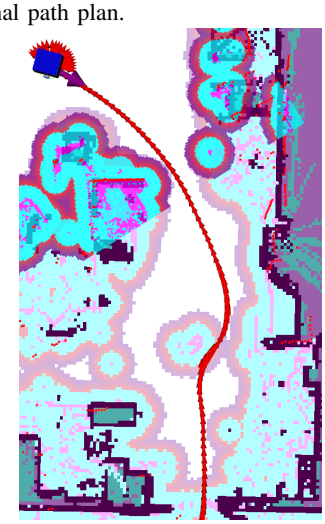
Fig. 5: Omnidirectional drive robot evading an extending obstacle in a wide room.

the desired goal utilizing their available configuration.

Since the differential platform does currently possess a more robust long lasting localization capability, more complex experiments further illustrating this were performed. Figure 7 shows the full travel history as the robot was gradually being sent to explore a floor with 3 rooms. The whole round trip was not completed in one continuous motion, but rather the robot was being consecutively sent to different goal positions throughout the map (as evidenced by some of the edge points on the path). Figure 7b shows exploration through a clear



(a) Original path plan. (b) Collision event.



(c) Goal reached.

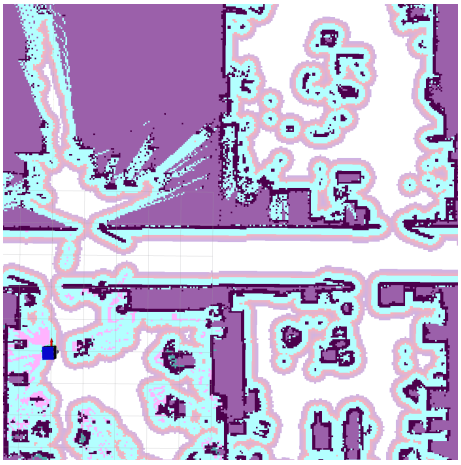
Fig. 6: Differential drive robot evading a drop-in obstacle in a wide room.

map without additional obstacles, while Figure 7c shows exploration with some unanticipated obstacles encountered along the way. Figure 6 provides a closer inspection of the obstacle situation from the top room. The evasion maneuver presented earlier in Figure 2 is actually also the hall section of 7c.

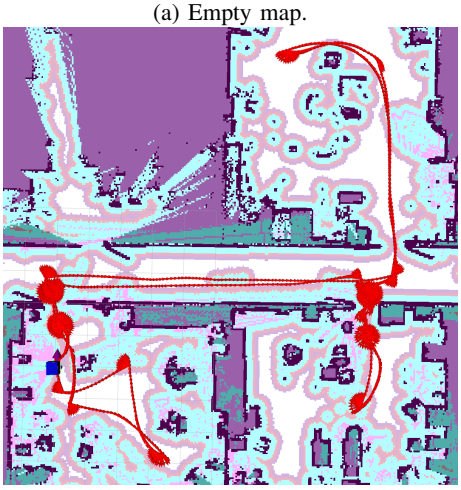
Some problems when attempting to cross doorways can be seen in the rotation-in-place motions, however the robot was always able to recover and pass through. This behavior could be eliminated further by decreasing the inflation around doorway edges, which would provide more confidence for the path planning algorithm.

IV. Conclusion

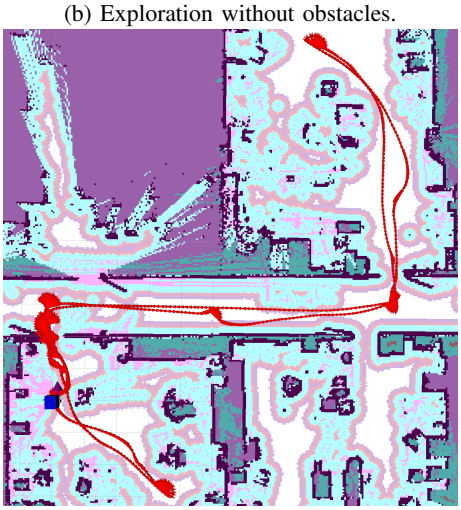
As was shown in chapter III, the system is deployable and currently fully functional on two different types of mobile robot platforms. Both of the robots are able to autonomously



(a) Empty map.



(b) Exploration without obstacles.



(c) Exploration with encountered obstacles.

Fig. 7: Differential drive robot exploring a 3-room map.

navigate within a global static map generated beforehand through *SLAM*, and are able to avoid obstacles in real time. Chapter II described the currently proposed structure. Its modular nature opens up a way to perform rapid prototyping of different algorithms that can be quickly tested on a given robot. It also provides a means to design completely new robotic platforms that can be quickly brought to life using the established concepts. It is clear that (if available) the usage of already distributed and proven solutions should be encouraged preferentially. However, if needed, new customized *nodes* with more modified algorithm implementations can be created and integrated into almost any level of the system. The navigation and environmental awareness capabilities of the system at present are quite satisfactory, however there is still a large window open for improvement.

For one, a more complete and in-depth understanding of all of the tools contained within the *Navigation2* package itself is still necessary in order to use the more complex algorithms efficiently, implement new algorithms that build upon the state-of-the-art, and to enable support for a broader range of sensor accessories. Secondly, the available *SLAM* algorithms are currently only used for offline manual global map creation. This already greatly accelerates the robot deployment, however it should be more closely integrated into the autonomous navigation process itself to enable online global map expansion in real time instead. This is especially relevant for previously unknown or open-spaced environments, which in principle can not utilize a complete global map. Thirdly, the system has not yet been deployed in more complex outdoor environments nor is it fully equipped to do so, which should also be of a higher focus for future work.

Finally, the research presented in this paper is so far only accounting for traditional wheel or belt driven robots. Quadrupedal robotic platforms are becoming an ever more popular tool for remote exploration, however a robust autonomous control and navigation solution respecting difficult terrain environments for these robots is much harder to achieve than for the more traditional platforms. New research utilizing latest developments in *Reinforcement Learning (RL)* algorithms are only now beginning to emerge, which opens up a wider range of possible research directions and enhancements to the current system, platform-and-algorithm-wise.

ACKNOWLEDGMENT

The completion of this paper was made possible by the grant No. FEKT-S-23-8451 - ”Research on advanced methods and technologies in cybernetics, robotics, artificial intelligence, automation and measurement” financially supported by the Internal science fund of Brno University of Technology. The work was supported by the infrastructure of RICAIP that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 857306 and from Ministry of Education, Youth and Sports under OP RDE grant agreement No CZ.02.1.01/0.0/0.0/17_043/0010085.

REFERENCES

[1] F. Burian, L. Zalud, P. Kocmanova, T. Jilek, and L. Kopecny, "Multi-robot system for disaster area exploration," in *WIT Transactions on Ecology and the Environment*, vol. 184. WIT Press, 2014, pp. 263–274.

[2] P. Gabrlik, T. Lazna, T. Jilek, P. Sladek, and L. Zalud, "An automated heterogeneous robotic system for radiation surveys: Design and field testing," *Journal of Field Robotics*, vol. 38, no. 5, pp. 657–683, 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.22010

[3] S. Macenski, T. Foote, B. Gerkey, C. Lalancette, and W. Woodall, "Robot operating system 2: Design, architecture, and uses in the wild," *Science Robotics*, vol. 7, no. 66, p. eabm6074, 2022. [Online]. Available: https://www.science.org/doi/abs/10.1126/scirobotics.abm6074

[4] S. Macenski, F. Martin, R. White, and J. Ginés Clavero, "The marathon 2: A navigation system," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020.

[5] T. Horelican, "Utilizability of navigation2/ros2 in highly automated and distributed multi-robotic systems for industrial facilities," *IFAC-PapersOnLine*, vol. 55, no. 4, pp. 109–114, 2022, 17th IFAC Conference on Programmable Devices and Embedded Systems PDES 2022 — Sarajevo, Bosnia and Herzegovina, 17-19 May 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896322003330

[6] T. Moore and D. Stouch, "A generalized extended kalman filter implementation for the robot operating system," in *Proceedings of the 13th International Conference on Intelligent Autonomous Systems (IAS-13)*. Springer, July 2014.

[7] D. V. Lu, D. Hershberger, and W. D. Smart, "Layered costmaps for context-sensitive navigation," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 709–715.

[8] A. A. Furtuna, D. J. Balkcom, H. Chitsaz, and P. Kavathekar, "Generalizing the dubins and reeds-shepp cars: Fastest paths for bounded-velocity mobile robots," in *2008 IEEE International Conference on Robotics and Automation*, 2008, pp. 2533–2539.

Device behavior analysis based on OPC UA

1st Ondrej Pospisil
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
xpospi89@vut.cz

2nd Radek Mozny
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
radek.mozny@vut.cz
Unit of Electrical Engineering
Tampere University
Tampere, Finland
radek.mozny@tuni.fi

***Abstract*—Device identification in industrial networks is a significant topic as there is a need to identify devices based on passive scanning. Machine learning (ML) techniques can prove useful in this approach. This paper discusses the applicability of different data sources for device identification and focuses on timing information from network communication, specifically from the open platform communication unified architecture (OPC UA) protocol. The aim was to explore the potential of device identification based on this data. In this study, Random Forest and XGBoost ML algorithms were utilized for device identification. In the case types of devices known in the learning phase, the proposed algorithm achieved an accuracy score of up to 0.99. However, for unknown device types, the accuracy score was only around 0.71. The findings suggest that using time information for device identification is a promising direction for further development.**

***Index Terms*—device behavior analysis, machine learning, opcua, random forest, xgboost**

I. Introduction

As the industrial market expands and operational technologies (OT) integrate with information technology (IT), the scope of industrial networks also widens. This convergence of OT and IT in industrial networks increases the risk of cyber attacks, making cybersecurity a critical issue in ICS industrial control systems (ICS). According to the CAGR metric, the global market for this sector is expected to grow at a compound annual growth rate (CAGR) of 7.2% during the period to 2027 [1]. Device behavior analysis is a powerful tool for enhancing cybersecurity in industrial control systems. By detecting anomalies within specific devices, this approach enables better identification of devices and helps to recognize potential vulnerabilities of this device. In this regard, the recently favored method utilized for identifying device modes is the machine learning (ML) approach, which shows promise in improving the accuracy of device behavior analysis and more reliable identification.

The concept of the proposed ML-aided approach is that each device can be identified by a unique fingerprint, similar to that of a person. This fingerprint can be extracted and stored from various types of data. The approach involves extracting information about a device's specific identity from two distinct categories. The first category comprises external data sources, such as network communication, clock skew, and electromagnetic signals. The second category includes in-device data sources, such as hardware events, information from system processors and oscillators, and data on resource usage (such as CPU). Additionally, software information, such as software signature, can also serve as a data source.

Identifying industrial facilities poses several challenges. The primary challenge is to devise a device identification approach that is environment-independent, allowing the method to be applied across various devices and applications. It is also critical to avoid disruption of the device operation during the identification process, necessitating the use of passive data acquisition methods. Moreover, active scanning could interfere with the process and damage the entire system, so precautionary measures must be taken. Another challenge brings the identification of suitable data sources, given the diversity of communication methods and protocols used by industrial equipment. Additionally, data sources that can be generalized across various scenarios should be identified. Limited data availability is another significant challenge, with industrial facilities often generating insufficient data to train ML models accurately. Accessing many data sources is challenging, given the complexity of the devices. Furthermore, the dynamic behavior of industrial devices due to factors like wear and tear, software updates, or environmental changes must also be taken into account.

This paper focuses on a data source that can be widely used across various devices and scenarios, namely time information from network communication. The concept is to identify an industrial plant model based on the time data, delays, and regular patterns. Specifically, we aim to identify programmable logic controller (PLC) models based on their time values, utilizing ML techniques to enable such identification.

II. Related work

Device behavior analysis can be divided into two categories: (1) behavioral analysis aimed at identifying the type or model

of the device, and (2) behavioral analysis aimed at detecting anomalous behavior. For the purposes of this paper, we will focus on the former approach. Specifically, we will discuss the data sources related to network communication. Recently, several papers have been published on this topic [2]–[5].

While there are several articles discussing the analysis of internet of things (IoT) devices, none of them are specifically focused on industrial devices. Moreover, these articles only address device type identification and do not consider pure time information within packets. In contrast, our paper is primarily concerned with utilizing timing parameters extracted from network communication. Although previous research has addressed this issue, it has been limited to device model recognition within the IoT domain and did not examine industrial devices.

Sharma et al. [6] conducted research on device type identification using Internet control message protocol (ICMP) timestamps. Their approach involved statistical methods and did not involve the use of ML.

Oser et al. [7] focus on identifying the type of IoT devices using transport control protocol (TCP) timestamp option time. However, their approach does not account for device models and is limited to IoT devices only.

Le et al. [8] also focus on identifying the type of IoT devices rather than their specific models. Their primary method involves calculating clock skew using Timestamps from the TCP protocol.

III. Methodology

A. Dataset and data preparation

We used a total of five physical PLC devices from Siemens for data collection. These devices included two models of the ET200SP, two models of the S7-1500, and one model of the S7-1200, as shown in Table I.

TABLE I
PLC device for dataset collection.

Device	Description
ET200SP_1	Only PLC without scenario
ET200SP_2	PLC with sensors in scenario
S7_1500_1	Only PLC without scenario
S7_1500_2	PLC with sensors in scenario
S7_1200	PLC with sensors in scenario

For data collection, we used the PLCs ET200SP_1 and S7_1500_1 without any sensors connected, and directly connected them to the computer. The ET200SP_2 device was running in a live scenario, and we captured its data through a virtual private network (VPN). In the current scenario, the device is connected to a router in the lab. The device has 4 output LEDs and 4 input buttons, which are used to control the LEDs. Physical access, through toggling the input switches, is required to operate the device. Similarly, the data from the S7_1500_2 device was captured within a live scenario via a switch. In the live scenario, the device serves as a control element in a beer production test environment (mini brewery). This environment enables the entire process of semi-automated brewing, with a number of input and output devices and sensors such as temperature measurement, motors, and liquid quantity monitoring. Lastly, we used the PLC S7_1200, which was also involved in a real scenario, and connected it directly to the computer to capture its data directly. The scenario involved the control of a motor by a driver. In this scenario, there are three devices with their own IP addresses: PLC, HMI, and driver. There are also several input and output sensors. Overall, our data collection involved a diverse set of PLC devices, including those in live scenarios and those directly connected to the computer, providing us with a comprehensive dataset for analysis.

1) Data collection: We collected the network communication packets of the open platform communication unified architecture (OPC UA) protocol from the device by capturing the communication between the OPC UA client and OPC UA server. In this setup, each PLC device worked as an OPC UA server, and we connected to these devices using the OPC UA client via the UAExpert application during testing. To achieve generalization of the model, we only captured specific message types within OPC UA and avoided subscribing to any OPC UA tags from the server. This approach ensured that the training dataset consisted only of the necessary message types. Therefore, we captured the following message types for the training dataset within OPC UA:

- *UA Secure Conversation Message: ReadRequest*,
- *UA Secure Conversation Message: ReadResponse*,
- *UA Secure Conversation Message: PublishResponse*,
- *UA Secure Conversation Message: PublishRequest*,
- *UA Secure Conversation Message: ServiceFault*.

To ensure the datasets' diversity, we collected data over several days and with different settings. Additionally, we captured data from live scenarios where PLCs were connected and communicating with sensors. Moreover, we captured data over VPN to test the accuracy of device identification over potentially affected traffic. This allowed us to evaluate the performance of our method in a more realistic scenario where network traffic may be subject to various types of interference.

2) Data preparation: We only used the time information from the received packets, resulting in the following data:

- *Time UTC* – the time of frame receiving based on the network interface.
- *OPCUA_Timestamp* – the time the OPCUA Client and OPCUA Server sent the message.
- *Frame_Time_Delta* – the time difference between two received packets from Time UTC.
- *TCP_ACK_RTT_Time* – the time in the response from the server.

These parameters alone do not provide sufficient information, so we processed them to create new parameters. We used the rolling average technique with the previous 12 values to process most of the parameters. Starting with *Time UTC*, we created five new values and discarded the original value. For each new value, we calculated the rolling average of the previous 12 values. We created the average of the *Time_UTC_Mean*

values, the maximum value of *Time_UTC_Max*, the minimum value of *Time_UTC_Min*, and the averages of the maximum and minimum values of *Time_UTC_Max_Mean* and *Time_UTC_Min_Mean*.

Next, we processed the *OPCUA_Timestamp* parameter, which is the most important parameter for our analysis. To ensure that this parameter is only dependent on the device and is not distorted by the type of OPC UA client used, we processed it only for packets sent from the OPC UA server. We created five new parameters based on this parameter. The first parameter is the difference between the two previous *OPCUA_Timestamp_PLC_delta* values. We then followed the same procedure as before to create the:

- *OPCUA_Timestamp_PLC_mean*,
- *OPCUA_Timestamp_PLC_min*,
- *OPCUA_Timestamp_PLC_max*,
- *OPCUA_Timestamp_PLC_delta_mean*

values. The *OPCUA_Timestamp_PLC_delta_mean* value was created from the *OPCUA_Timestamp_PLC_delta* values.

Finally, We processed the *Frame_Time_Delta* and *TCP_ACK_RTT_Time* values in the same way as before, creating the following parameters:

- *Frame_Time_Delta_Mean*,
- *Frame_Time_Delta_Min*,
- *Frame_Time_Delta_Max*,
- *TCP_ACK_RTT_Time_Mean*,
- *TCP_ACK_RTT_Time_Min*,
- *TCP_ACK_RTT_Time_Max*.

All of these values are presented in Table II, which includes a total of 19 input features that were created and used in the initial phase. However, during the initial training attempts, there was a significant amount of overfitting and poor performance on the test datasets. To address this issue, we had to remove some features from the dataset. The "After drop" column in the table indicates which features remained after the others were dropped.

3) Datasets: For this work, four datasets were created: one for training and three for testing in different scenarios.

The *training dataset* was created from strictly separated data collected on separate days from the other datasets. It was composed only of data from *ET200SP_1* and *S7_1500_1* devices, which were not involved in any real operation scenario and are different from the *ET200SP_2* and *S7_1500_2* devices utilized in later testing phase. Additionally, data from the *S7_1200* device was included in this dataset. The dataset was balanced with respect to the number of samples for each label. Figure 1 shows the distribution of labels in this dataset.

The *test_set_1* was created using data collected at different time periods and on different days from the ET200SP_1, S7_1500_1, and S7_1200 devices connected to the OPC UA Client under different conditions. The purpose of this scenario was to test the device recognition capability of the model in the particular environment in which it was trained, with the aim of improving cybersecurity in that environment. By training the model on a specific type of device, the learned device behavior can be used to detect anomalies on that device. This is because the device's behavior exhibits certain patterns, and variations in these patterns can be observed by the trained model. The *test_set_1* was created for this purpose. The size of this set is 1473 rows.

The *Test_set_2* was composed solely of data collected from *ET200SP_2* and *S7_1500_2* devices that were not part of the training set. Data from the S7_1200 device was also

TABLE II
TRAINING PHASE INPUT PARAMETERS OF SELECTED ALGORITHMS.

Feature	After drop	Description
Time_UTC_Mean	✓	rolling_mean\break (Time UTC)
Time_UTC_Max		rolling_max\break (Time UTC)
Time_UTC_Min		rolling_min\break (Time UTC)
Time_UTC_Max_Mean	✓	rolling_mean\break (Time_UTC_Max)
Time_UTC_Min_Mean		rolling_mean\break (Time_UTC_Min)
OPCUA_Timestamp_PLC_delta	✓	diff\break (OPCUA_Timestamp)
OPCUA_Timestamp_PLC_mean	✓	rolling_mean\break (OPCUA_Timestamp)
OPCUA_Timestamp_PLC_min	✓	rolling_min (OPCUA_Timestamp)
OPCUA_Timestamp_PLC_max	✓	rolling_max (OPCUA_Timestamp)
OPCUA_Timestamp_PLC_delta _mean	✓	rolling_mean (OPCUA_Timestamp_PLC _delta)
Frame_Time_Delta		without change
Frame_Time_Delta_Mean	✓	rolling_mean (Frame_Time_Delta)
Frame_Time_Delta_Min	✓	rolling_min (Frame_Time_Delta)
Frame_Time_Delta_Max	✓	rolling_max (Frame_Time_Delta)
TCP_ACK_RTT_Time		without change
TCP_ACK_RTT_Time_Mean		rolling_mean (TCP_ACK_RTT_Time)
TCP_ACK_RTT_Time_Min		rolling_min (TCP_ACK_RTT_Time)
TCP_ACK_RTT_Time_Max		rolling_max (TCP_ACK_RTT_Time)

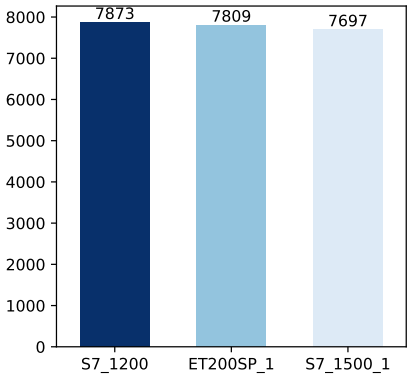


Fig. 1. Training dataset labels.

included in this set. The purpose of this testing scenario was to determine if the given parameters could be used to identify the model of a device using the generic device model trained by the algorithm. The set consisted of 985 rows.

The *Test_set_3* comprises all devices, but it was created specifically to address poor predictions in identifying the *S7_1500 device*. More information on the rationale for creating this set can be found in Section V. This test set has only two labels as the *S7_1500* and *ET200SP* devices were merged into a single label. The dataset contains 8943 lines.

B. Algorithms

We have a classification problem and have tested several classical ML algorithms to find the best solution. The classification algorithms we evaluated include random forest (RF), logistic regression, k-nearest neighbors, XGBoost (XGB), CatBoost, decision tree, AdaBoost, and support vector machines. After testing these algorithms, we found that XGB and RF were the most successful. As a result, we will focus our discussion on these two algorithms. We have included Table III to show the hyperparameters we used to obtain the best results.

TABLE III
HYPERPARAMETER TUNING FOR RF AND XGB ALGORITHMS.

Algorithm	Hyperparametrs
XGB	objective = 'multi:softmax' num_class = 3 eval_metric= 'merror' subsample = 0.5 reg_lambda = 0 reg_alpha = 0 n_estimators = 50 max_depth = 1 max_delta_step = 5 learning_rate = 0.01 gamma = 0.1 colsample_bytree = 1.0 alpha = 0 lambda = 1
RF	n_estimators = 10 min_samples_split = 50 min_samples_leaf = 5 max_depth= 3 criterion = 'gini' bootstrap = True

IV. RESULTS

The results of the two best-performing models in the classification problem are presented in Table IV. As previously stated, the trained models were evaluated on three test sets that represent different scenarios. The first scenario involved recognizing known devices to track changes in their behavior, and both the XGB and RF models performed well, achieving an accuracy score of approximately 0.99. In the second scenario, which involved recognizing the same device model but with an unknown device in the training set, the accuracy score for both models was lower, around 0.71. In the last scenario, where the unknown test data for the training model was merged as one device, the XGB algorithm achieved a high accuracy score of 0.8964, while the accuracy score for the Random Forest algorithm was lower, at 0.7373.

TABLE IV
FEATURES OF RF AND XGB ALGORITHMS IN CLASSIFICATION PROBLEM.

Scenario	Algorithm	Acc	Prec	Rec	F1
Test_set_1	XGB	0.9952	0.9953	0.9952	0.9952
	RF	0.9939	0.9939	0.9939	0.9939
Test_set_2	XGB	0.7076	0.5993	0.7076	0.6290
	RF	0.7147	0.8727	0.7147	0.6457
Test_set_3	XGB	0.8964	0.9078	0.8965	0.8747
	RF	0.7373	0.8666	0.7373	0.7709

Figure 2 shows a large overlap in the results obtained by both training algorithms. The figure on the left represents the results of the Random Forest algorithm, while the figure on the right shows the results of the XGBoost algorithm. Both algorithms accurately identified most of the devices, with only minor differences in the identification of the ET200SP device. Overall, the devices were identified with good accuracy.

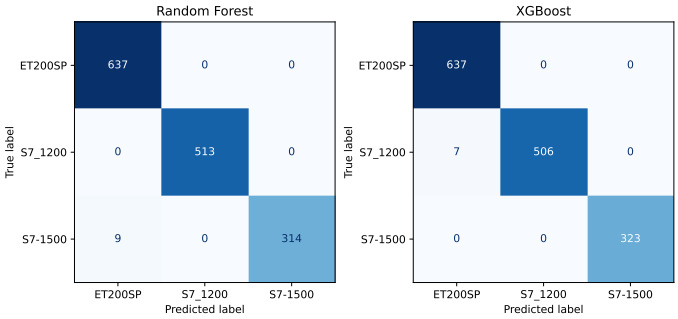


Fig. 2. Test_set_1 results.

Figure 3 highlights an important issue where the algorithm frequently misclassifies the S7-1500 device as an ET200SP device. We suspect that this is because the ET200SP device shares the same hardware configuration as the S7-1500 device, leading to confusion for the algorithm and resulting in poor recognition of the ET200SP device. To test this theory, we developed a test set which is described below.

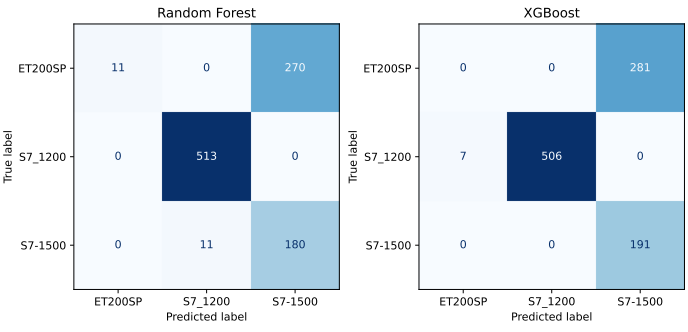


Fig. 3. Test_set_2 results.

As previously mentioned, the third test scenario involved evaluating the performance of the identification process when treating the ET200SP and S7_1500 devices as a single device due to their identical hardware configuration. Figure 4 shows

that the RF algorithm is not particularly suitable for this approach, as it produced a large number of outliers.

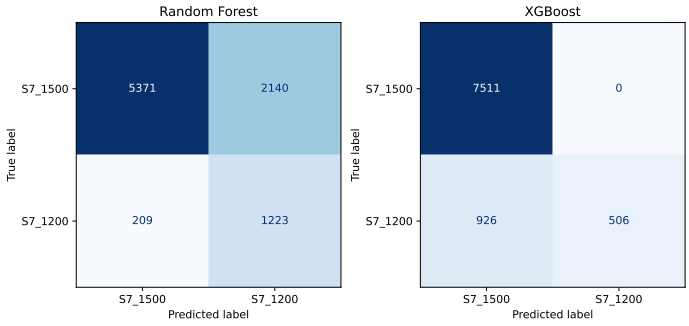


Fig. 4. Test_set_3 results.

Instead, the XGB algorithm is better suited for this purpose. To improve accuracy, future work should involve using a more diverse and extensive training set as well as a larger test set.

V. Discussion

The confusion matrices revealed that the main issue in the identification process was predicting the S7_1500 device incorrectly. To address this, we developed a special test set called Test_set_3. The challenge in distinguishing between the ET200SP and S7_1500 devices is that they have identical hardware configurations, making it difficult to differentiate between them. Therefore, with the parameters we used, the identification accuracy was not very high. Even active scanning tools that scan the network can return the value of S7_1500 instead of ET200SP, further compounding the identification issue. While our solution can accurately identify the ET200SP device, the problem persists with the S7_1500 device. As a result, we created the Test_set_3 to evaluate the model's accuracy when treating both devices as a single device with the same label.

One of the limitations of this study is the current lack of training data available to refine the device identifications. Increasing the size of the training set may improve the performance of neural networks and produce better results. However, a significantly larger test set would be required for this purpose. Another significant limitation was the limited number of devices used in the study. We only used five devices, with two identical models being tested twice. To improve and expand on this work, we plan to include more devices in future studies. The goal is to have at least three devices of the same model and use the training data from two of these devices to enhance the identification capabilities.

VI. Conclusion

In this study, we aimed to explore the feasibility of identifying industrial PLC devices based on time data obtained from network communication, specifically from the OPC UA protocol. As the amount of temporal information available was limited, we augmented the features with newly computed ones from the existing features. Our study demonstrated the potential of using timing data as input parameters for device identification. We then selected appropriate parameters from these newly created features and trained a variety of classification algorithms. Among these algorithms, we identified RF and XGB as the top performers. To evaluate the effectiveness of our approach, we conducted experiments on three different test sets. Our results showed higher success rates for Test_set_1 and Test_set_3 scenarios, whereas the performance of Test_set_2 was lower due to reasons previously discussed. One limitation of this study is the lack of training data, which could be addressed by collecting more data for refining the identification model. Furthermore, we plan to expand the study by incorporating more devices, which could enhance the identification capabilities.

In future work, we aim to expand our proposed solution to include more PLC models, which would enhance the training potential and improve the testing results. To improve the identification accuracy between the S7_1500 and ET200SP models, it would be beneficial to have data collected from at least two of each model in the training set to refine their behavioral patterns. Additionally, we plan to collect more data for the training set to create a Neural Network and compare its performance with classical ML algorithms.

Acknowledgment

The presented research is a part of the project reg. no. FW07010004, financially supported by the Technology agency of the Czech Republic under the 7. TREND programme.

The authors thank Dr. Radek Fujdiak and Dr. Pavel Masek from the Brno University of Technology for their expertise and assistance throughout all aspects of this study and for their help in writing the manuscript.

References

[1] "Industrial Control Systems (ICS) security Market by Component (Solutions and Services), Solution, Service, Security Type (Network Security, Endpoint Security, Application Security, and Database Security), Vertical and Region - Global Forecast to 2027", *Markets and Markets*, 2022.

[2] X. Liu, Y. Han, and Y. Du, "IoT Device Identification Using Directional Packet Length Sequences and 1D-CNN", *Sensors*, vol. 22, no. 21, 2022.

[3] F. Yin, L. Yang, Y. Wang, and J. Dai, "IoT ETEI: End-to-End IoT Device Identification Method", *2021 IEEE Conference on Dependable and Secure Computing (DSC)*, pp. 1-8, Jan. 2021.

[4] C. Duan, H. Gao, G. Song, J. Yang, and Z. Wang, "ByteIoT: A Practical IoT Device Identification System Based on Packet Length Distribution", *IEEE Transactions on Network and Service Management*, vol. 19, no. 2, pp. 1717-1728, 2022.

[5] I. Hafeez, M. Antikainen, A. Y. Ding, and S. Tarkoma, "IoT-KEEPER: Detecting Malicious IoT Network Activity Using Online Traffic Analysis at the Edge", *IEEE Transactions on Network and Service Management*, vol. 17, no. 1, pp. 45-59, 2020.

[6] S. Sharma, A. Hussain, and H. Saran, "Experience with heterogenous clock-skew based device fingerprinting", *Proceedings of the 2012 Workshop on Learning from Authoritative Security Experiment Results*, pp. 9-18, Jul. 2012.

[7] P. Oser, F. Kargl, and S. Lüders, "Identifying Devices of the Internet of Things Using Machine Learning on Clock Characteristics", *Security, Privacy, and Anonymity in Computation, Communication, and Storage*, pp. 417-427, 2018.

[8] F. Le, E. Valdez, and P. -C. Cheng, "Counting Devices: Revisiting Existing Approaches in Today-s Settings", *2019 IEEE International Conference on Big Data (Big Data)*, pp. 4032-4037, 2019.

Multiprotocol Energy Load Management System: A Customer-Side Approach

Lukas Benesl*, Radovan Juran†
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
*lukas.benesl@vut.cz, †radovan.juran@vut.cz

***Abstract*—The trend towards decentralized energy production is rapidly increasing along with the growing variability of the loads connected to the grid, making the concepts of Smart Grid significantly irreplaceable. However, the visions of communication-aided energy distribution sometimes encounter the problem of real-world implementation, as is also the case with the communication channel from the energy distributor to the customer and backwards. Incompatibility of communication protocols in currently available smart electricity meters and loads at the side of the customer does not allow direct means of information exchange, thus we propose a multiprotocol energy load management system, which aims to fill this communication gap, capable of bridging the omnipresent, widely used DLMS/COSEM and Modbus protocols, which we aim to further evaluate in a testbed based on real-world devices. On top of that, the article provides a brief analysis of the main load elements that the distributor is able to regulate according to the legislation.**

***Index Terms*—Smart Grid, e-mobility, charger, wallbox, photovoltaics, inverter, RS-485, DLMS/COSEM, protocol, PLC, BPL**

I. Introduction

The amount of renewable energy sources being deployed on the customer's side steeply grows, which is beneficial as it reduces costs per unit of consumed power and allows to sell the excess produced energy back to the grid. However, this decentralization of production can greatly affect the quality of electricity in the rest of the local grid and also the stability of the distribution system as a whole [1]. Legislation obliges the distributors to manage production and consumption, but the customer has the opportunity to influence not only his own consumption but also the production of energy. However, the distributor will also be able to limit the charging power of the wallbox in order to prevent overloading of the distribution system, theoretically bring the grid to a possible blackout [2].

This leads to the need of having means of communication between the distributor and the customer for direct control and regulation or to provide certain feedback and set prioritization; ideally a device capable of filling this gap in the communication chain. On top of that, in case of widespread deployment, the potential of such a device lies also in the usability of the collected data for accurate monitoring of consumption and system load with the possibility of predicting electricity consumption, manage community energy sharing, and efficient use of energy accumulation and storage, be it static batteries or electric vehicle (EV) charging – wallbox. Due to the limited native functions of currently deployed electricity meters, it is not possible to collect an utilize any detailed data directly *in situ* on the consumption point for energy monitoring and providing any feedback information in terms of control and regulation. There is also a lack of communication interoperability between end-devices of various types and vendors, ranging from domestic power plants to e-mobility charging stations, and thus insufficient means to be able to effectively ensure enough energy flexibility based on individual consumption points (residences, commercial buildings, industrial facilities, etc.). The existing infrastructure does not allow the creation of a communication channel with energy production and consumption devices for control or data collection using the electricity meter based on distributor commands or local evaluation [3].

With the aim to build a testbed for evaluating methods of load management based on real-world, commercially available charging stations, we propose a testing prototype of a device to fill the gap in the communication chain between the distributor's requirements and customer's needs, providing vendor-agnostic interoperability of communication and data aggregation capabilities for additional analysis.

The paper is structured as follows, Section II introduces the analysis of the customer-side device. Section III presents the device design itself, closing with future work visions in IV.

II. Market analysis of the Loads

Before the design of the device itself, it was necessary to perform market analysis, which is presented in this section, where available wallboxes II-A, photovoltaic (PV) inverters II-B, and other potentially controllable loads, all with their corresponding interfaces and protocols available II-D are summarized and discussed.

A. Wallboxes

This sub-section provides an overview of home charging stations (wallboxes) that are available for the Czech and European markets. Tab. I also includes wallboxes that can be delivered including an electric car. In order to analyze the communication interface and protocol, the parameters of the physical layer were searched, i.e. what resources the charging station itself is equipped with and what communication protocol they use.

TABLE I
OVERVIEW OF WALLBOX MANUFACTURERS FROM THE POINT OF VIEW OF INTERFACES AND PROTOCOLS

Wallbox manufacturer	Physical layer	Protocol	Source
Terra	PLC, RS-485, ETH, 4G, Wi-Fi	OCPP 1.6, Modbus	[4]
BMW iWALLBOX	Wi-Fi, RS-485	Modbus	[5]
PCE Wallbox	RS-485, Wi-Fi	OCPP, Modbus	[6]
Webasto Next	GSM, RS-485, ETH, Wi-Fi	OCPP 2.0, Modbus	[7]
WALLBOX eMH2	RS-485, USB, ETH	OCPP 1.6	[8]
WALLBE ECO 2.0	ETH, RS-485	Modbus	[9]
Tesla Gen 3	Wi-Fi, RS-485	(unknown)	[10]
OlifeEnergy	ETH, Wi-Fi, GSM, RS-485	OCPP 2.0, Modbus	[11]

Based on the market research analysis the most used interface for communication are wired technologies, typically RS-485 and Ethernet. There are only three communication protocols used by the listed charging stations: Modbus TCP, Modbus RTU, and OCPP (Open Charge Point Protocol) in version 1.6 or 2.0. Communication between the charging station and another element is most often realized using the OCPP protocol with a backend system. The manufacturer then gives access to charging statistics and charging settings. Some charging stations can be connected through a Home Energy Management System (HEMS), which enables the charging of an electric car from PV plants or charging using dynamic load distribution. However, the key finding of our analysis is that no commercially available solution is not considering the DLMS/COSEM protocol. This is our main motivation, because the DLMS/COSEM is a widespread protocol adopted by smart meters, and is therefore the key connecting element to the distributor.

B. PV inverters

In 2009 and 2010, when there was a shortage of inverters on the market, new vendors started to emerge rapidly, but any standardization or unifying communication protocols were not considered a priority for manufacturers, primarily due to the rapid need of fulfilling the market demand, hence foiling any opportunities to develop and deploy vendor-agnostic systems and solutions ensuring interoperability for future integration. The fact that each inverter has a different communication protocol was taken by many manufacturers as an advantage, offering proprietary solutions, such as Solar-Log, which is based on proprietary knowledge with a signed NDA (Non-Disclosure Agreement) contract for the use of the protocol as one of the few practically exclusive manufacturers. The analysis of inverters is given in Tab. II.

Based on the analysis of the interfaces and protocols, it can be concluded that the RS-485 interface is used the most in the off-the-shelf products with the Modbus RTU or Modbus TCP protocol is utilized. On the other hand, the DLMS/COSEM protocol is not considered any commercially available solution. Currently, the communication of the inverters via a serial line using Modbus RTU only with a secondary electricity meter (e.g. EASTRON SDM630-Modbus) is considered.

C. Other loads

Another load on the customer's side can be water heating (boiler) and household heating. HDO (Mass Remote Control) is still used to switch this load. The control unit sends out a pulse several times in a row at a certain time and the HDO receiver picks up this pulse and opens or closes the relay/contactor, causing a change of state. This service is deprecated, communication is one-way only, and the command that was supposed to happen cannot be confirmed.

Communication between the Home Area Network (HAN) interface of the customer's smart meter and the elements of consumption and production at the customer. The potential users of the result are therefore:

- Distribution companies – compensation for insufficient HDO or use for elements of consumption and production beyond the number of relays of a smart meter.
- End customers – for the involvement of production and consumption in the electricity market. The customer has the option to dispose of the excess electricity himself and sell it at spot prices on the electricity market.
- Merchant and aggregator – PLC/BPL will be the communication link between the customer (with a smart meter) and the trader, i.e. the electricity market.

D. Protocols and interfaces

The entry-point for a distributor to the wiring installation of a site is an electricity meter (smart meter), which typically utilizes DLMS/COSEM protocol, which is specifically designed to communicate with energy meters and similar devices. It allows the transfer of various data including electricity consumption measurements, connection status, voltage quality, and other information. It also supports many other functions such as device status feedback, programming, and firmware updates. The DLMS/COSEM protocol is used for communication between smart meters and central control systems. The main advantage of this protocol is that it enables the communication between different types of electricity meters and control systems, which ensures interoperability and reduces implementation and maintenance costs. The protocol also improves the efficiency and accuracy of data collection, which can lead to reduced power consumption and energy costs.

TABLE II
OVERVIEW OF PV INVERTER MANUFACTURERS FROM THE POINT OF VIEW OF INTERFACES AND PROTOCOLS

PV inverter manufacturer	Interface	Protocol	Source
Fimer/ABB	ETH, RS-485, Wi-Fi	Modbus RTU/TCP	[12]
Fronius	ETH, RS-485, WLAN	Modbus RTU/TCP	[13]
SMA	WLAN, RS-485	Modbus	[14]
Solar Monitor	ETH, RS-232/422/485	Modbus, DNS, UDP	[15]
RCT Power Storage DC 6.0	RS-485	Modbus	[16]
Infinisolar 10K-PAR-B	RS-485	Modbus	[17]

1) DLMS/COSEM protocol: DLMS/COSEM is a widely used communication protocol in intelligent energy management systems. The protocol is used for communication between electricity meters and energy management systems (Smart Grid and Smart Metering), enabling monitoring and control of energy consumption in real-time. Furthermore, it enables the transmission of energy consumption data between energy meters and control systems, monitoring and managing energy consumption to optimize energy efficiency.

To identify and manage data objects in DLMS/COSEM to be able to collect, store, and analyze energy usage data from a variety of different sources, the Object Identification System (OBIS) codes are defined. By using a standardized system for identifying data objects, DLMS/COSEM ensures interoperability and compatibility between different devices and systems.

Based on the outputs of the NAP SG study "Technical Specification of Customer Interface" [18], the RS-485 interface and the DLMS protocol are considered for the customer interface of the electricity meter. In order to use DLMS/COSEM for communication between smart meters and charging stations, DLMS UA has launched its "Electric Vehicle Charging Station" (EVCS) working group. The association encourages participants in the e-mobility ecosystem, including manufacturers of charging stations and electric vehicles, e-mobility operators, electricity suppliers or distributors, to participate in the DLMS EVCS working group and use the proven and recognized DLMS industry standard as well as the DLMS certification system. Extending the DLMS standard to the EVCS area and the e-mobility industry will unify communication standards and prepare for direct charging regulation and control [19].

2) Modbus protocol: Modbus is a serial communication protocol used to connect industrial devices. The protocol enables data transfer between the control system (server) and peripherals (client) in industrial devices. Data is transmitted serially via RS-232/422/485 lines. The Modbus protocol enables the control of various industrial devices, including sensors, frequency converters, valves, pumps, control systems, and other elements. The protocol also supports multiple functions such as reading and writing data, reading and writing registers, forcing events, timing, and more. The Modbus protocol is very robust and reliable and is therefore very suitable for use in industrial applications. The analysis showed that wallboxes and photovoltaic inverters most often use the RS-485 serial line, on which the Modbus communication protocol is used.

The Fig. 1 shows the Automated Meter Management (AMM) communication chain for smart metering, introducing the possibility of end device regulation, altogether forming the core idea of our intended testbed.

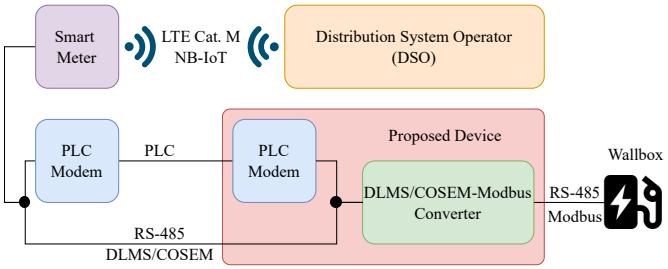


Fig. 1. Testbed featuring the devices in the communication chain.

The communication chain that we are dealing with consists of a channel that allows a full flow of information between the Distribution System Operator (DSO) and the customer. It opens a direct connection through the smart meter (SM), which interfaces the loads via the proposed protocol converter. The DLMS/COSEM to Modbus protocol converter translates the information contained in the specific OBIS codes into the message format used by the Modbus protocol comprehensible to the loads.

Additionally, we take into account the utilization of the Power-Line Communication (PLC) technology. Having a device capable of leveraging PLC can be an advantage, as it is a cost-effective way to establish communication networks in the case of older buildings or in general any specific places where requirements to install a dedicated communication infrastructure can not be met.

This brings us to the main motivation and sets the objectives of proposed work:

1) **communication** – filling the gap of the missing part in the communication chain implementing a protocol converter with data aggregation capabilities,
2) **integration** – deployment in a testbed for testing the interprotocol information exchange between real-world devices,
3) **evaluation** – running various real-time scenarios, data collection and analysis.

III. The Proposed Device

The motivation to create this device stemms from the need to bridge the described gap in communication chain and integration of this solution into a testbed ecosystem. Although there are off-the-shelf solutions, they do not meet the described needs – cost, size, application specific design, user interface, protocol mapping, etc. Sub-chapter III-A offers a more detailed view and compares available commercial solutions. The block diagram and description of the proposed device are presented in subsection III-B.

A. Commercially available solutions

As Tab. III shows, there are some existing available solutions able to serve as the needed protocol converter to some extent, but they are relatively of high cost and may not be able to convert protocols in both ways.

TABLE III
Existing commercial devices

Device	Interfaces	Price [€]	Source
EXCELVISOR ET-DCU-xC	1x ETH - RJ45 1x RS-485 - RJ45	(unknown)	[20]
MarCom DLMS Gateway	3x USB 3x ETH - RJ45 1x RS-232 - DB9 1x RS-485 - DB9	≈ 900	[21]
FANOX SYNC 2000 M1	2x RS232/485 - RJ45 1x ETH - RJ45	≈ 550	[22]

For the testbed, it is necessary to create specific objects according to OBIS, which will store the necessary information for controlling the given end element.

Available commercial devices are usually not cost-effective for use in a testbed. Also, their size is not very suitable for mounting on a DIN rail. Not all products can be mounted on a DIN rail. The size of the device itself is indicated by only one manufacturer. Powering the device is a solution using an external source. Only one device allows you to use the 230 V power supply.

Due to the disadvantages mentioned above, it was more appropriate to consider the design own device with a user interface into which it will be possible to enter and change any functionality as needed. In addition, the proposed device will be able to use an alternative communication interface – PLC. This can be very helpful in a situation where it would be difficult/inconvenient to install a serial communication cable.

B. Device design

The whole device is designed as a part of a larger tested, therefore we approached the whole concept with scalable prototyping in mind. The whole idea of the proposed device is depicted in Fig. 2.

The core of the proposed system is the Raspberry Pi Compute Module 4 (CM4), which thanks to its versatility, compact form factor, and variety of interfaces is an accessible embedded development platform for convenient hardware and

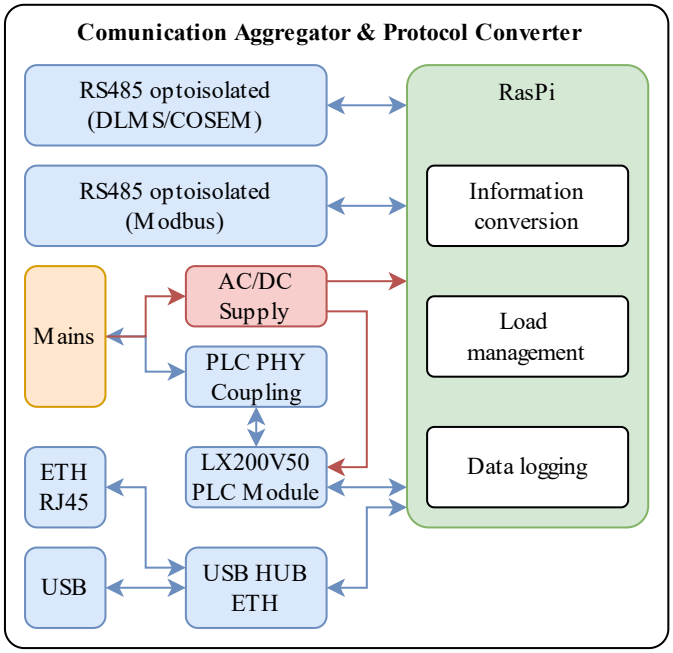


Fig. 2. Converter DLMS to Modbus – Block schematic diagram.

software prototyping. It was selected intentionally before any narrowly-specific microcontroller, as it brings a large performance reserve. It aims to be capable not only of performing the main task in the specified communication chain for conversion of information for functional control of the electric car charging process, but also to act as a data aggregator for monitoring and verification of communication parameters on the subjected interfaces and test the interoperability of the protocols in question. It features the following interfaces:

- 2x screw terminal RS-485,
- 1x screw terminal for 230 AC & PLC,
- 2x RJ45 Ethernet (one via PLC),
- 2x USB.

The proposed system is seated in a DIN-rail mount housing to fit in a standard switchboard. It does not require an external power source as it is intended to utilize PLC and thus it can be powered directly from the mains (230 V AC). Inside there is an AC/DC converter, supplying the required DC voltage for the CM4, which is a single +5 V, +3.3 V and +1.8 V to peripherals. The PLC is managed by Qualcomm QCA7550 system-on-a-chip transceiver, utilizing snap-in LX200V50 module by WisLink [23], which bridges communication between the PLC and the CM4 via Ethernet. The key role play the two optically isolated screw terminals for RS-485 interfaces for the DLMS/COSEM and Modbus communication.

IV. Discussion

The main contribution of the work is the creation of a testbed and verification of the method of communication or regulation of the end element on the customer's side.

At first glimpse, the proposed platform is packed with a significant computing performance reserve for a simple

protocol-to-protocol conversion task. However, we aim to create an easy-to-use temporary solution for quick prototyping and deployment into the intended testbed with real commercial devices, which would allow not only to perform the validation of the procedure of data crossing protocols and messaging topologies, but also the control and management of customer-side workloads.

However, in the future, we aim to expand the capabilities of the proposed device to a device we envision as more general Smart Box, will perform many more demanding tasks in real-time. These functionalities can include dynamic load balancing (simultaneous charging of several cars), protection against overloading of the home network, optimization of the use of produced solar energy, acceleration of the EV charging process, working with the data of the produced energy and directing the delivery in the appropriate direction, given the current state of load networks (batteries, appliances, charging, etc.)

A fundamental requirement beyond the connectivity of the box with the electricity meter on the secondary interface is, in particular, sufficient computing power, which will enable the implementation of more advanced functionalities and a sufficient level of cryptographic security within the framework of cyber security.

V. Conclusion

The trend of growth in the number of PV and charging stations is currently very current. Many households can now supply electricity to the distribution network and the distributor cannot regulate these new sources. For this reason, the issue of enabling communication was outlined, which is presented in Fig. 1. To be able to regulate end devices, it is necessary to use a protocol converter. The designed device shows bring a bridge for future-proof interoperability.

Another continuation of the work in this issue is the creation of a device prototype that will be able to read data from a smart meter and, according to the internal logic settings, convert it into specific messages for the charging station, as it is intended in the testbed.

Acknowledgment

The authors would like to acknowledge the valuable contribution and materials provided by the supervisor, assoc. prof. Ing. Petr Mlýnek Ph.D., in the completion of this paper.

References

[1] M. Karimi, H. Mokhlis, K. Naidu, S. Uddin, A.H.A. Bakar. (2016). *Photovoltaic penetration issues and impacts in distribution network – A review.* Renewable and Sustainable Energy Reviews. 53. 594-605. DOI: https://doi.org/10.1016/j.rser.2015.08.042.

[2] EUROPEAN COMMISSION. *EU Solar Energy Strategy.* EUR [cit. 2023-03-06]. Available from: https://bit.ly/3F6U42J

[3] *Ministerstvo průmyslu a obchodu* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3L4Fc8A

[4] Terra ac wallbox. [online]. *ABB* [cit. 2023-03-06]. Available from: https://bit.ly/3Zqvthh

[5] *BMW Wallbox Plus/Connect.* [online]. BMW [cit. 2023-03-06]. Available from: https://bit.ly/3SW0dUX

[6] Charging stations wall mounted. *PC Electric* [online] [cit. 2023-03-06]. Available from: https://bit.ly/3ZIEiml

[7] Webasto Next - Autonabijeni.cz. *Autonabíjení.cz - Vše pro Váš elektromobil* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3YoI5nV

[8] WALLBOX eMH2. *Conrad* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3ITwYxF

[9] WALLBE Driving eMobility. *ESL e-mobility* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3JeoAdo

[10] Tesla Wall connenctor 3. generace – příručka. *Tesla* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3KVIDP3

[11] OlifeEnergy Wallbox [online]. *SECTRON* [cit. 2023-03-06]. Available from: https://bit.ly/3kQDVaF

[12] Fimer Spa. *Fimer* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3ZmOBwE

[13] Měnič 3-fázový měnič FRONIUS SYMO 6.0-3-M LIGHT 6KW. *eltrox.store* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3muKsIA

[14] SMA SUNNY TRIPOWER 3-fázový invertor 10.0 10kW. *eltrox.store* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3SU4yrC

[15] Start / 60 / 300 / 1000 - Solar Monitor. *Solar Monitor* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3YxyyLu

[16] RCT Power Storage 6kW DC 6.0 -DC-Coupled Hybrid Inverter for 3 phases. *ETRONIX CENTER* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3JdKhdv

[17] InfiniSolar 10KW. GBC Montáže s.r.o. *nemakej* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3YoT9kX

[18] Studie z projektů Národního akčního plánu pro chytré sítě (NAP SG) – MPO. *Ministerstvo průmyslu a obchodu* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3YqUho4

[19] DLMS to incorporate smart EV charging in standard. *SMART ENERGY* [online]. [cit. 2023-03-06]. Available from: https://bit.ly/3YqQebs

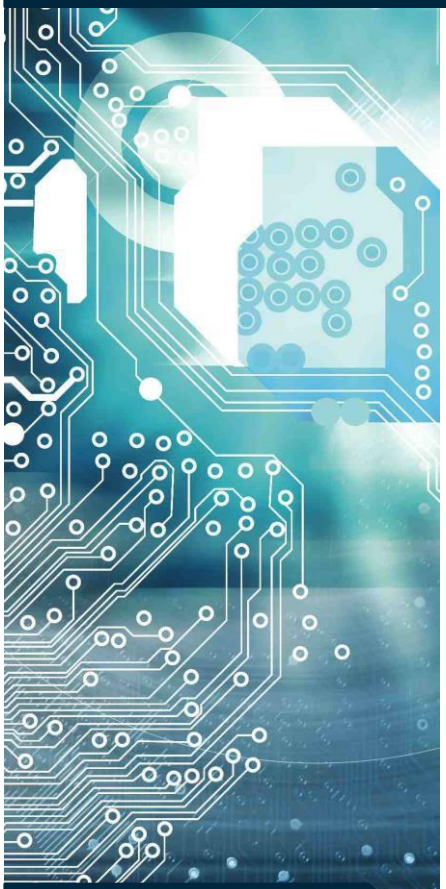
[20] DATA CONCENTRATOR UNIT(DCU). *excelvisor* [online]. [cit. 2023-03-06]. Available from: https://rb.gy/xhgfsu

[21] ADVANCED DLMS GATEWAY -MODBUS TCP/RTU. *MarCom* [online]. [cited 06.03.2023]. Available from: https://rb.gy/c1zmps

[22] SYNC 2000. *FANOX* [online]. [cit. 2023-03-06]. Available from: https://rb.gy/v6otzs

[23] *RAK Wireless: LX200V50 Power Line communication(PLC) Module* [online]. [cit. 2023-03-12]. Available from: https://store.rakwireless.com/products/lx200v50-power-line-communicationplc-module

Získej víc s **IEEE**



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 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 či Siemens. Značka IEEE se 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. U nás má aktivity společnosti IEEE na starosti Československá sekce IEEE, https://www.ieee.cz.

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

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. IEEE také 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. IEEE podporuje publikační činnost a svým členům umožňuje volný přístup k licenci Overleaf PRO.

www.ieee.org

IEEE

Point cloud obstacle detection with the map filtration

1st Lukáš Kratochvíla
Department of Control and Instrumentation
Brno University of Technology
Brno, Czechia
ORCID: 0000-0001-8425-323X

***Abstract*—Obstacle detection is one of the basic tasks of a robot movement in an unknown environment. The use of a LiDAR (Light Detection And Ranging) sensor allows one to obtain a point cloud in the vicinity of the sensor. After processing this data, obstacles can be found and recorded on a map. For this task, I present a pipeline capable of detecting obstacles even on a computationally limited device. The pipeline was also tested on a real robot and qualitatively evaluated on a dataset, which was collected in Brno University of Technology lab. Time consumption was recorded and compared with 3D object detectors.**

***Index Terms*—3D point cloud, obstacle detection, robot, environment map, LiDAR**

I. Introduction

Robotics is a sector that is experiencing a huge boom nowadays. Different types of industries e.g. automotive, aerospace, chemical, and others, as well as the social care field, are increasingly discovering, adopting, and seeking opportunities to robotize various human tasks. However, the main driver of progress is an industry. The need to automate manufacturing processes, optimize production, and move objects requires specialized equipment capable of working autonomously.

One of the first steps towards this goal is to create a robot capable of avoiding obstacles. The robot must not only sense its environment but also understand it, map it, be able to navigate through it, and react to changes in it. For these tasks, the ROS (Robot Operating System) environment has been created to implement different nodes that can communicate with each other and transfer data. ROS architecture is depicted on Fig 1. An essential part of such a robot capable of avoiding obstacles is a localization and navigation system. If the robot could not locate itself, it would not be able to navigate through the map and perform the tasks given. On the other hand if the robot localizes itself and knows where it is going, it must be able to map new obstacles it encounters and avoid them if necessary.

There are several ways in which we can perceive the robot's environment. Some rely on visual information, in which a large amount of data is hidden, but it is necessary to understand what is in the scene. Others explore the surrounding environment directly using distance-measuring sensors such as LiDAR, stereometry or ultrasound. This paper focus on data from a LiDAR sensor.

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.

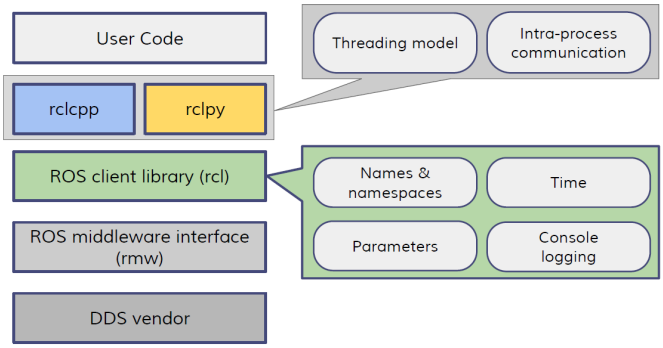


Fig. 1: The ROS architecture overview [1]. Whole library consists of five main parts. At the top of the pyramid is used code in C++ or python. Next level is responsible for translation of the user code. Level below is client library with namespaces, logging, parameters and time clocking. On the bottom is communication provider.

The work is structured as follows. The section II describes papers dealing with similar topics. The section III describes methods used in a new pipeline. The pipeline design in the following section IV is proposed based on this research. The section V discusses the implementation of the pipeline and the results in the experimental setup. The section VI summarizes the achievements of the whole paper.

II. Related articles

Various methods are used to detect objects in a point cloud and some will be mentioned. Currently, the development focus on deep learning models. The models can be divided into two categories, i.e. single-step and multi-step.

The first group [2]–[6] works directly with the raw point cloud and searches for various objects in it. [2] builds detection on feature extraction, convolutional layers and finding suitable candidates for output. [3] takes a similar approach but modifies convolution for sparse input data. [4] leverages previous experience from 2D SSD (Single Shot Detector) model. [5] uses graphical representation of data for computing output.

The second group [7]–[10] performs multiple steps to detect an object. [7] first detects parts of the object and then aggregates these parts to outputs. [8] also performs two steps, first detecting based on the raw data and second fine-tuning the detections with a second model. [9] uses model consisting of three steps. In the first step, features are generated based on the defined structures. In the second step, information is compressed into a more compact representation, and in the last step, detections are sought. [10] based on structured features, it uses them to create segmentation and classification of objects.

There is also a group of methods that try to achieve detection using classification on a 2D image. [11]–[13] try to use classification on an image to achieve 3D detection. [11] create a representation of the image using attributes from a convolutional network. Afterwards it calculates the angle of the object, the size of the object in 3D, and the confidence of detection based on these features. [12] use a fully connected network to estimate the 2D detection, orientation of the object (direction for expansion to 3D), make a cubic assumption of the 3D object and extract the features. In the second step, they extract the 3D detection using a 3D model based on the extracted features. [13] try to reconstruct objects in 3D from 2D image and create 3D detections based on them.

III. Explanation of terms

The *Voxel Grid* is a representation of points in space that allows for a significant way to enable point manipulation. This representation divides space into small blocks and each block is represented by a center point. If the number of points in a given block is sufficient, the block will be implemented in the Voxel Grid, otherwise it will not. This method also allows to perform data filtering. This method significantly reduce number of points in a point cloud. Therefore it is absolutely necessary if we want to perform the calculations on a power-limited device.

The RANSAC (Random Sample Consensus) [14] method is one of the simplest methods for finding correspondence points. The method selects a random subset of the data and tries to find a pattern, then tests the remaining points and calculates the error. This procedure is repeated multiple times and for the result is chosen a model with the minimal error. This method can also be used to find surfaces in a point cloud.

KDtree (short for k-dimensional tree) [15] is a method that uses binary trees to represent the data and makes it easier to work with the data. The method is based on creating a binary tree that ideally captures the data. The search in this representation is simpler than in an unsorted data.

FLANN (Fast Library for Approximate Nearest Neighbors) [16] is used to find the clusters using KDtree. FLANN is a library implementing the nearest neighbor approximation algorithm in a multidimensional space. It also allows to find the optimal cluster with the help of a projection into the floor plane of the object.

IV. Proposed pipeline

Many methods process point clouds, others follow the 3D object reconstruction route. As can be deducted from related work, all of these approaches require significant computational power. Therefore, in this section, I will propose a pipeline implementing obstacle detection given that it is desired to implement this task on a computationally constrained device.

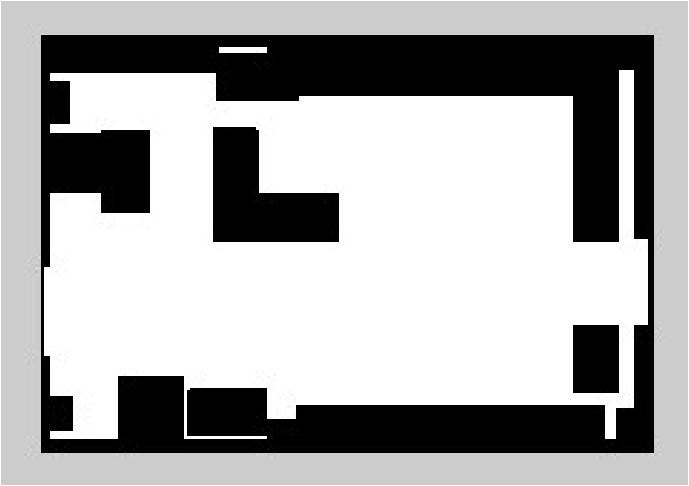


Fig. 2: A map example. This example of lab map was created with graphic editor.

The point cloud data will be first preprocessed. For pre-processing, methods for outlier removal, data reduction and floor removal will be used. Voxel Grid will be used for the first two tasks. This method is simple enough to use on the entire point cloud and brings a significant speedup in the entire post-processing. The RANSAC method will be used to find the floor. This method is slightly computationally intensive, so we could consider a fixed floor separation, but we could lose information about, for example, a hole in the surface. Subsequently, the data will be filtered based on the occupancy map to keep the new observations up-to-date. The last point of the pipeline will be a cluster search that will group points into blocks. It will also allow to find out the ideal rotation of the blocks based on observations, so that they do not occupy space that is not observed.

The map can be created in advance, from knowledge of the space, or it can be created in progress from sensing the environment using the ROS package *navigation2*. The Fig 2 depicts a map of a sensed lab. For the map generation, the map created from the occupancy cells will be used. Occupancy grid is *navigation2* package format for a map. The map must have a known transformation to the sensor and robot reference frame. The robot can modify the map as it moves by incorporating new obstacles.

V. Experiment

In this work, I have performed an experimental implementation of the introduced pipeline. For the experimental setup, I chose ROS2 foxy distribution [17], because it is currently supported version of ROS. It was installed on the Raspberry Ri with Ubuntu 18.04 OS (Operating System). The whole pipeline was attached to a robot that could be



Fig. 3: The robot overview. LiDAR MID-70 is in front. Raspberry Pi is hidden inside. Battery and WiFi router are on the side. The robot was equipped with more facilities than is described in this paper.

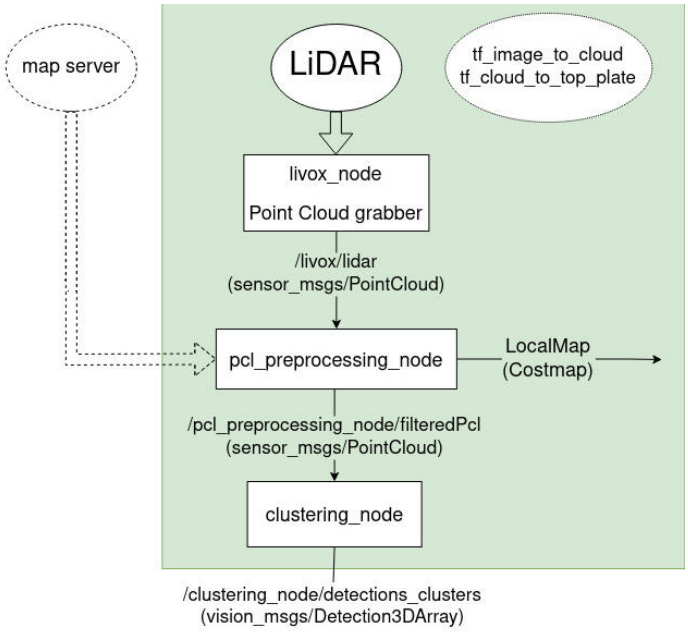


Fig. 4: Diagram with pipeline structure. The whole pipeline is computed on a computationally constrained device. In this diagram one can see messages used for communication between nodes.

controlled remotely with a temporary power source. The goal was to achieve real-time detection of obstacles in the robot's surroundings as it moves, while recognizing objects located in front of the robot for timely stopping in case of a collision within the robot's trajectory. The robot is depicted on Fig 3.

A. Hardware

Pipeline consisted of a MID-70 sensor from Livox and a Raspberry Pi computing unit with 4GB RAM. The MID-70 is LiDAR sensor, which consists of a laser beam source and a spreading mirror that produces a random scan of the area. The sensor senses a cone with a 70 degree FOV (Field Of View) and is capable of sensing from 5cm to tens of meters depending on the tilt of the installation and the reflectivity of the material being sensed. The Raspberry Pi is a credit card sized computing unit capable of running an OS with connectors for HDMI (High-Definition Multimedia Interface), USB (Universal Serial Bus), Ethernet, CSI (Camera Serial Interface), DSI (Display Serial Interface) and GPIO (General-Purpose Input/Output).

Furthermore, the robot was constructed as a two-wheeled machine with a differential chassis. A battery was used as the power source. A wifi-router was added on the robot for communication with the control. Robot was also equipped with more facilities, but they are not important for this work.

B. Dependencies

Following ROS2 packages were used. Package *navigation2* for navigation and working with a map. Package *tf2* for transformation from one frame to another and *rviz2* for message vizualization. You can see the whole pipeline diagram in Fig 4.

C. Software

Each node was created as a separate node in the ROS2 environment. This concept allows to implement differently demanding operations on different HW, or it allows to run a pipeline with multiple different settings. To implement the clustering and data preprocessing methods I have used the PCL (Point Cloud Library) [18]. The PCL is open-source library specialized to work with point cloud.

The *livox_ros2_driver* was used as the first node that collects the sensor data. This software is created directly by Livox as a support for developers. It allows to collect data from multiple sensors at the same time. It puts the data into ROS messages that have a clear indication of which frame of reference they come from and what their timestamp is. The node calculates the timestamp from the timestamp differences of the incoming packets. The visualized raw data are depicted in Fig 5a and Fig 5d.

The next node in the sequence is *pcl_preprocessing_node*. This node is responsible for appropriately processing the data so that clustering can be performed afterwards. The node can detect and remove the floor by finding the floor equation. Subsequently, it allows to reduce the amount of points in the point cloud by using a Voxel Grid. It also allows points to be transformed, with *tf2* package, from one reference frame to another and filtered according to a map (in which known obstacles are entered). The point clouds are filtred with map

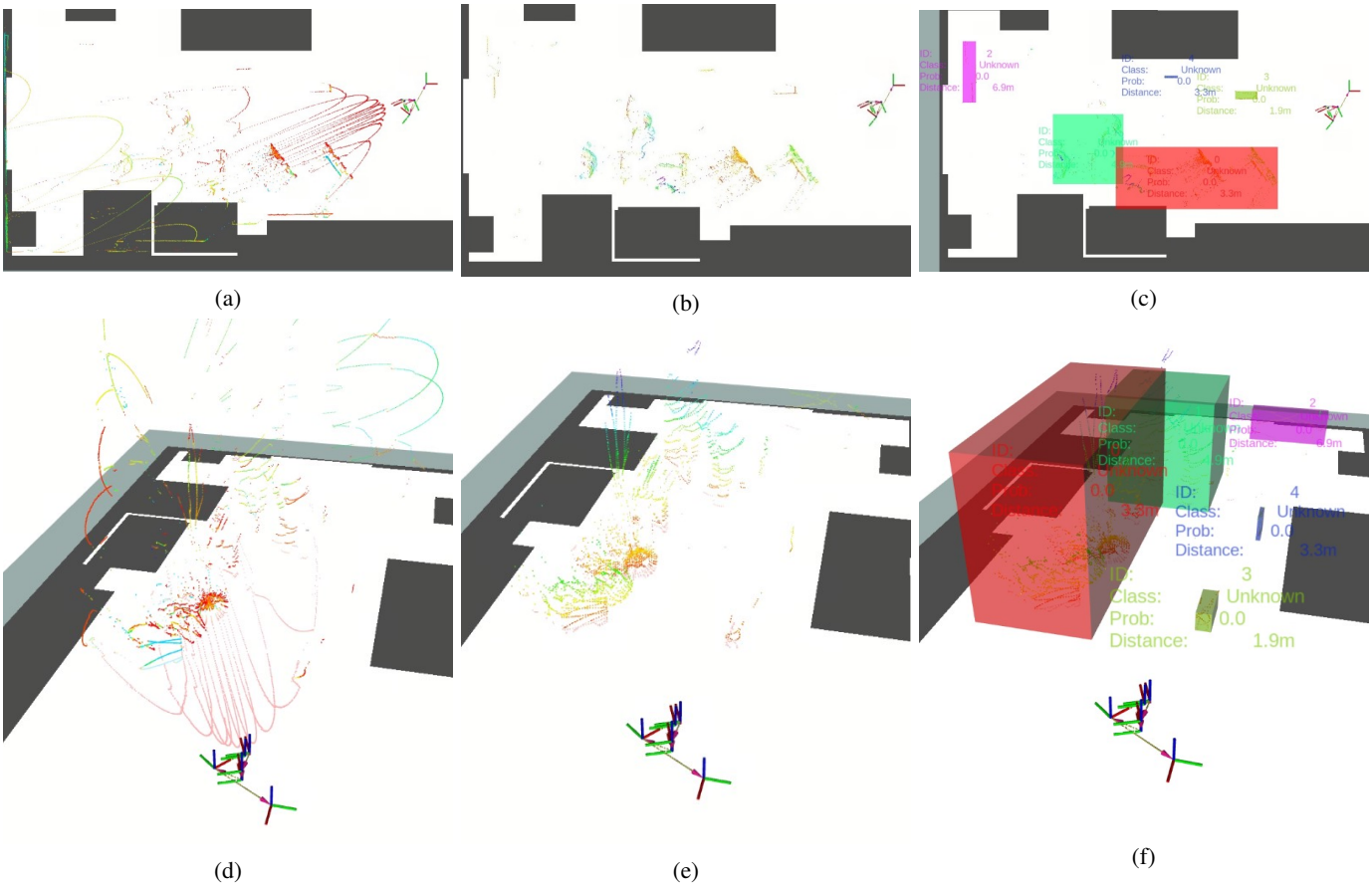


Fig. 5: Example of processed point cloud from two perspectives. *(a)* Raw data from sensor (bird's-eye view), *(b)* preprocessed data with *pcl_preprocessing_node* (bird's-eye view), *(c)* detections found by *cloustering_node* (bird's-eye view), *(d)* Raw data from sensor (perspective view), *(e)* preprocessed data with *pcl_preprocessing_node* (perspective view), *(f)* detections found by *cloustering_node* (perspective view).

by comparison of each cloud point with corresponding point in map. If the point in map is occupied, the point is removed from point cloud. The node also allows the creation of a local map, where the sensed part (of the environment only) is included. The input to this node is ROS messages from the previous node. The visualized preprocessed data are depicted in Fig 5b and Fig 5e.

The last node in the pipeline is *clustering_node*. This node allows to create objects from clusters in a point cloud by use of FLANN. The object is rotated around the center point and the content of the projected rectangle is correspondingly recomputed. When the minimum content is found, the remaining points are recalculated. The input to this node is preprocessed data and the output is ROS messages containing detections. The visualized detections are depicted in Fig 5c and Fig 5f.

I have added a visualization node to the entire pipeline for rendering using the *rviz2* package. This node allows to convert non-visualizable (DetectionArray) messages into visualizable (MarkerArray) ones. One node that stands aside from our pipeline but is necessary for pipeline implementation is *map_server*. It is responsible for working with the map. Receive changes from other nodes and provide the map on request. This node can create a map from a predefined one, or based on messages from the pipeline. If the obstacles disappear it is necessary to reboot this node.

An important aspect of the implementation was the creation of configuration files for easy parameter modification and the creation of startup scripts. Each node has its own parameters that are editable in a single configuration file. The startup scripts allow the configuration files to be loaded and it also allows individual nodes to be started.

D. Results

Several experiments were performed in the laboratory of Brno University of Technology. The aim of the experiment was to move the robot around the lab, detect obstacles and capture data for evaluation. A sensor was attached to the robot along with a computing unit, power source and controls. The following scenarios were prepared: sudden obstacle detection, static obstacle and map error.

The sensing frequency of the LiDAR was 10 Hz. A total of 17,000 point cloud samples were measured during 3 measurements. These samples were qualitatively evaluated, e. g.

TABLE I: Comparison of time consumption for two selected model and my pipeline. Note that both 3D object detectors were processed on GPU (Graphical Processing Unit), but pipeline was computed on CPU (Central Processing Unit)

Method	Time[ms]
CBGS [6] (on GPU)	40
Second [3] (on GPU)	171
Pipeline (My on CPU)	11

depicted in Fig 5. In Tab I is presented time consumption for two selected 3D object detectors and my pipeline. The pipeline allows to find obstacles in point cloud.

VI. Conclusion

The aim of this paper was to familiarize the reader with the possibilities of finding obstacles in space using a LiDAR sensor. These sensors are usually used in robotics to sense the surrounding environment. The methods that can be used are discussed. E. g. 3D object detectors could be use for obstacle searching, If we have enought computational power. On the other hand I want to search for obstacles on a computationally constrained device a new pipeline has been proposed. The pipeline consists of data preprocessing and clustering. The proposed pipeline was qualitatively verified on measured data. In the experiment section, the implementation of such a method was described along with the achieved results. The pipeline was tested on a computationally constrained device and also compared with 3D object detectors. The pipeline is suitable for obstacle search in real-time.

Follow-up work will focus on extending the pipeline to include obstacle classification. The development will look at different responses to different types of obstacles. If a robot passes a moving human it should react differently than if it is moving around a shelf.

References

[1] H. Kutluca. (2020) Robot operating system 2 (ros 2) architecture, software architecture foundations, medium. Accessed: 2020-03-12. [Online]. Available: https://medium.com/software-architecture-foundations/robot-operating-system-2-ros-2-architecture-731ef1867776

[2] Y. Zhou and O. Tuzel, "Voxelnet: End-to-end learning for point cloud based 3d object detection," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4490–4499.

[3] Y. Yan, Y. Mao, and B. Li, "Second: Sparsely embedded convolutional detection," *Sensors*, vol. 18, no. 10, 2018. [Online]. Available: https://www.mdpi.com/1424-8220/18/10/3337

[4] A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, and O. Beijbom, "Pointpillars: Fast encoders for object detection from point clouds," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 12 689–12 697.

[5] W. Shi and R. Rajkumar, "Point-gnn: Graph neural network for 3d object detection in a point cloud," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 1708–1716.

[6] B. Zhu, Z. Jiang, X. Zhou, Z. Li, and G. Yu, "Class-balanced grouping and sampling for point cloud 3d object detection," *arXiv preprint arXiv:1908.09492*, 2019.

[7] S. Shi, Z. Wang, J. Shi, X. Wang, and H. Li, "From points to parts: 3d object detection from point cloud with part-aware and part-aggregation network," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 8, pp. 2647–2664, 2021.

[8] S. Shi, X. Wang, and H. Li, "Pointrcnn: 3d object proposal generation and detection from point cloud," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 770–779.

[9] Z. Yang, Y. Sun, S. Liu, X. Shen, and J. Jia, "Std: Sparse-to-dense 3d object detector for point cloud," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 1951–1960.

[10] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, ser. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, p. 5105–5114.

[11] A. Mousavian, D. Anguelov, J. Flynn, and J. Košecká, "3d bounding box estimation using deep learning and geometry," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 5632–5640.

[12] B. Li, W. Ouyang, L. Sheng, X. Zeng, and X. Wang, "Gs3d: An efficient 3d object detection framework for autonomous driving," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 1019–1028.

[13] J. Ku, A. D. Pon, and S. L. Waslander, "Monocular 3d object detection leveraging accurate proposals and shape reconstruction," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 11 859–11 868.

[14] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, p. 381–395, jun 1981. [Online]. Available: https://doi.org/10.1145/358669.358692

[15] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Commun. ACM*, vol. 18, no. 9, p. 509–517, sep 1975. [Online]. Available: https://doi.org/10.1145/361002.361007

[16] M. Muja and D. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration." in *VISAPP 2009 - Proceedings of the 4th International Conference on Computer Vision Theory and Applications*, vol. 1, 01 2009, pp. 331–340.

[17] S. Macenski, T. Foote, B. Gerkey, C. Lalancette, and W. Woodall, "Robot operating system 2: Design, architecture, and uses in the wild," *Science Robotics*, vol. 7, no. 66, p. eabm6074, 2022. [Online]. Available: https://www.science.org/doi/abs/10.1126/scirobotics.abm6074

[18] R. B. Rusu and S. Cousins, "3D is here: Point Cloud Library (PCL)," in *IEEE International Conference on Robotics and Automation (ICRA)*. Shanghai, China: IEEE, May 9-13 2011.

Autonomously Controlled Car Track

Michal Pomajba
grammar school
Purkyňovo gymnázium, Strážnice, Masarykova 379, příspěvková organizace
Sudoměřice, Czech Republic
michalpomajba@seznam.cz

***Abstract*—This paper, Autonomously Controlled Car Track, discusses the behavior of a model car on a car track with a focus on speed control. The aim of the work is to develop a system for controlling the speed of the model car depending on the profile of the track. The practical part measures the behavior of the car model, the principle of the program structure and the development of the model. The work is based on the assumption that when changing direction, the car must reduce its speed to be able to traverse the path. The practical part tries to prove this assumption. This is to be achieved by controlling the power input of the motor. The result of the work should be an application that allows the model to traverse the track in the shortest possible time.**

Keywords—model car, PWM, car track

I. Introduction

Autonomous driving is a much discussed topic nowadays. In everyday life we are surrounded by different types of autonomous systems, and it is the creation of one simple autonomous system that I focus on in this paper. The aim of this thesis is to develop Hardware and Software for a model car on a car track. This application allows the model to drive through any set up car track in the shortest time possible. For each area of the racetrack, the maximum speed of the model car is limited in some way. Whether it is the maximum power of the engine that powers the model, the maximum speed in a curve, or the need to reduce speed well in advance of a curve due to the inertia of the model. For this reason, it is necessary to determine the maximum motor input for each section of the track. This theoretically achieve the lowest possible lap time for a given car track and car model. The basic principle of controlling the car model is to divide the microcontroller program into a measurement sequence and an operation sequence. In the measuring sequence, the microcontroller receives the required data and based on this data it controls the motor power in the operating sequence.

II. Hardware

The heart of the application is the ATmega328P microcontroller connected to the Arduino Pro Mini development board [1]. The peripherals of this microcontroller include a gyroscope with MPU-6050 chip [2] and Bluetooth module HM-10 [3]. The car model is driven by a DC motor. The input power to this motor is controlled by TIP31A switching transistor. There are more options to control the motor power. In this project I use PWM output to switch the transistor [4]. The value to set the PWM output ranges from 0 to 255. 0 is equivalent to 0% and 255 is equivalent to 100% of motor power. There is a single pole power supply line on each side of the car track rail. The model car is connected to this line by two metal brushes. To safely connect the model car in the reverse direction of the car track, I had to add a 2W10 diode bridge to the circuit. The car track's transformer generates a voltage of 14.8 V. This voltage is too high to power the voltage regulator of the Arduino Pro Mini board, so I reduce the voltage to 9 V with the L78S09CV voltage regulator [5]. The whole application connected on the model can be seen in the Fig. 1. In the Fig. 2 you can see the circuit diagram of the whole application. The car track I use is analog type (for which is speed of model car controlled by changing power supply voltage) and its base is one type of straight lines and one type of curved parts with 60 degrees angle. There are always two rails on each part, so the curves will need to be differentiated into inside and outside.



Fig. 1: A model car of an autonomously driven car track. The model contains all components necessary for autonomous operation. Photo by the author.

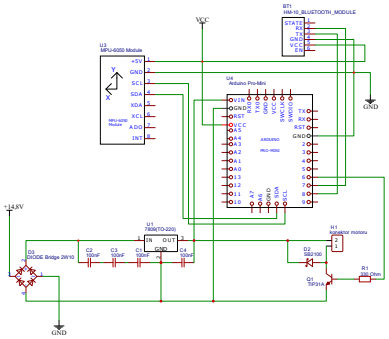


Fig. 2: Overall circuit diagram. Created in EasyEDA.

III. MEASURMENT

A. Measuring sequence speed

The task of the first part of the measurement is to determine the appropriate speed of the measurement sequence of the model. I construct a car track that contained various lengths of curves. The model car drives around the track and sends gyro data via Bluetooth wireless communication. I then process this data and create plots of the dependence of the gyroscope value on time. When the speed is too low, the gyro value is too low in the curves and when going through a long curve, the value decreases rapidly. Too high speed causes the model car to skid in the curves and this negatively affects the gyro data and furthermore threatens to drop the model car off the track. Based on the gyroscope data and observations, I select a speed at PWM value 65 (it equates to roughly 25% of the maximum motor power) as the ideal speed for the measurement sequence. At this speed, the rise and fall of the gyroscope readings at the beginning and end of the curve is most discernible and does not threaten to drop the model car off the track.

B. Measuring techniques

The first thing to understand is how the car model can detect a curve. If the value of the raw Z-axis gyro data exceeds 1.5, the car model goes through the curve. If the value does not exceed this threshold, the car model travels in a straight line. I use the timer to measure the length of the section with the higher and lower value. Due to increased friction and vibration in the corners, there are fluctuations in the gyroscopic values. This phenomenon needs to be treated. The treatment is as follows. If a decrease in value below gives threshold occurred, the microcontroller writes the timer value to a variable and reset the timer. If the value is increased again within 0.2 seconds, the microcontroller recognizes that the car model is still going through the same curve and adds the new timer value to the original timer value when the gyro value is decreased again. A time of 0.2 seconds is not long enough to get the model across one part of the straight line into the next curve. Therefore, if the time is not greater than 0.2 s and the gyro value has again exceeded the 1.5 limit, it is certain that the model is staying in the curve. This process can be repeated several times in succession until the gyro value remains low for at least 0.2 seconds. The last time the value is lowered, the timer is added for the last time and the sum of all timer values during one curve is stored in a variable by the microcontroller and considered as the curve time.

In simple terms, the way this system works is that if the microcontroller records two or more curves during one curve based on the gyro fluctuation, it adds the times of those curves, including the short recorded straight lines in between. The whole principle of the measurement program is shown in the block diagram in Fig. 3.

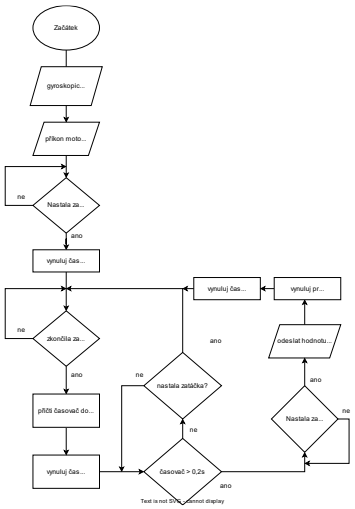


Fig. 3: Block diagram of the program for measuring straight line's and curve's times at a measuring speed PWM 65. Created in diagrams.net.

C. Measuring sequence

I measure the curve and straightaway times at PWM 65. I always let the model car go through 10 curves to get the engine up to speed and stabilize. I average the times of the next 20 curves and straight lines and enter them in Tab. 1 and Tab. 2. The tables already contain the data from the other measurements in Sections D and E. Tab. 1 and Tab. 2 are divided into a measurement sequence and an operational sequence. In the Tab. 1 you can see the times and speeds of the curves. These curves are divided by the number of parts that make up the curve. Next, the curves are divided by rail type into inside and outside. In this measurement, there are always two straight blocks before the curve. The Tab. 2 shows the times and speeds of the straight lines, which are divided according to the length and type of rail preceding the curve. Each straight line is always preceded by a three-part curve as you can see in Fig. 4. Based on the measured times and values, I then determine the number of curves of the autocross track and the high-speed times on the straight lines.

Tab. 1: Curve's times for different type curves at measuring or operating speed.

curve type	rail	measurement sequence		operational sequence	
		speed PWM	average time t [s]	curve speed PWM	average time t [s]
one-piece	outside	65	0,294	71	0,251
	inside	65	0,248	71	0,197
two-piece	outside	65	0,624	71	0,515
	inside	65	0,485	71	0,439
three-piece	outside	65	0,981	71	0,816
	inside	65	0,773	71	0,698
four-piece	outside	65	1,379	71	1,15
	inside	65	1,253	71	1,102

Tab. 2: Times of the different types of straight lines in a measuring or operating sequence.

straight line type	previous curve	measurement sequence		operational sequence			
		speed PWM	average time t [s]	high speed PWM	time of high speed t [s]	brake speed PWM	average time t [s]
One-piece	outside	65	0,355	255	0,08	15	0,308
	inside	65	0,313	255	0,09	15	0,283
two-piece	outside	65	0,621	255	0,15	15	0,487
	inside	65	0,712	255	0,16	15	0,364
three-piece	outside	65	0,878	255	0,21	15	0,71
	inside	65	0,896	255	0,23	15	0,608
Four-piece	outside	65	1,098	255	0,28	15	0,848
	inside	65	1,145	255	0,31	15	0,788



Fig. 4: Demonstration of the construction of the car track during a part of the measurement. Photo by the author.

D. Curves speed

In the next measurement, I determine the maximum speed at which the model car could safely go thought the curve. I always set up an axisymmetric car track in which there is always only one type of curves and straight lines. I gradually increase the speed until the model car is drifting in the curve. I find that the maximum motor input at which the model car can travel around the track without drifting is PWM value 71 (it equates to roughly 28% of the maximum motor power). Based on further testing, I find that this speed is the same for all types of curves.

E. Measurement of the operating sequence

In the last part I focuse on acceleration and deceleration on the straight lines. In order to keep the time for the car model to complete the straight lines as low as possible, I have to increase the motor power to the maximum value at the beginning of the straight line and decrease the motor power to the minimum value for a few moments before the end. After a few measurements, I find that the model car gradually accelerates with this control and only after a few laps does the speed settle down.

If I set a shorter period of high speed, the model car does not reach the corner but stops before it. Conversely, if I set a longer high-speed time, the car model gradually reaches such a high speed that it is unable to stay in the car track. I solve this problem by increasing the braking power to PWM 15 (it equates to roughly 6% of the maximum motor power). With this braking power there is less chance of the model car coming to a complete stop before the curve. This increases the time at which the model car would have a suitable speed for corner entry. The measurement is carried out as follows.

When the microcontroller detects the end of the curve, it sets the motor to maximum power (PWM value 255) and resets the timer. It drives at maximum speed for a very short period of time and then set the motor input power to the braking value of PWM value 15. If the timer exceeds the value of 2 s, it increases the maximum speed time by 0.01 s and sets the motor input power to PWM 71. By doing this, it determines that it fails to come all the way into the curve and increases the maximum speed duration for the next lap. If the microcontroller detects the curve, it increases the motor input to PWM 71 to safely negotiate the curve. The whole principle of the measurement program is shown in the block diagram in Fig. 5. With the model car running like this, I measure the straight line's and curve's times. I always measure 20 straight line's or curve's times. From these times I calculate an average, which I insert into Tab. 1 and Tab. 2.

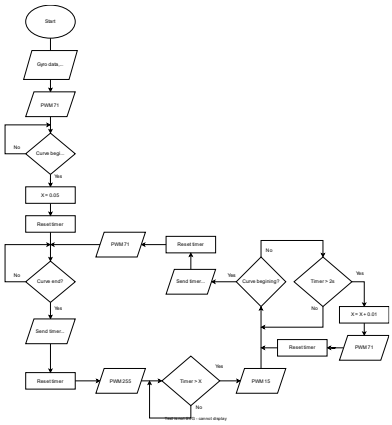


Fig. 5: Block diagram of the program for measuring the appropriate time of maximum speed for different types of straight lines. Created in diagrams.net.

IV. Measuring the Number of Curves

For the autonomously controlled version of the system to work properly, the number of curves per lap of the track must be determined. One possibility is to create a landmark that the system would recognize. The number of curves between passing this point would be equal to the number of curves of one lap of the car track. However, this solution would require an additional microcontroller peripheral.

My simple algorithm for measuring the number of curves consists of comparing the measured curves lengths by sequences of conditions (see in demonstration of the code). Due to external factors, the times of the same curves are not always the same, and therefore the microcontroller calculates a maximum deviation of 0.17 s when making the comparison. I determine this time based on the average difference between the different curves. The whole program for measuring the curves consists of several parts. Each part of the program is designed for one number of curves on the car track. If I have a given number of curves on a car track, I can assume that the curves start to repeat after that number of curves. If these curves match, the car track has just the given number of curves. If the curves do not match in this program, they match in another part of program that assumes a different number of curves in the car track.

The number of individual sequences of conditions depends on the maximum possible number of curves in the car track. In my case, the maximum number of curves in the car track is equal to seven, so there are always seven conditions for

matching curves in each part of program. This limitation on the number of curves is due to the lower computational power of the microcontroller and the lower number of variables used.

```
int positive_deviation = 0.17;
int negative_deviation = -0.17;
sum = (z1 - z4);
if ((sum > negative_deviation) && (sum < positive_deviation)){
  sum = (z2 - z5);
  if ((sum > negative_deviation) && (sum < positive_deviation)){
    sum = (z3 - z6);
    if ((sum > negative_deviation) && (sum < positive_deviation)){
      sum = (z4 - z7);
      if ((sum > negative_deviation) && (sum < positive_deviation)){
        sum = (z5 - z8);
        if ((sum > negative_deviation) && (sum < positive_deviation)){
          sum =(z6 - z9);
          if ((sum > negative_deviation) && (sum < positive_deviation)){
            sum = (z7 - z10);
            if ((sum > negative_deviation) && (sum < positive_deviation)) {
            bluetooth.println("3 curves");
            }
          }
        }
      }
    }
  }
}
```

Demonstration of the part of the program for recognizing three turns in one lap of a car track. Created in PlatformIO.

V. AUTONOMOUS VERSION

The final version of this project already includes fully autonomous driving. When the model car is placed on the racetrack, it goes through 7 curves for starting and steadying the speed. It then drives through 7-17 curves (depending on the number of curves of one lap of the car track) at the measuring speed (PWM 15). Based on the measured data, it assigns data to each straight line for proper acceleration and braking. After the measuring sequence, the program controlling the car model switches to the operating sequence in which it increases and decreases the motor power according to the assigned data. It compares the newly measured straight line's times with the previously determined straight line's times and adjusts the lengths of the high speed straight lines accordingly. The principle of the measuring sequence program is shown in the block diagram in Fig. 6. The principle of the operating sequence program is shown in the block diagram in Fig. 7. The achieved single lap time should be close to the minimum possible time for this car model with the same setup.

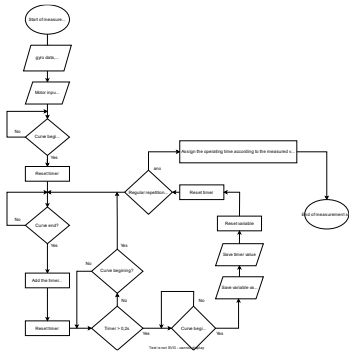


Fig. 6: Block diagram of the measurement sequence of the autonomous program for the car model. Created in diagrams.net.

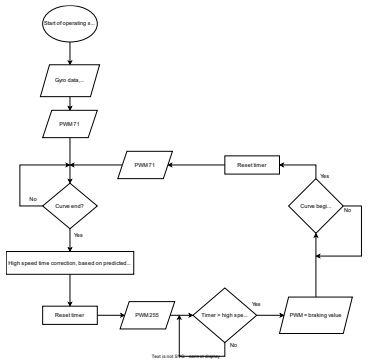


Fig. 7: Block diagram of the autonomous program operation sequence for a car model. Created in diagrams.net.

RESULTS

As part of my paper, I achieve an autonomously controlled model car on a car track. When the model is put on the car track, it runs several laps to measure the data from the car track and measure the number of curves per lap. Based on this data, it determines the appropriate motor power input for each section of the car track to minimize the single lap time through the car track. The program then switches to an operating sequence in which it will control the motor based on the assigned power inputs for each section. This will achieve the lowest possible lap time for a given grip, power and weight of the car model. The program is capable of refining the time of the high-speed during operation. If there is a change in the operating conditions of the car track, (e.g. change in surface grip), the program is able to react to the change and refine the high-speed time.

This project could be used as a simple example of autonomous driving, or even as an autonomous counterpart when racing model cars on a racetrack. In following work I would like to compare the algorithm and regular manual model car driving.

ACKNOWLEDGMENT

I would like to thank my advisor, doc. Ing. Tomáš Frýza Ph.D. for his valuable advice, substantive comments and great patience. This work was supported by JMK.

REFERENCES

[1] Arduino Pro Mini, 2022. In: Arduino [online]. 15. 12. 2022 [cit. 2023-02-10]. Available from: https://docs.arduino.cc/retired/boards/arduino-pro-mini

[2] Arduino gyroskop + akcelerometr. Dratek.cz [online]. [cit. 2023-02-10]. Available from: https://dratek.cz/docs/produkty/0/134/1500635992.pdf

[3] Bluetooth modul HC-10 4.0 BLE klon. Dratek.cz [online]. 2016 [cit. 2023-02-10]. Available from: https://dratek.cz/docs/produkty/0/567/1464644818.pdf

[4] Basics of PWM (Pulse Width Modulation). Arduino Documentation [online]. [cit. 2023-02-10]. Available from: https://docs.arduino.cc/learn/microcontrollers/analog-output

[5] 2 A positive voltage regulators [online]. [cit. 2023-02-10]. Available from: https://www.vpcentrum.eu/index.php?route=product/file&product_id=88515&file_id=2306032

Bipedal Walking Robot

1st Sebastian Matoušek
Gymnázium Matyáše Lercha
Brno, Czech Republic
seb.matousek@gmail.com

***Abstract*—This paper presents the design, manufacturing and control of a walking bipedal robot capable of independent movement and stabilization. The mechanical construction for the robot was 3D modeled and printed out, and custom electronic circuits were made and hand soldered. As a part of this project a computer application was developed for controlling and monitoring the robot.**

***Index Terms*—walking robot, bipedal robot, intelligent servomotor**

I. Introduction

Today, in the field of robotics, many robots are developed with the hopes of them once easing or even replacing human work. Since the environment we live in is mostly constructed for people, it can be challenging for systems utilizing wheels or tracks to navigate an environment or carry out a task which would be simple to do for a human, who can move on two legs. Therefore, in some cases, bipedal robots can be much more effective.

This thesis focuses on the creation of such a bipedal robot. After summarizing the goals of this project, the reader will be acquainted with the construction of the robot, next his electronics and the application made for controlling and monitoring the robot will be introduced.

II. Goals of the project

The main focus of the project was to build a walking robot. Most importantly, this robot should be capable of smooth, stable movement similar to human walking.

The construction would feature two human-like legs, and it was supposed to be easily extendable and modifiable.

To make the electronics of the robot reliable, custom PCBs would be designed with all of the components hand soldered.

A computer application should be developed for controlling the robot and monitoring its parameters, allowing for straightforward teaching of all the movements to the robot.

III. Robot Hardware

Before beginning the development of robot electronics, a rough idea of the mechanical construction had to be though out. The hardware of the robot should consist of two legs, carrying the robot, and a torso, which would hold all the electrical parts and function as a platform for further add-ons.

A. Servomotors

The LX-225 Serial Bus Servomotors chosen for this project act as the robots muscles and there are 13 of them in the whole design. Detailed parameters can be found in Table I. These servomotors are controlled not by PWM signals as is standard, but rather communicate using the UART protocol. This enables multiple servomotors to be chained together and be controlled by a single output pin rather than having 13 separate control pins.

The servomotors have a range of functions, for example being capable of reading the position of their shaft, moving at different speeds, reading housing temperature, setting movement limitations and offsets. These functions enable a versatile usage of the servomotors, while all the functions are controlled by UART communication.

B. 3D Model and robot construction

For purposes of later modifications and also the manufacturing of parts, a 3D model of the robot, which can be seen on Fig. 1, was created in Solidworks 3D CAD. Later, most of the parts were 3D printed in PLA material and the robot was built. Using 3D printing resulted in strong but flexible, lightweight parts.

IV. Robot Electronics

Together with the mechanical parts, the electronics of the robot were developed. The final electrical circuit consists of the following parts:

- Raspberry Pi Zero W
- ATmega328P microcontroller
- MPU6050 inertial measuring unit
- 13 LX-225 servomotors
- 2x 103AT-2 thermistors

TABLE I
LX-225 Servomotor parameters

Nominal Voltage	7,4 V
Working Current(Stall)	100 mA (4A)
Torque	25 kg.cm at 7,4 V
Turning angle	0–240deg
Weight	63 g
Dimensions	40 x 20.14 x 51.1 mm
Baud Speed	115200



Fig. 1. Rendered 3D Model of the robot in Solidworks.

A. Raspberry Pi Zero W

This microprocessor was chosen to drive the robot and communicate with the computer application. It is possible to connect to the robot wirelessly over SSH to work with the robot. Tanks to its compact dimensions it could be fitted onto the back of the robots torso. It can also communicate with the servomotors using the UART protocol.

B. Communication with the servomotors

A module was sold together with the servomotors, used for the translation of full-duplex UART communication – where data transferred by two wires, so that both sides can communicate simultaneously – to half-duplex – where data is sent over one wire. However, this module was functioning unreliably, preventing repeated accurate readings from the servomotors.

Therefore, a solution had to be devised. At first, the idea was to simply connect the two wires and filter the communication using software, since the the servomotors and the control unit did not need to communicate simultaneously, so they could just use one wire.

Another problem arose, however, because the Raspberry wasn't capable of picking up all the communication. Therefore, an ATmega328P chip was used to send the communication further.

This brought yet another problem however, as the ATmega328P microcontroller uses push-pull UART transmission, holding the line high at 5 V. The servomotor weren't able to respond while the line was pulled high, so two mosfet transistors were introduced to the circuit to separate the transmit pin (the pin holding the line high) of the microcontroller from the rest of the communication, which worked as a solution.

C. PCB with communication electronics

All these electronic parts described in the previous subsection needed to somehow be included on the robot. For this, a custom printed circuit board was designed, as seen on Fig. 2.

1) ATmega328P: This microcontroller unit was included on the board for communicating with the servomotors. It also serves other purposes, such as reading out measurements from the IMU included on this board, used for static stabilization, or reading out values from the thermistors attached on the accumulators of the robot.

Pins for an external programmer are soldered on the board, to enable the programming of the microcontroller.

2) MPU6050 IMU: For the purposes of stabilization, an inertial measuring unit was included on the board. This IC is capable of gathering information from a gyroscope input, an accelerometer and a compass.

D. PCB with power management

A second PCB was designed to accommodate for the need of power management part in the robots circuit. Both this PCB and the first one have the same outline as the Raspberry Pi microprocessor PCB does, allowing them to be stacked on top of each other using special, double ended, pins as HAT modules.

This board has 4 layers, to better carry all the current going to the servomotors, whose pins are attached on it. However, although each servomotor is rated to up to 4 A of current, the whole robot rarely draws more than 3 A.

The electronic circuits on the PCB include an LDO regulator for powering the Raspberry and ATmega328P circuits, and a charging circuit made for charging the accumulators of the robot.

The full board can be seen on Fig. 3.

1) LDO regulator: The NCP59300 regulator embedded on the board transforms the voltage coming from the accumulators, corresponding to about 8 V, to 5 V used to power the controller electronics. The current going to those was measured to be about 350 mA, while the regulator would suffice to supply up to 3 A of current.

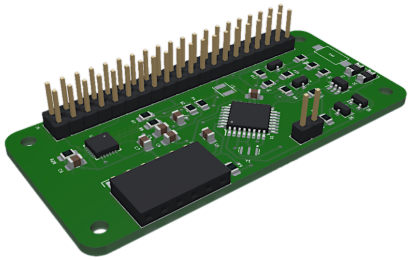


Fig. 2. Rendered image of the PCB with communication electronics on board. The ICs on this board are the ATmega328P microcontroller and the MPU6050 IMU.

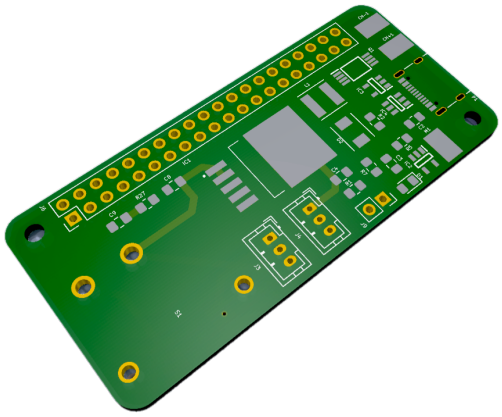


Fig. 3. Rendered image of the power management PCB. A four layer design was used to better conduct all the current to the servomotors.

2) Charging circuit: The charging circuit is divided into a few sections responsible for charging the accumulators and the monitoring of said charging.

The IC HY2120 is the main component of the charging circuit, controlling the charging process using a double-channel power mosfet. An operational amplifier is used to determine whether the accumulators have been charged and controls an LED signaling to the user.

The charging circuit can be powered from the USB-C port soldered on the board, or the robot can be charged wirelessly. The robot is placed on a charging pad and utilizing magnets in the pad and his soles, two coils in his soles align with two coils in the charging pad and the robot starts charging. Due to current restrictions the charging current using wireless charging is comparable with that of the wired charging.

V. Control Application

For a microprocessor or microcontroller to function as needed, it requires a program to be written. The program for the Raspberry Pi was programmed in Python and the program for the ATmega328P chip in the Arduino coding environment, however, the details of the programs will not be discussed here. Rather, the computer application developed for controlling and monitoring the robot will be introduced. The application is shown on Fig. 4.

The control application was programmed in C++ using Qt Creator, and has the following important features:

- It lets the computer communicate with the robot, controlling it
- It has three tabs for switching between servomotor control and limit and offset settings, mainly used to make small changes to the robots position
- An editor for creating position sequences for the robot and their playback
- An integrated application for creating trajectories which the robot should walk on.
- An option to view accumulator and servo housing temperature

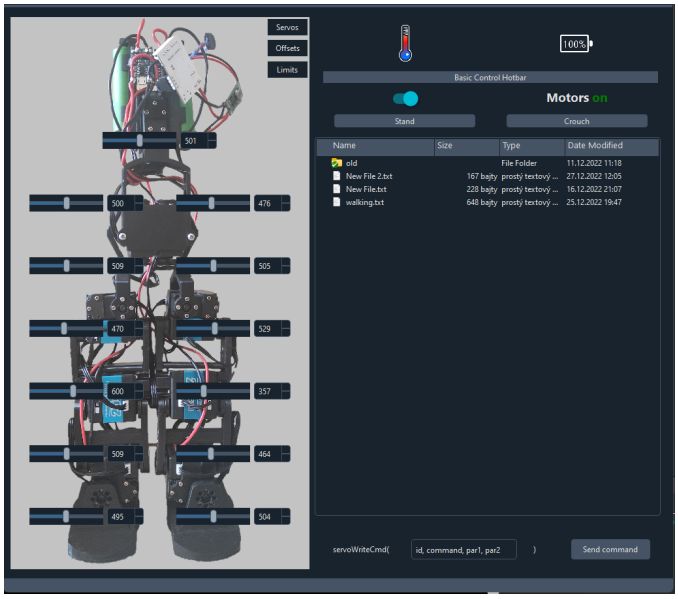


Fig. 4. Computer application developed for easier control of the robot and monitoring.

- A command line to send direct UART commands to specified servomotors

By developing this application, easier control and monitoring of the robot was achieved, allowing to teach the robot to walk effectively, using the functions embedded in the application.

The robot was then taught how to walk by saving all the the position sequences which form a step into a library. The robot now takes these sequences from the library and after adjusting them accordingly, so that he turns by the requested angle, he uses them to move around as specified by the map in the control application.

VI. Results and Testing

The final robot is over 40 cm tall, with a leg span of around 15 cm. Its mass with installed accumulators is around 1400 g.

When the robot was completed, testing of its movement capabilities began. Firstly, it was determined that it can be powered from its accumulators for a little bit under one hour of continuous movement. The movement in it self is also quite reliable, although when asked to turn 90° on the spot, the robot usually turns a few degrees less.

VII. Conclusion

An image of the finalized robot can be viewed on Fig. 5. Two legged and capable of independent movement, the created robot meets all goals that were set for this project. Utilizing CAD software, a 3D model of the robot was made and later used for printing the individual parts of the construction. The custom printed circuit boards ensure reliability of all the robots electronics even in the case of a fall during testing. A computer application developed for the monitoring and controlling of

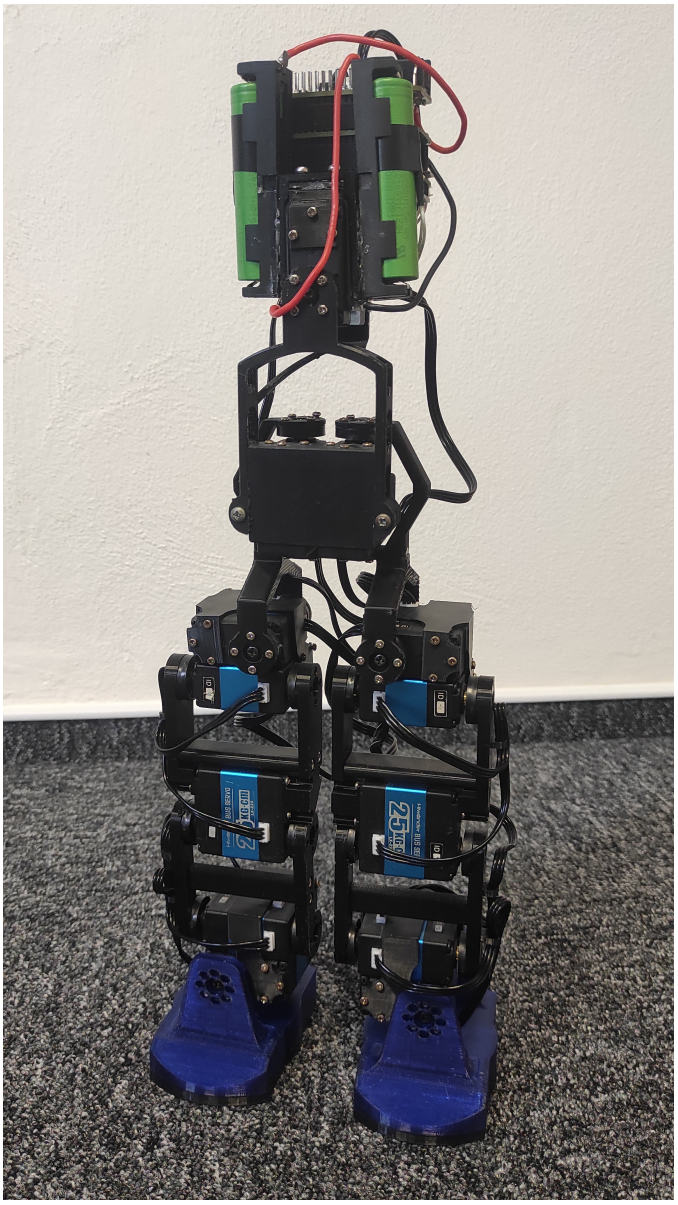


Fig. 5. The completed bipedal robot.

the robot oversees the robots movement and allows the user to determine the actions of the robot.

In the future, arms and a head will be developed and fitted on the legs and torso, making a complete humanoid robot.

VIII. Acknowledgment

I would like to thank Ing. Marek Žák who was my supervisor and consulted me with this project.

Electric field web simulation

https://mariododo.github.io/electric-field-simulator/

Střední škola Teleinformatiky, Ostrava, p.o.

1st Mário Daniel Doskočil
Karviná, Czech Republic
mddoskocil@pm.me

2nd Monika Pustková
Opava, Czech Republic
monika@pustkova.cz

3rd Martin Holánik
Ostrava, Czech Republic
martin.holanik06@gmail.com

***Abstract*—This paper describes a web-based simulation of an electric field primarily designed for interactive lessons in lower education.**

***Index Terms*—Simulation, HTML5, Web, Svg, Javascript, Coulomb's law**

I. Introduction

The goal of this project is to create a teaching tool on the theme of "Electric Field," primarily for physics classes. The tool displays electric charges and their corresponding interactions using field lines, compass needles, or animations, enabling students to visualize and understand the concepts easily. User-selectable parameters, such as the intensity, length, and quality of field lines, or the colors of all rendered elements, allow for customization of the simulation. The simulation result can be exported to an SVG file for further use. The entire simulation is developed using JavaScript, without any external libraries. Moreover, the calculations used for the simulation can also be applied in other related applications.

II. Description

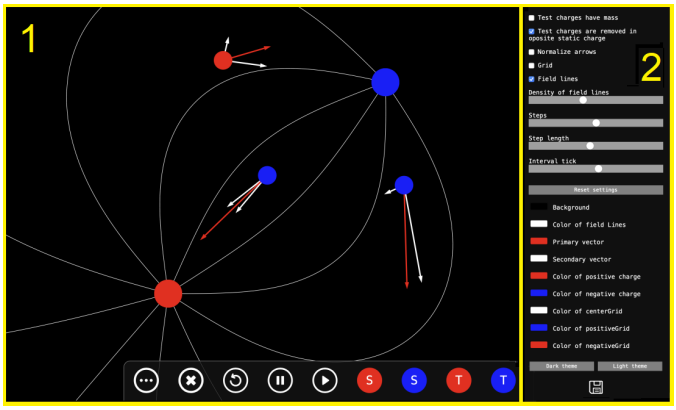


Fig. 1. View of the simulation

A. Basic view

In the basic view, you can see a black background with a bar at the bottom of the screen. On the bar, there are buttons (in order from left): advanced menu, remove all, restart, stop, and play. Additionally, static and test charges of opposite charge are in different colors.

B. Work with website:

1) Addition of charges: To add a charge, drag it from the bar to the black area. Field lines will appear around positive charges, while force vectors will be displayed around test charges in the direction of static charges of opposite charge. The resultant force is displayed as a red vector.

2) Removal of charges: You can remove the charges by dragging them back to the bar. Additionally, charges will be automatically removed once they reach 200 pixels behind the black area in every direction.

3) Action buttons:

- Play - Starts the animation
- Pause - Stops the animation
- Reset - Returns the charges to their previous position
- Remove all - Removes all elements from the black area
- Advance settings - Opens advance settings menu

C. Advanced settings

- Test charges have mass - Enables test charges to act on the field
- Test charges are removed in opposite static charge
- Normalize arrows - Gives the same length to all vectors
- Grid - Toggle the compass needle grid
- Field Lines - Toggle field lines
- Density of fields lines - Amount of the initial field lines points
- Step length - Length of one field line rendering step
- Interval tick - Sets delay for the interval that drives animation
- Reset settings - Resets all settings to default values
- Change the color of background, field lines, vectors, charges and needle grid
- Dark/Light - Sets dark or light theme (for example for projection on a board)
- Save as SVG - Saves simulation to the .svg format

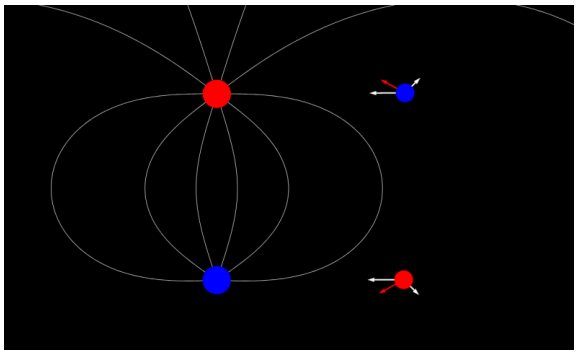


Fig. 2. Default settings view

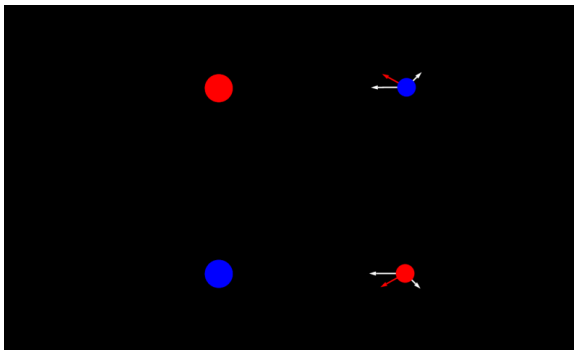


Fig. 3. Field lines and the needle grid disabled

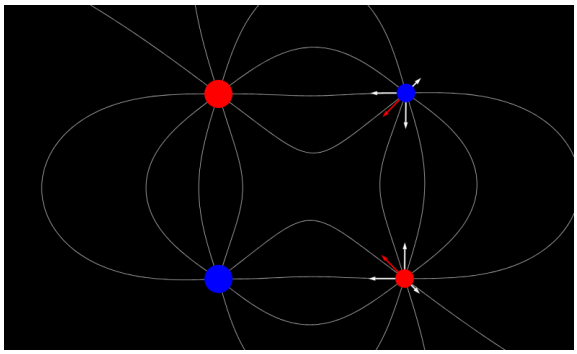


Fig. 4. Test charges have mass

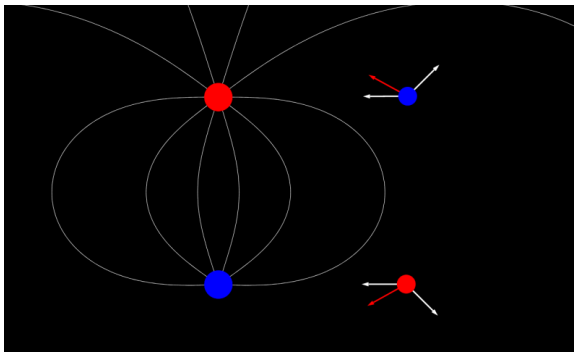


Fig. 5. Normalized vectors

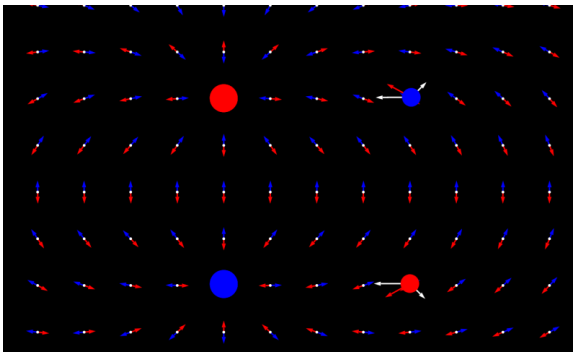


Fig. 6. Compass needle grid without the field lines

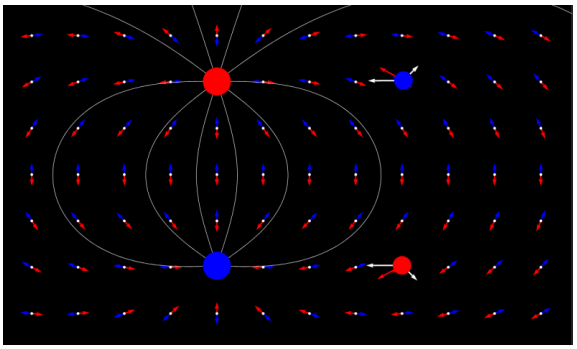


Fig. 7. Compass needle grid with the field lines

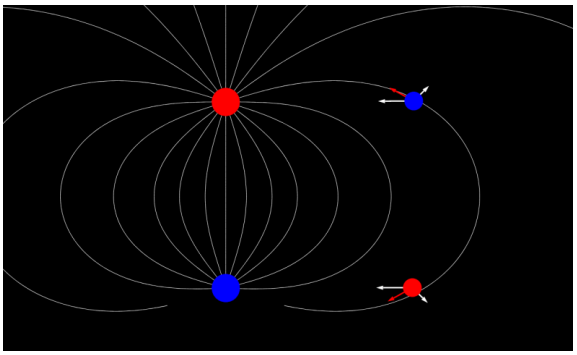


Fig. 8. Step length set to the maximum value

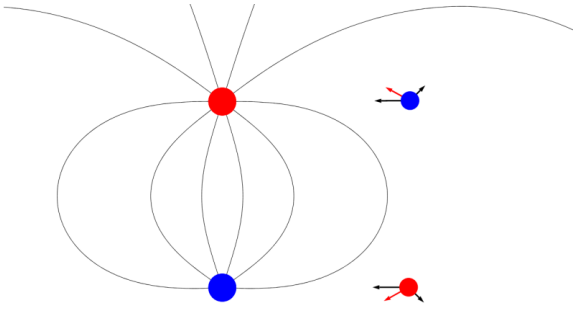


Fig. 9. Light mode

III. Algorithms

Charges Q are objects with the following attributes:

- x - The x coordinate of the charge on the screen.
- y - The y-coordinate of the charge on the screen.
- N - A boolean value that determines the sign of the charge.

Calculations involving angles are performed using radians because the standard JavaScript Math library utilizes them.

A. Calculating force

To calculate the force between two charges, Coulomb's law is used.

$$F = k_e \cdot \frac{|Q_1| \cdot |Q_2|}{r^2}$$

The Coulomb's law is used in a modified form where $k_e = 1$ and $Q_1, Q_2 = 1000C$ which results in:

$$F = \frac{10^6}{r^2}$$

B. Calculating the distance r

To calculate the distance between two charges, we can use the Pythagorean theorem. To calculate the overhangs, we have to take the absolute value of the difference between their respective coordinates.

$$x = |Q_{1x} - Q_{2x}|$$
$$y = |Q_{1y} - Q_{2y}|$$
$$r = \sqrt{x^2 + y^2}$$

C. Resultant force on a charge

The end coordinates of the resultant force U_r can be calculated by summing up all the individual end coordinates U_p of the forces acting on a charge.

$$U_r = \begin{bmatrix} U_{rx} \\ U_{ry} \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{f-1} U_{pix} \\ \sum_{i=0}^{f-1} U_{piy} \end{bmatrix}$$

Where f is the number of forces acting on a specific charge.

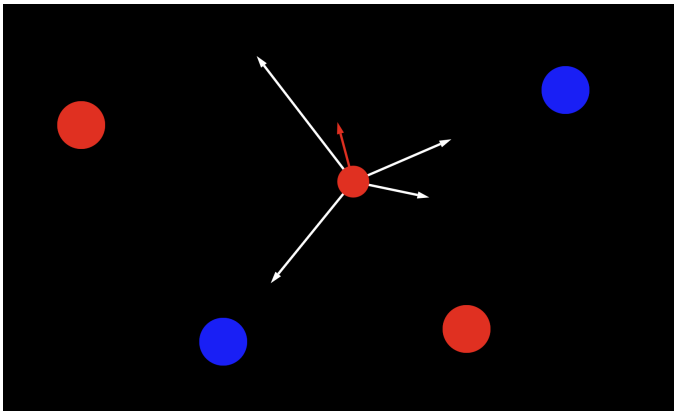


Fig. 10. The white arrows represent the individual forces acting on a charge, while the red arrow represents the resultant force.

D. Generating field lines

The number of field lines n is determined by the user. To start drawing the field lines around a positive charge, we need to distribute the initial points I of the field lines around the charge.

$$\alpha_i = 2\pi - i \cdot \frac{2\pi}{n}$$

$$I_i = \begin{bmatrix} I_{ix} \\ I_{iy} \end{bmatrix} = \begin{bmatrix} Q_{ix} + r \cdot \cos(\alpha_i) \\ Q_{iy} + r \cdot \sin(\alpha_i) \end{bmatrix}$$

Where $i \in \mathbb{N}_0 \wedge 0 \leq i < n$ and r is the radius of the charge. Then we can calculate the force at each initial point and normalize the vector representing the force, so that we obtain the direction of the next step for the field line.

$$u_n = \frac{u}{|u|}$$

But since we are interested in specific coordinates, we can rewrite the formula as follows, where λ is the length of one field line rendering step:

$$L_i = \begin{bmatrix} L_{ix} \\ L_{iy} \end{bmatrix} = \lambda \begin{bmatrix} \frac{I_{ix}}{\sqrt{I_{ix}^2 + I_{iy}^2}} \\ \frac{I_{iy}}{\sqrt{I_{ix}^2 + I_{iy}^2}} \end{bmatrix}$$

We will get a new point L_i, which is the end coordinate of one field line step.

Then, we set I_i to be equal to L_i and repeat the whole process for the user-defined number of steps.

E. Needle grid

For this type of visualization, we have to create a 2D matrix of points.

$$G = \begin{bmatrix} G_{00} & \cdots & G_{0j} \\ \vdots & \ddots & \vdots \\ G_{i0} & \cdots & G_{ij} \end{bmatrix}$$

$$i, j \in \mathbb{N}_0 \wedge 0 \leq i < \frac{\text{screen-height}}{100} \wedge 0 \leq j < \frac{\text{screen-width}}{100}$$

Each point G_{ij} has a default offset of 20 pixels from the screen border. The gaps between the needles are 100 pixels.

$$G_{ij} = \begin{bmatrix} 20 + j \cdot 100 \\ 20 + i \cdot 100 \end{bmatrix}$$

For each point G_{ij}, we have to calculate the force acting on it, normalize that force to get the direction of the needle, and calculate the end coordinate of the needle.

To create the second needle in the opposite direction, we take each end coordinate of the first needle and replace it with its opposite value.

IV. Feedback and use in practice

During the realization of the project, the teachers gave us feedback. Originally, the application was only in dark colors. One of the problems was that it was poorly visible during projection. That's why we added an option to customize the colors of all objects. For quick changes, we also added a light and dark mode button.

We have made the final application available to teachers of electrical engineering and physics, who were happy that we were interested in the topic within their specialization. The final version of the application has already been used in electrical engineering classes.

According to teachers, the application is interesting for students as they do not only see the subject from the point of view of drawn lines on the blackboard, but also observe how the charges dynamically interact with each other.

For future updates we would like to add a calculation of the voltage between charges. The higher the voltage, the brighter the color of the background will be. Additionally, we plan to include the dependence of the voltage on the environment, such as in a vacuum, air, water, etc.

V. Conclusion

We have developed an educational tool that has proven to be practical and effective in enhancing the learning experience of students. This tool is freely accessible to the public, and its source code is also available, making it easily modifiable. Our project has already been utilized in real-world settings and has been evaluated as successful in achieving its objectives.

VI. Acknowledgment

Thanks to our supervisor, Ing. František Knápek, for his patience and time he dedicated to us. Special thanks also belong to our electrical engineering teacher, Ing. Václav Jedlička, for introducing us to this topic and explaining it. We are glad that he uses our tool in his classes and we believe that it helps students to understand the lessons about electric field better. Finally, thanks to our class teacher Mgr. Petra Johnová for introducing us to this project.

References

[1] D. Halliday, R. Resnick, and J. Walker, Fyzika: vysokoškolská učebnice obecné fyziky. V Brně: VUTIUM, 2000.

[2] "JavaScript — Dynamic client-side scripting - Learn web development — MDN," Mar. 22, 2023. https://developer.mozilla.org/en-US/docs/Learn/JavaScript

[3] "SVG: Scalable Vector Graphics — MDN," Mar. 06, 2023. https://developer.mozilla.org/en-US/docs/Web/SVG

Design Methodology for 3D Printing of Spines

Ondřej Buček
Gymnázium Brno-Řečkovice
Brno, Czech Republic
xbuceo@gyrec.cz

doc. Ing. Jan Mikulka, Ph.D.
Ústav teoretické a experimentální elektrotechniky
Vysoké učení technické v Brně
Brno, Czech Republic
mikulka@vutbr.cz

Abstract: ***This work deals with the methodology for three-dimensional printing of the spine model. This is because the model of the patient's spine can be used to give the doctor a better idea of the procedure being performed before and during the surgery. The images of the spine part without surrounding soft tissue taken by computed tomography and magnetic resonance imaging were available for processing. Classifiers were trained on the selected images to segment the images of both modalities, with greater success in segmenting the computed tomography images. The proposed classifier model was then used to segment a comprehensive set of images from which a digital spine model could be constructed. In the last phase of the project, the digital model was printed on a 3D printer.***

Keywords: 3D printing, image segmentation, CT, MR, SVM, random forest

INTRODUCTION

Scoliosis, a pathological deviation of the spine, is the most common spinal deformity in childhood. Only its timely treatment can significantly affect problems in adulthood. Unfortunately, the results of conservative therapy, especially in cases of larger spinal curvatures, are not always sufficient, and surgery is necessary for greater degrees of curvature. Until now, surgeries have been performed by doctors studying an online screen, where the operator transforms a three-dimensional model from computed tomography or magnetic resonance imaging reconstructions according to their instructions. However, this is not very convenient because the doctor must understand the operator, and in addition, not all necessary details may be visible in a simple spine model on the monitor. To better understand the operation and its execution, a printed model of the patient's spine can be used. With it, the doctor can optimally prepare for the operation or subsequently use it in the operating room. A printed model can be taken to the operating room and used for real-time orientation or to discuss surgical procedures with colleagues. The motivation for the project is, therefore, to significantly increase the efficiency of planning (the ability to cut into the model before the operation itself) and the course of the operation (better orientation during the operation).

The aim of the project was to design and verify a methodology for performing three-dimensional printing of spines. Images obtained using computed tomography (CT) and magnetic resonance (MR) were available, showing a portion of the spine without surrounding tissues. The images were taken at the Faculty Hospital Brno Bohunice. The goal was to segment individual images with suitable segmentation methods on the background and track bone tissue in the form of a binary mask, which would be suitable as a digital template for three-dimensional printing of the model of the spine.

I. METHODOLOGY

There were spinal image data available in DICOM format acquired using CT and MR. The aim of the work was to find a method to 3D print the spine for improving diagnosis and simplifying orientation during pediatric spine surgery. As the person is not exposed to ionizing radiation when acquiring spine images using MR, one of the goals was to train a classification model for image segmentation that could recognize the vertebrae in the image, but on spine images without surrounding soft tissue. Of all the images available were chosen 3 CT images and 3 MR images which were already co-registered.

First, the segmentation of CT images was performed, resulting in binary masks from which the borders of the vertebrae were obtained. These were then used to create a training set for the classifier to segment the MR images. Subsequently, this classifier was to be used for segmenting the entire spine or only a few vertebrae from which a computer-generated 3D model would be created. This model would then be converted into the correct format and 3D printed.

II. SEGMENTATION

Segmentation is an operation that divides an image into individual objects. During this process, sections are extracted from the image for use in subsequent image processing. Therefore, segmentation serves as tool that is necessary for further advanced techniques, such as object detection or classification in the image. The result of segmentation is an image in which individual areas of interest are marked. In this work, we will therefore separate individual vertebrae of the spine from the rest of the image. The resulting image will retain its original dimensions, and individual sub-images will not overlap with other segmented sub-images. The segmented parts of the image are usually marked using a specific color. Therefore, the resulting image is a binary mask that aggregates all segmented points belonging to the segmented object.[1] [2]

Trainable segmentation is then a classifier whose input is *Training Features* (tab. 1), which are image data that are computed in a certain way and give the classifier more information to improve the quality of segmentation, unlike, for example simple thresholding, where the only input information is the pixel intensity. This is followed by the so-called model learning, then validation and finally testing, which is the final stage of using the model to classify the unknown pixels into one of two groups (background vs. tissue). [3]

For image processing were used specialist program Fiji, that is mainly used for pictures of medical character. For trainable

segmentations then were used its Plugin called Trainable Weka Segmentation (TWS)(fig 1).

Tab 1: *Training Features*

Training Features	Rovnice
Mean	$\bar{x} = \frac{1}{n} \cdot \left(\sum_{i=1}^{n} x_i \right)$
Gaussian blur	$G_{(x)} = \frac{1}{\sqrt{2\pi_\sigma}} e^{\frac{x^2}{2\sigma^2}}$
Minimum	$\min(\boldsymbol{x})$
Median	$\text{median}(\boldsymbol{x})$
Maximum	$\max(\boldsymbol{x})$
Variance	$r = \frac{1}{n} \Sigma (x_i - \bar{x})^2$

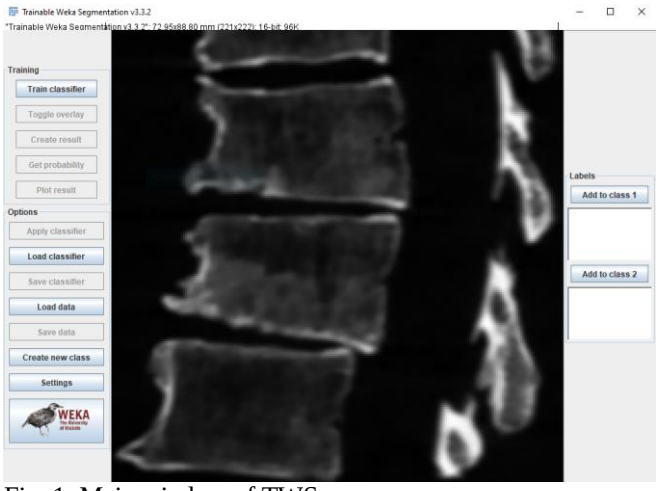


Fig. 1: Main window of TWS

Segmentation of CT pictures

Segmentation was performed on 3 spine CT images acquired in DICOM format. Individual images were already modified and provided by the author of [1]. Firstly, there were manually marked vertebra and background were marked and divided into 2 classes. It was then necessary to select the individual segmentation preferences in the *Settings*. From the *Training Features* were used all types and as classifier type were chosen *FastRandomForest*, which is already automatically preselected when the Plugin is started. The rest of the parameters remained unchanged. Subsequently, the segmentation was started. The whole process was repeated on all 3 images. In its result were 3 binary masks (fig. 2).

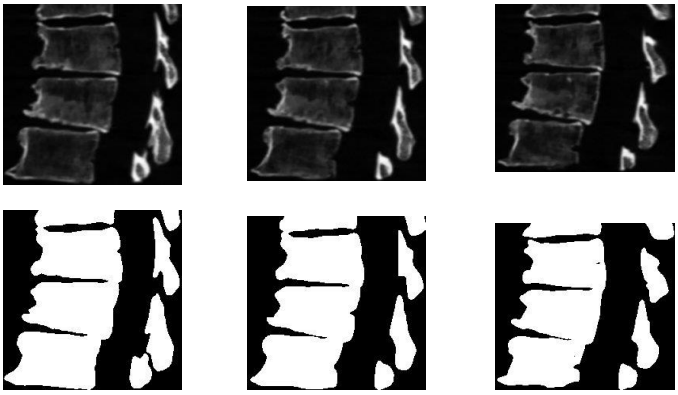


Fig. 2: Three CT images and their created binary masks

Then the accuracy of the segmentation had to be set. This was done by caculationg coeficient Dice.

$$D(\boldsymbol{A}, \boldsymbol{B}) = \frac{2 \cdot N_{TP}}{2 \cdot N_{TP} + N_{FP} + N_{FN}}$$

Therefore, a Ground truth images to all three images was created. That is, how the segmentation should turn out if its accuracy was 100%. This was determined and created in a painting program, where the vertebra was manually marked in white and a binary image was then created by appropriate thresholding. The number of pixels of every region was then counted using Fiji. First, however, the individual pixels had to be assigned to the corresponding regions. This was done using the *Image Calculator*, which is used to perform arithmetic and logical operations between two images. The reagions of pixels were created by the following operations:

- N_{TP} : I_{GT} AND I_{result}
- N_{FP} : (I_{GT} XOR I_{result}) AND I_{GT}
- N_{FN} : (I_{GT} XOR I_{result}) AND I_{result}

Coefficient Dice was then calculated, and it was in average about 0.96. That is very accurate, because value of Dice coefficient is in interval <0, 1>, where 1 is absolute accuracy.

Trainable segmentation on MR pictures

Because the segmentation of CT pictures was successful, the binary mask could be used for next steps. In every of them was created *Selection*, which is process where was every vertebra bordered. Then every selection was saved and used for creating training set for classifier in 3 MR pictures. All MR segmentations were done in TWS.

First segmentation was done on all 3 MR images together. From vertebra selections were used almost every besides the ones which could also mark a big amount of background pixels. That was because the MR images weren't hundred presently co-registered with CT pictures. From *Training Features* were used again every available and as classifier type was used *FastRandomForest*, that was successful for CT images. As result were again 3 binary masks, but from them were clear, that segmentation wasn't successful, and classifier wasn't enough

accurate. That was because vertebra in MR images have similar intensity of pixels as the background. So, for next segmentations there had to be used better classifier type, thus *LibSVM* was chosen. This is a Support Vector Machine classifier. There were done a lot of segmentations by this classifier, in which were used different number of Selections, different combinations of *Training Features* or even were changed more specific details of *LibSVM*, for example kernel type or SVM type, but nothing helped to successfully train a good classifier model that could be used for segmenting big amount of MR images.

Trainable Segmentation of CT picures

Because there was not much time left there had to be plan changed and 3D digital model had to be created from segmented CT images. Firstly, there had to be again classifier trained, so there were used all 3 CT images. Area selection of background and vertebra was created manually as before. As classifier type was chosen *FastRandomForest*, because it was successful before. From *Training Features* was chosen combination of *Mean*, *Variance* and *Median*. After segmenting value of *Dice* was again calculated. Average value of Dice was 0,953, so do classifier model could have been saved and used for segmenting number of CT images from where also came 3 CT images which were used in past. After segmentation of number of images was done, 1 binary mask to every segmented image was created. From all these binary masks was then created digital 3D model of Spine part (picture 3).

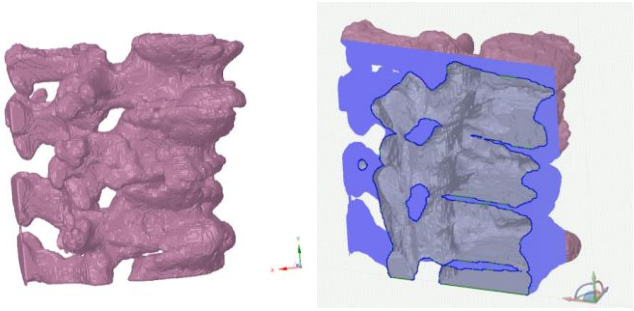


Fig. 3: STL model of Spine part (*Design Spark Mechanical*

III. 3D PRINTING

The FDM printer selected for printing was called *Ultimaker*. The *Ultimaker Cura* program was used to prepare the model and to select individual parameters. The first step was to select the position in which the model was to be printed. Then correct parameters of printing had to be chosen, such as material - ABS, the layer height – 0.15 mm or type of support, speed of printng – 40m/s or layer thickness. Printing process took about 24 hours and was very succesfull, but there are some inaccuracies. The classifier that was used was not perfectly accurate, and the printing material and 3D printer itself were not designed for printing tissue. Additionally, the model only included hard tissue.



Fig. 4: Printed model of Spine part

CONCLUSION

In this work, a methodology for three-dimensional printing of the spine model was proposed and tested. This can greatly simplify the planning and progress of surgery for a patient with scoliosis.

A suitable classifier was successfully trained to segment CT images of the spine without surrounding soft tissues. This was subsequently implemented on a full set of CT images of a portion of the spine, and a three-dimensional model was created from the resulting binary mask. The model was then converted to STL format and after setting the correct print parameters, 3D printing was successfully performed.

REFERENCES

[1] SVOBODA, Jan. *Bakalářská práce - Přesná segmentace obrazových dat*. Brno, Česká republika, nedatováno. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav telekomunikací.

[2] MATULA, Jan. *Diplomová práce - Segmentace chrupavčité tkáně ve 3D mikro CT snímcích myších embryí*. Brno, 2019. VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ.

[3] ImageJ Wiki. *ImageJ Wiki* [online]. [vid. 2022-12-28]. Dostupné z: https://imagej.github.io/



BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
TECHNICKÁ 3058/10, 616 00 BRNO, CZECH REPUBLIC

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

www.eeict.cz

www.fekt.vut.cz